Further cleaning efforts..

Adding the Lat and Lng to the Tableau sheets led to some schoold being mapped to other countries like India.

Upon investigation, these were a few schools from Vermont and 3 schools from Florida.

Looking at the three schools in Florida,

name state city district county lat lng

*<chr>* *<chr>* *<chr>* *<chr>* *<chr>* *<dbl>* *<dbl>*

1 Rise Academy School Of Science And Technology Florida NA Broward NA 26.3 80.2

2 Florida Cyber Charter Academy At Clay Florida NA Clay NA 30.2 30.2

3 Florida Cyber Charter Academy At Osceola Florida NA Osceola NA 30.2 28.3

Given that these schools are from the US, they should have a -ve longitude.

Looking them up on google maps gave the following results. As it looks like the Latitudes were correct but the Longitudes were wrong. In case of rice academy, the latitude was just marked +ve instead of -ve. But in the other two cases, the

Rice academy Florida 26.269098297855223, -80.20490036825754

Florida Cyber Charter Academy At Clay 30.199796517314102, -81.57098807790892

Florida Cyber Charter Academy At Osceola 30.201496427168053, -81.57152127116356

There were 83 schools from Vermont that had incorrect Longitudes. The Latitudes are in the correct range. Vermont Longitudes are supposed to be in the -70s. However, from the following table one can see that the Latitudes have been copied as Longitudes.

| **name** | | **state** | | **city** | **district** | | **county** | | **lat** | **lng** | **mmr** | | **overall** | |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  |  | |  | |  |  | |  | |  |  |  | |  |
| **1** | Addison Central | | Vermont | | *NA* | *NA* | | Addison | | 44.08794 | 44.08794 | 100.00000 | | 100.00000 |
| **2** | Burke Town | | Vermont | | *NA* | *NA* | | Caledonia | | 44.62461 | 44.62461 | 100.00000 | | 100.00000 |
| **3** | Canaan Schools | | Vermont | | *NA* | *NA* | | Essex | | 44.99539 | 44.99539 | 100.00000 | | 98.80240 |

Upon investigating the source data (from the links that were used to get the latitude longitude info on) for the two states the longitude values for the total of 87 are +ve as well.

count(clean\_states[clean\_states$lng>0,])

# A tibble: 1 × 1

n

*<int>*

1 87

Looking at the full dataset, there seem to be the same 86 records from florida and Vermont that have +ve longitude.

kk=vacc\_rec[vacc\_rec$lng>0,c("name","state","district","county","city","lat","lng")]

> kk=kk[!is.na(kk$name),]

> count(kk)

# A tibble: 1 × 1

n

*<int>*

1 86