Schema for 550

**Data we have**

1. Zika Cases

* With the zika data we want to calculate the total number of cases, the number of cases per month, and predict the number of cases in the future (say year 2020)

Zika(

report\_date: date,

location: varchar(),

location\_type: varchar(),

data\_field: varchar(),

data\_field\_code: varchar(),

time\_period: varchar(),

time\_period\_type: varchar(),

value: int,

unit: varchar()

)

2. Temperature Data

* Temperature data source: <https://www.ncdc.noaa.gov/cag/>
* Get US states and territories monthly average temperature for 2016 (time interval during which zika data are available)
* Combine with mosquito data, as features to analyze and predict zika cases

Temp(

Date: int(),

Value: float()

)

3. Mosquito Data

* <https://www.epa.gov/climate-change-science/future-climate-change>
* Get latitude and longitude of the occurrence data of Aedes aegypti and Aedes albopictus in the United States and US territories
* We can use information about the temperature tolerances for each of the vectors
* With predicted temperature data, we can predict if the vectors will be there (regression in SQL)

Mos(

Vector: String()

Lat: float

Long: float

Country: String()

Primary Key: (Vector, Lat, Long)

Foreign Key: (Country)

)

4. LatLong Data

* Lat and long of each state and the US territories
* <http://www.latlong.net/> scrape data from here

Latlong(

State: string(2)

StateName: string(25)

Lat: float()

Long: float()

)