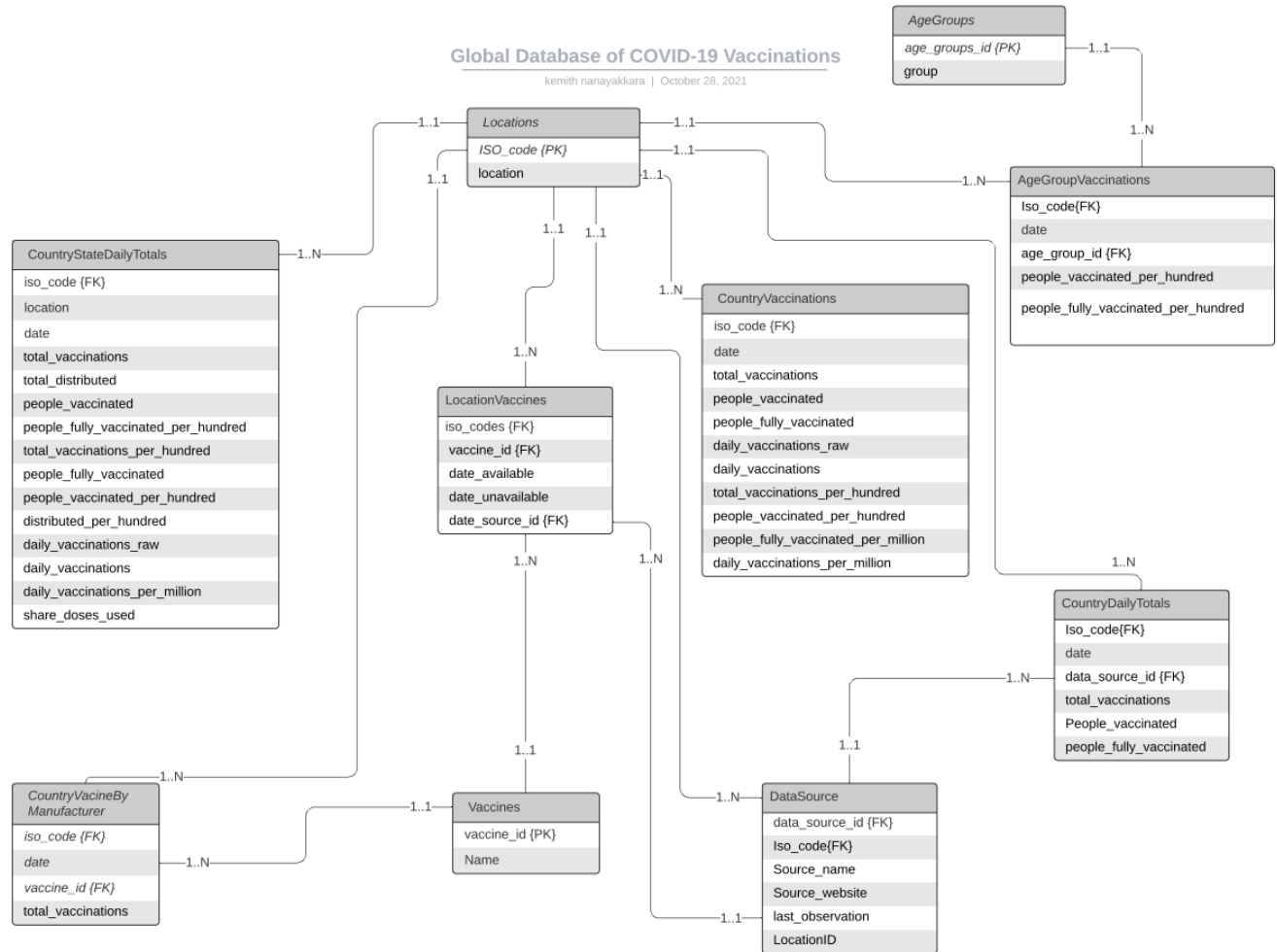


# Part B: Designing the Database

## ER diagram for Global Database of COVID-19 Vaccinations



### Assumptions

- iso\_code is the strongest link in the data provided, csv files not using iso\_code have a location which is a country name and can be converted using a query. EXCEPT: for us\_state\_vaccinations.csv where location is a state; however knowing from the file name its US data when importing into CountryStateDailyTotals we can add the iso\_code column with the value USA.
  - Eg: iso\_code = ISO 3166-1 alpha-3
- based off all data in csv files i will be tracking the date a country started using a vaccine in the LocationVaccines table using date\_available as well as tracking the date inwhich the data stopped showing the vaccine.
- Countries are using the same names for each vaccine.

- Age Groupings are the same across all countries and therefor can be recycled.
- us\_state\_vaccinations.csv could be apapted in the future to track states in other countries so the table for this data should be universal. So I created a table CountryStateDailyTotals to handle this type of data with the additional column iso\_code.