Diagram, schematic

Description automatically generated**ER Diagram**

**Assumptions:**

1. Not all countries have vaccinations records

**Part B: Designing the Database**

Initial schema:

**vaccinations\_by\_age\_group**(location, date, age\_group, people\_vaccinated\_per\_hundred, people\_fully\_vaccinated\_per\_hundred)

**vaccinations\_by\_manufacturer**(location, date, vaccine, total\_vaccinations)

**vaccinations**(location, iso\_code, 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)

**locations**(location, iso\_code, vaccines, last\_observation\_date, source\_name, source\_website)

**united\_states**(location, total\_vaccinations, date, vaccine, source\_url, people\_vaccinated, people\_fully\_vaccinated, total\_boosters)

**us\_state\_vaccinations**(date, location, 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)

**australia**(date, total\_vaccinations, people\_fully\_vaccinated, people\_vaccinated, vaccine, location, source\_url, total\_boosters)

**israel**(date, total\_vaccinations, people\_vaccinated, people\_fully\_vaccinated, total\_boosters, location, source\_url, vaccine)

**france**(date, total\_vaccinations, people\_vaccinated, people\_fully\_vaccinated, total\_boosters, vaccine, location, source\_url)

Step 1: Map Strong Entities

**vaccinations\_by\_age\_group**(location, date, age\_group, people\_vaccinated\_per\_hundred, people\_fully\_vaccinated\_per\_hundred)

**vaccinations\_by\_manufacturer**(location, date, vaccine, total\_vaccinations)

**vaccinations**(location, iso\_code, 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)

**locations**(location, iso\_code, vaccines, last\_observation\_date, source\_name, source\_website)

**united\_states**(location, total\_vaccinations, date, vaccine, source\_url, people\_vaccinated, people\_fully\_vaccinated, total\_boosters)

**us\_state\_vaccinations**(date, location, 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)

**australia**(date, total\_vaccinations, people\_fully\_vaccinated, people\_vaccinated, vaccine, location, source\_url, total\_boosters)

**israel**(date, total\_vaccinations, people\_vaccinated, people\_fully\_vaccinated, total\_boosters, location, source\_url, vaccine)

**france**(date, total\_vaccinations, people\_vaccinated, people\_fully\_vaccinated, total\_boosters, vaccine, location, source\_url)

Step 2: Map Weak Entities

None

Step 3: Map 1:1 Relationships

None

Step 4: Map 1:N Relationships

**united\_states**(location, total\_vaccinations, date, vaccine, source\_url, people\_vaccinated, people\_fully\_vaccinated, total\_boosters, iso\_code\*)

**us\_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, location\_country\*)

**australia**(date, total\_vaccinations, people\_fully\_vaccinated, people\_vaccinated, vaccine, location, source\_url, total\_boosters, iso\_code\*)

**israel**(date, total\_vaccinations, people\_vaccinated, people\_fully\_vaccinated, total\_boosters, location, source\_url, vaccine, iso\_code\*)

**france**(date, total\_vaccinations, people\_vaccinated, people\_fully\_vaccinated, total\_boosters, vaccine, location, source\_url, iso\_code\*)

Step 5: Map M:N Relationships

**contains\_age\_groups**(location\*, iso\_code\*)

**contains\_manufacturers**(location\*, iso\_code\*)

*Combining into one relation* ->

**contains\_age\_groups\_manufacturers**(location\*, iso\_code\*)

Step 6: Multi-valued Attributes

None

Step 7: Map higher-degree relationships

None

Final Schema:

**vaccinations\_by\_age\_group**(location, date, age\_group, people\_vaccinated\_per\_hundred, people\_fully\_vaccinated\_per\_hundred)

**vaccinations\_by\_manufacturer**(location, date, vaccine, total\_vaccinations)

**vaccinations**(location, iso\_code, 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)

**locations**(location, iso\_code, vaccines, last\_observation\_date, source\_name, source\_website)

**united\_states**(location, total\_vaccinations, date, vaccine, source\_url, people\_vaccinated, people\_fully\_vaccinated, total\_boosters, iso\_code\*)

**us\_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, location\_country\*)

**australia**(date, total\_vaccinations, people\_fully\_vaccinated, people\_vaccinated, vaccine, location, source\_url, total\_boosters, iso\_code\*)

**israel**(date, total\_vaccinations, people\_vaccinated, people\_fully\_vaccinated, total\_boosters, location, source\_url, vaccine, iso\_code\*)

**france**(date, total\_vaccinations, people\_vaccinated, people\_fully\_vaccinated, total\_boosters, vaccine, location, source\_url, iso\_code\*)

**contains\_age\_groups\_manufacturers**(location\*, iso\_code\*)

**Normalisation**

First Normal Form (1NF)

All the relations satisfy the 1NF as every value within each tuple component is atomic.

Second Normal Form (2NF)

All relations except the us\_state\_vaccinations relation only contain one primary key and no composite keys, they are all in 2NF. This is because all non-primary key attributes are entirely dependent on the single primary key of their respective relations.

The us\_state\_vaccinations relation contains a composite key; however, it is also in 2NF. This is because all attributes are functionally dependent on the entire composite key, namely “date” and “location\_state”.

Third Normal Form (3NF)

All relations are in 3NF as no non-primary key attributes are functionally dependent on any non-primary key attributes.

A relation named “contains\_age\_groups\_manufacturers” is created to store each country’s official names and ISO codes so that other relations can refer to this relation for country names or ISO codes if required.