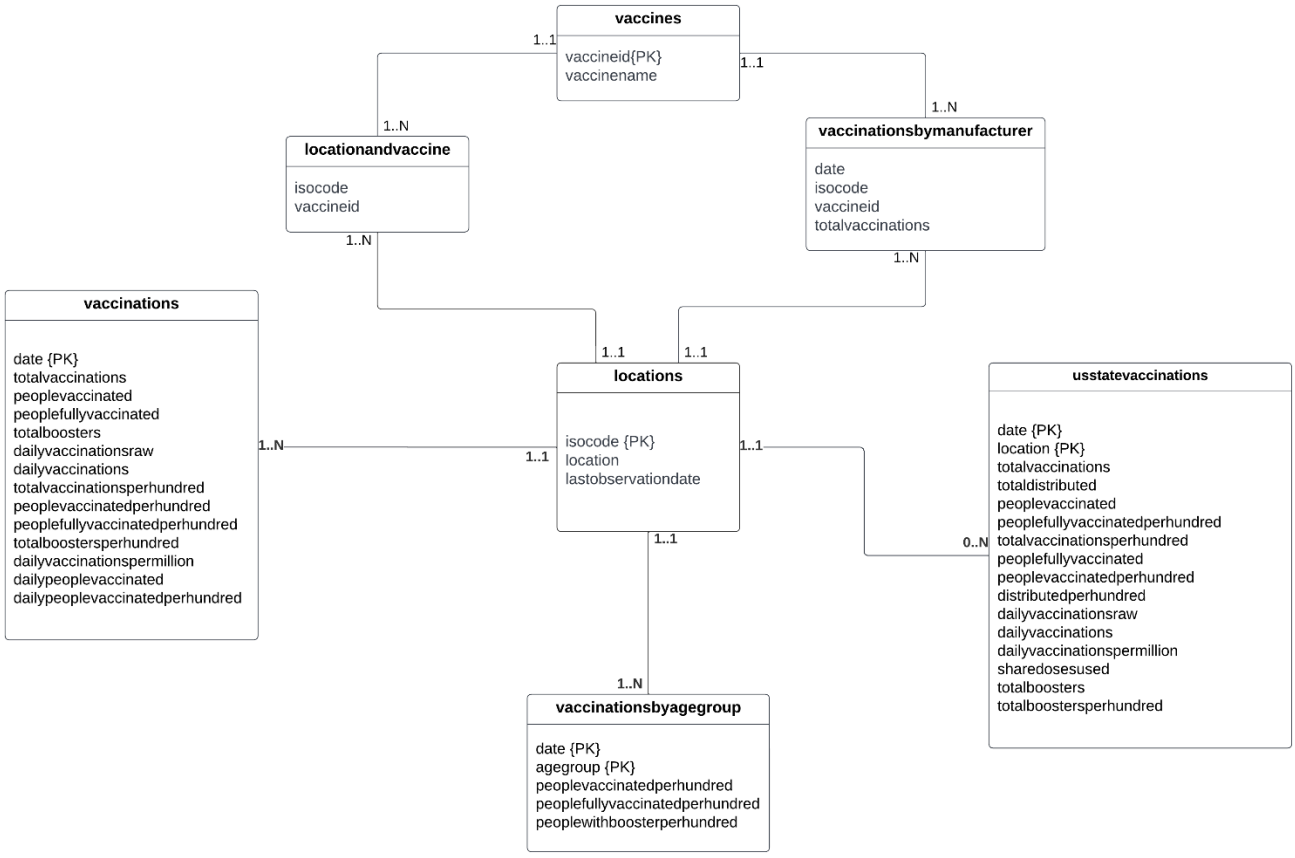
Pranav Karnth Mannur
s3828461

Part B – Model and Design



Assumptions/ Business rules:

1. Each Location can administer many vaccinations and each vaccination can be administered in a location.
2. The location will store 0 to many US state vaccination location.
3. Each location can have one to many vaccinations by age group data and every category in vaccinations by age group will have at most one location.
4. Each location can administer many vaccines and multiple vaccine can be present in one single location.
5. Each Vaccine can have one to many manufacturers.

Data Preprocessing and Normalization Challenges:

1. Connecting vaccinations, vaccinationsbyagegroup, usstatevaccinations, vaccinemanufacturer with location entity using isocode as the foreign key.

2. *Locations table has a Multivalued attribute: Vaccines:*

   We remove this attribute from the Locations table and make it into a new entity named Vaccines with vaccine id as primary key and vaccine name. We use another entity named locationandvaccine which has iso_code from locations and vaccineid from vaccines. **By 7 step mapping process and Normalization.**

   Resulting Changes:
   Locations(isocode, location, lastobservationdate, sourcename, sourceurl)
   Locationandvaccine(isocode*, vaccineid*)
   Vaccines(vaccineid, vaccinename)

Final Schema:

Locations (isocode, location, lastobservationdate, sourcename, sourceurl)

vaccinationsbyagegroup (date, agegroup, isocode*, peoplevaccinatedperhundred, peoplefullyvaccinatedperhundered, peoplewithboosterperhundred)

vaccinations (date, isocode*, totalvaccinations, peoplevaccinated, peoplefullyvaccinated, totalbooster, dailyvaccinationsraw, dailyvaccinations, totalvaccinationsperhundred, peoplevaccinatedperhundred, peoplefullyvaccinatedperhundred, totalboostersperhundred, dailyvaccinationspermillion, dailypeoplevacccinated, dailypeoplevaccinatedperhundred)

vaccines (vaccineid, vaccinename)

locationandvaccine (isocode*, vaccineid*)

vaccinationsbymanufacturer (date, isocode*, vaccineid*, totalvaccinations)

usstatevaccinations (date, location, isocode*, totalvaccinations, totaldistributed, peoplevaccinated, totalvaccinationsperhundred, peoplefullyvaccinatedperhundred, totalvaccinationsperhundred, distributedperhundred, dailyvaccinationsraw, dailyvaccinations, dailyvaccinationspermillion, sharedosesused, totalboosters, totalboostersperhundred)