Identifying Areas at Risk for COVID-19 and Other Natural Disasters

DSI -11- San Francisco

Alicia Rodriguez – Brooks Jessup – Griffin Talan

Problem Statement

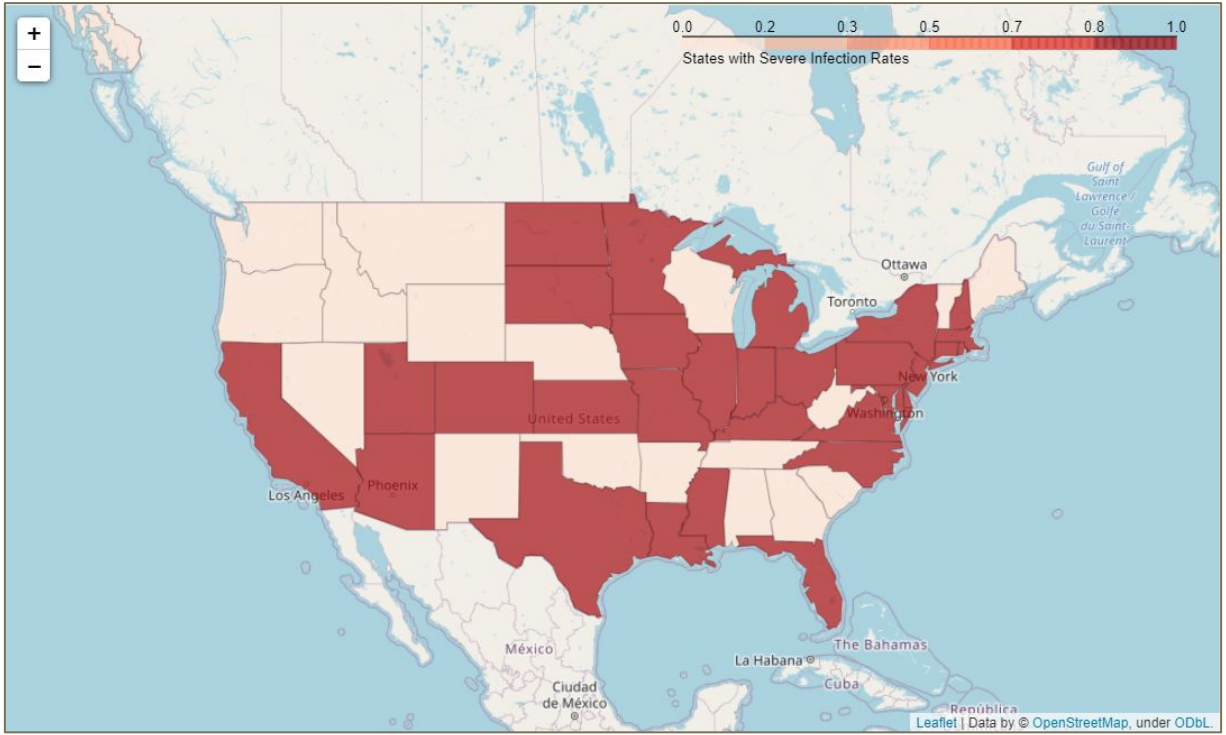
Convergence between COVID-19 with natural disasters risk spread.

Agenda

- Covid- 19
- -Data Collection
- -Data Analysis
- Natural Disasters
- -Nationwide –States
- California –Counties
- COVID/Natural Disaster History
- COVID/NATURAL Disaster 2020 Projected
- Conclusions and Next Steps

COVID -19

Identifying Areas at Risk for Covid-19



Steps

1. Data Collection
2. Data Evaluation
3. Data Wrangling
4. Data Analysis
5. Future Steps

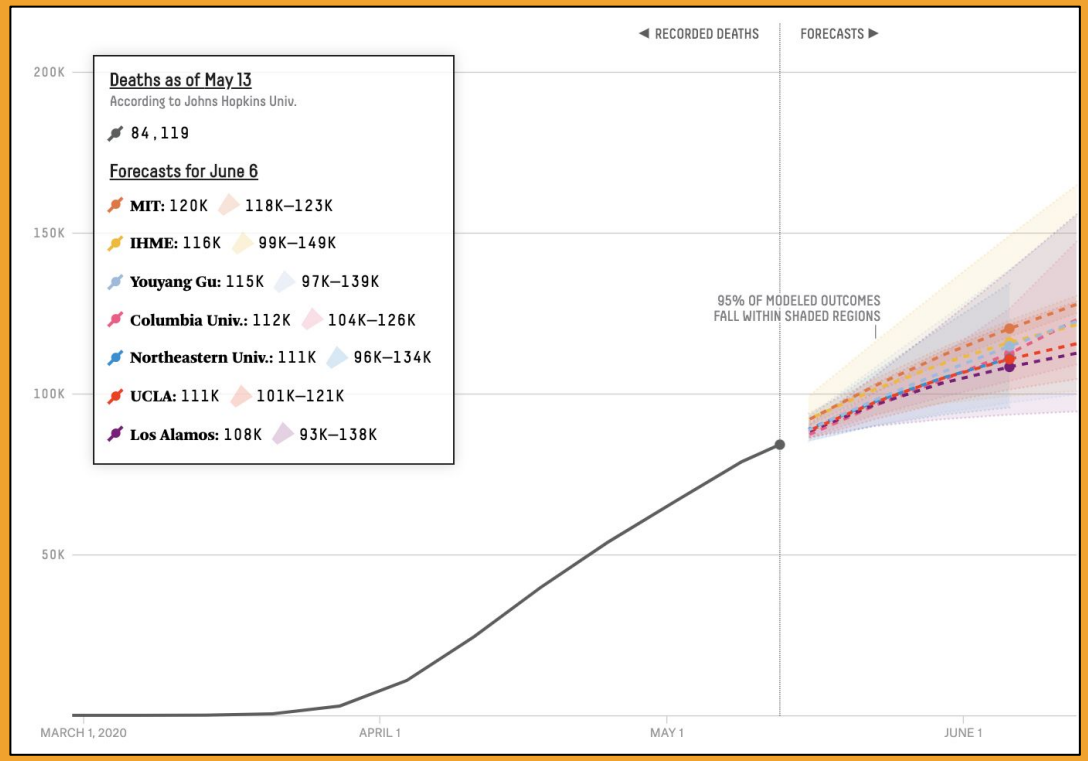
Step 1: Data Collection

The “wild west” of COVID data
- “A tidal wave of data”

Divergent projections
- Fatality rate variability
- Policy shifts

Historic uncertainty
- Reporting inconsistencies
- Insufficient testing

Selection criteria
- Expertise
- Transparency
- Thoughtfulness
- Distance of projection



Source: *FiveThirtyEight*

Sources:
Satchit Balsari , Caroline Buckee and Tarun Khanna, “Which Covid-19 Data Can You Trust?” *Harvard Business Review* (May 8. 2020).
Kevin Dayaratna and Norbert Michel, “The Challenges of Forecasting the Spread and Mortality of COVID-19”, *The Heritage Foundation* (April 15, 2020).
Maggie Koerth, Laura Bronner and Jasmine Mithani, “Why It’s So Freaking Hard to Make a Good COVID-19 Model,” *FiveThirtyEight* (March 31, 2020).

Step 1: Data Collection

1. Historic dataset (US & CA)
Center for Systems Science and Engineering (CSSE), *Johns Hopkins University*
https://github.com/CSSEGISandData/COVID-19

2. Projected dataset (US)
Institute for Health Metrics and Evaluation (IHME), *University of Washington*
http://www.healthdata.org/covid

3. Projected dataset (CA)
Mailman School of Public Health (MSPH), *Columbia University*
https://github.com/shaman-lab/COVID-19Projection

Step 2: Data Evaluation

Non-linear regression algorithm:

$$D(t;\alpha,\beta,p) = \frac{p}{2}\Psi(\alpha(t-\beta)) = \frac{p}{2}\left(1 + \frac{2}{\sqrt{\pi}}\int_0^{\alpha(t-\beta)} \exp\left(-\tau^2\right) d\tau\right)$$

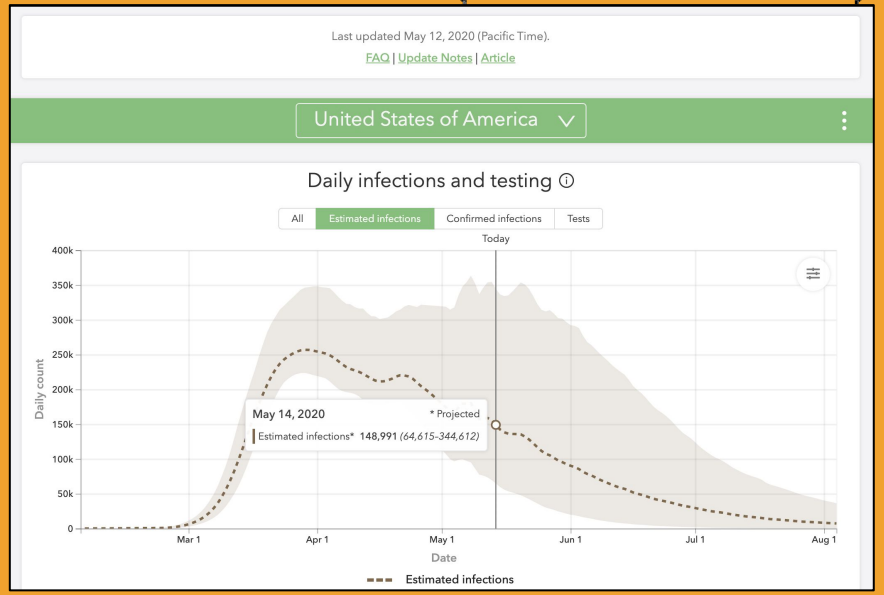
1. IHME Projection Model (U Washington)

Scope: National (Global)
Units: States
Dates: May 12th – Aug. 5th
Data: CSSE + case data
Audience: Governments and hospitals
Aim: Determine impact on health systems
Focus: Track the effect of reopening
Model: "Multi-stage Hybrid"

- Non-linear regression to fit curve of death rates
- Susceptible-Exposed-Infected-Recovered (SEIR)

Features:

- Mobility from anonymous cell phone data



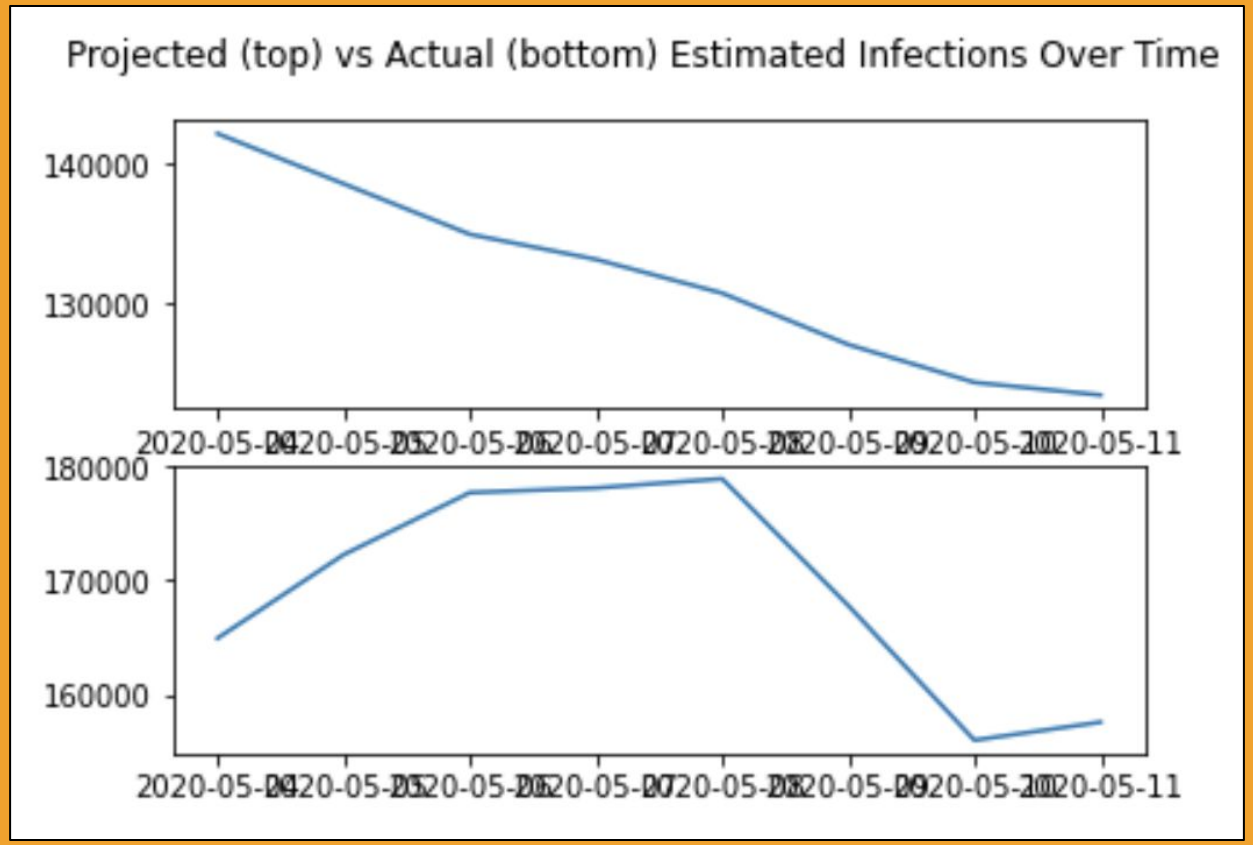
Sources:
http://www.healthdata.org/covid/faqs
https://www.medrxiv.org/content/10.1101/2020.04.21.20074732v1
http://www.healthdata.org/covid/updates

Step 2: Data Evaluation

1. IHME Projection Model
(U Washington)

Model Evaluation
(May 4th – May 11)
Metrics:

- RMSE 2647
- R2 Score 0.5353



Step 2: Data Evaluation

2. MSPH Projection Model (Columbia)

Scope: California (National)

Units: Counties

Dates: May 10th – June 20th

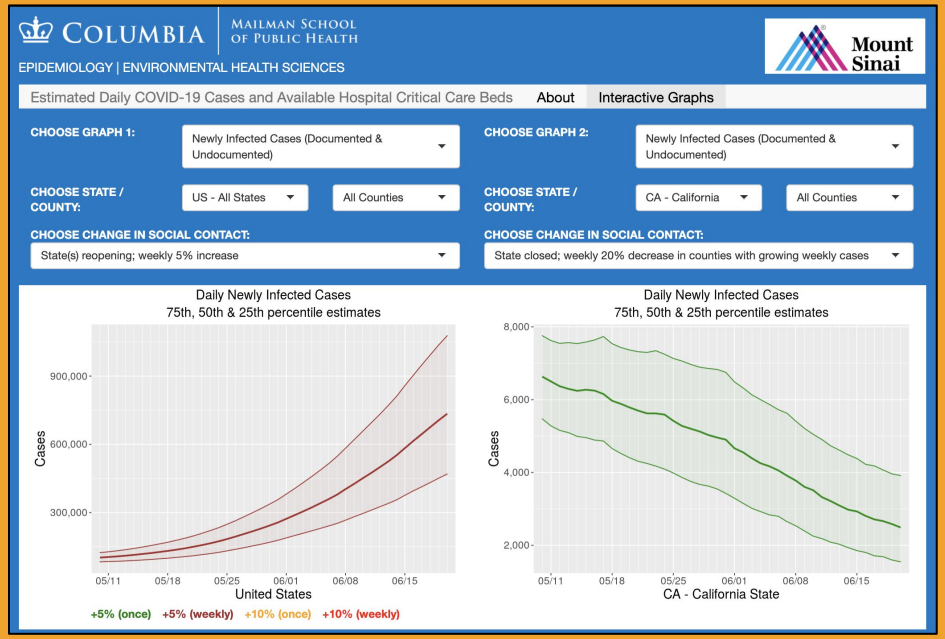
Data: USAFACTS (non-profit)

Model: Metapopulation model

- Simulates transmission of virus

Features:

- Different scenarios for rates of contact reduction.



Sources:

https://behcolumbia.files.wordpress.com/2020/05/yamana_etal_reopening_projections.pdf

http://www.columbia.edu/~jls106/pei_shaman_200324_projections.pdf

Step 3: Data Wrangling

General EDA

- Clean datasets
- Mostly reformatting
- IHME missing values

How to measure severity?

- Infections (vs deaths)
- Work back from deaths
- Variable fatality rates
- Scaled to 10 thou residents
- Threshold = 1/10 thou

Cleaned Dataframe with IHME Projections

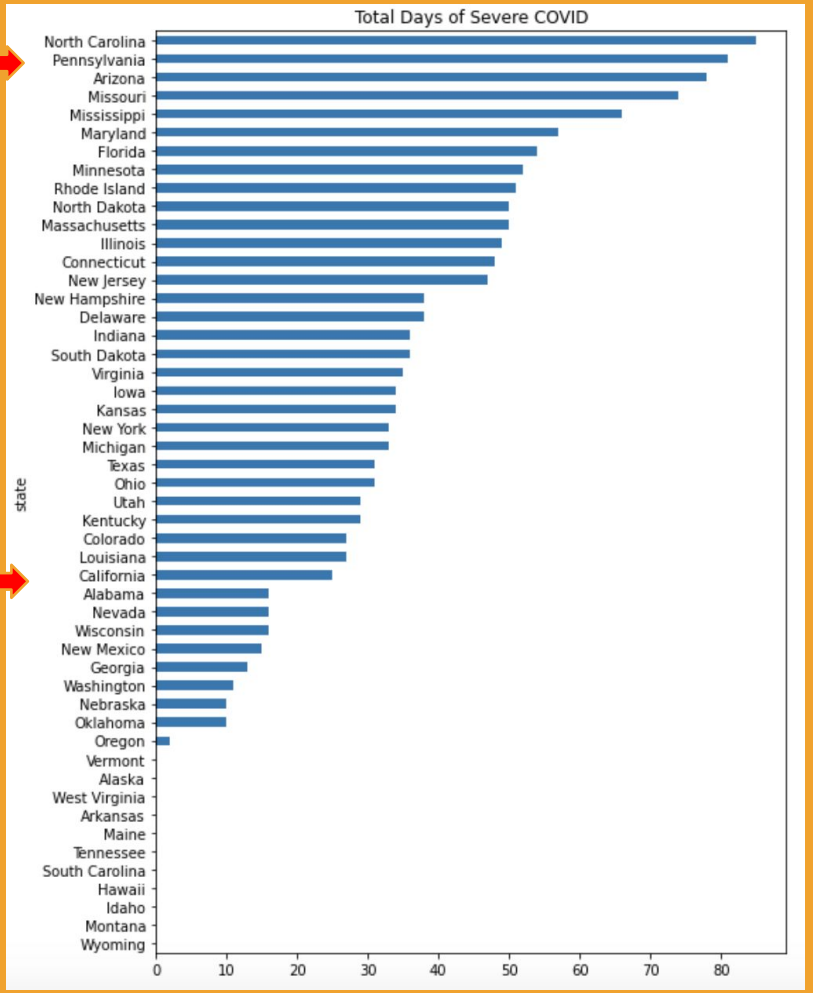
	state	date	est_new_infections	est_population	daily_infection_rate	severe
361	California	2020-06-02	4489.988918	39512223	1.136354	1
362	California	2020-06-03	4313.496519	39512223	1.091687	1
363	California	2020-06-04	4175.010712	39512223	1.056638	1
364	California	2020-06-05	4028.328279	39512223	1.019514	1
365	California	2020-06-06	3894.099070	39512223	0.985543	0
366	California	2020-06-07	3752.315407	39512223	0.949659	0
367	California	2020-06-08	3610.985939	39512223	0.913891	0
368	California	2020-06-09	3533.655590	39512223	0.894320	0
369	California	2020-06-10	3440.330730	39512223	0.870700	0
370	California	2020-06-11	3301.861736	39512223	0.835656	0

Step 4: Data Analysis

National Projections

1. State rankings by severity

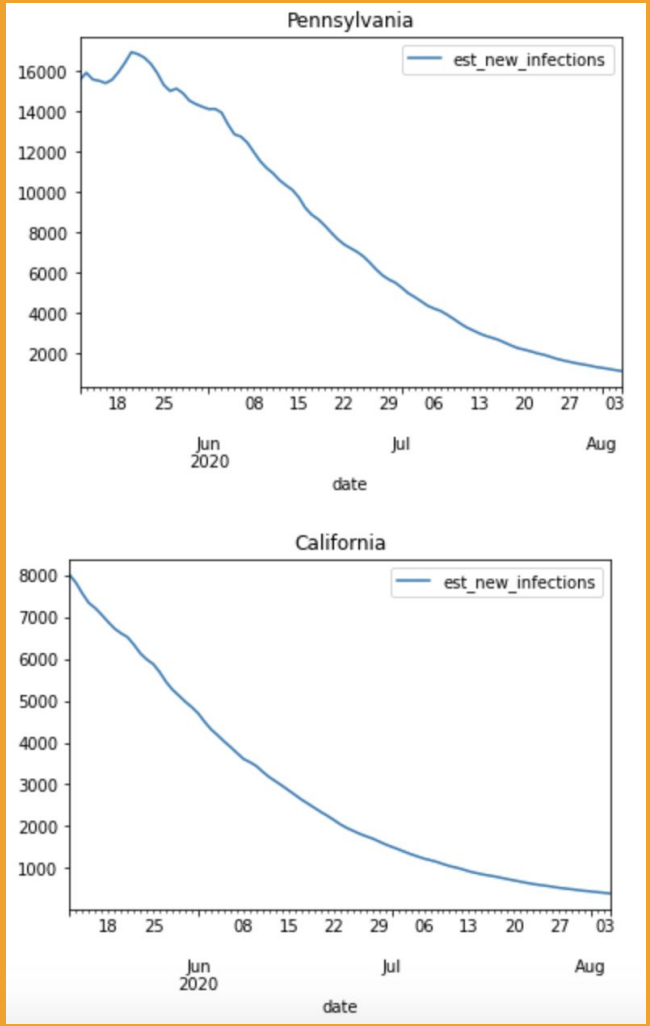
2. Daily infection rates



Step 4: Data Analysis

National Projections

1. State rankings by severity
2. **New infections daily**

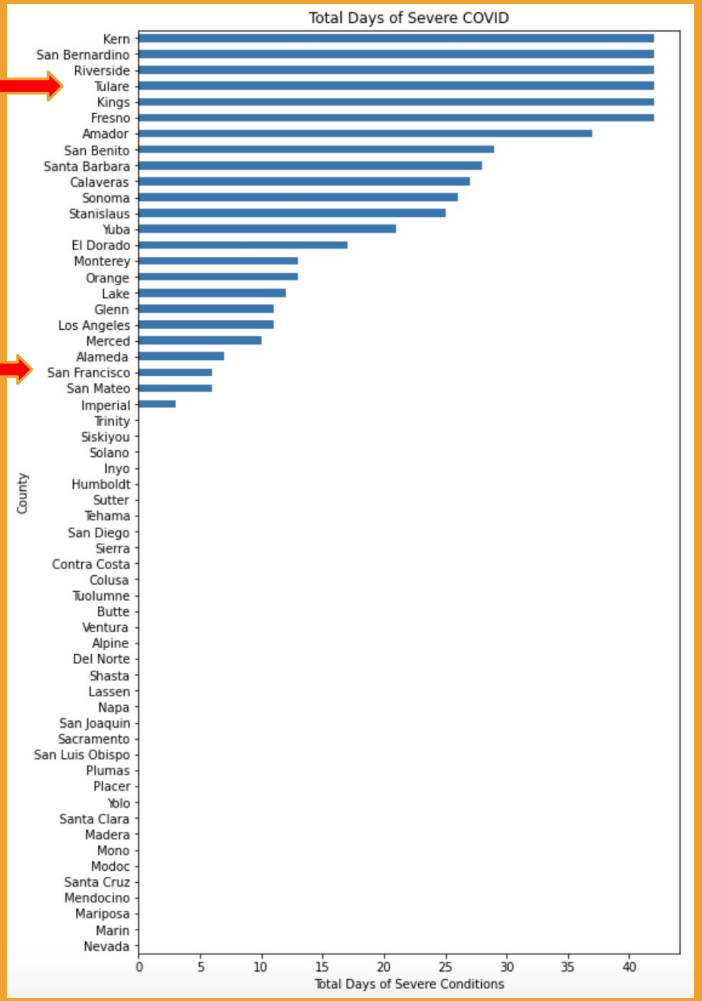


Step 4: Data Analysis

California Projections

1. County rankings by severity

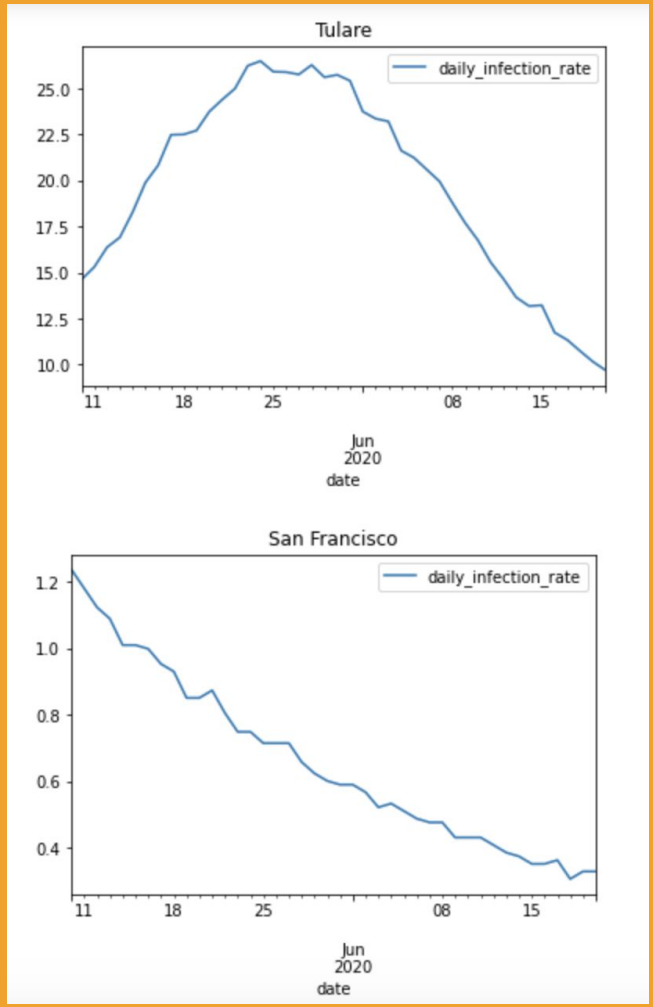
2. Daily infection rates



Step 4: Data Analysis

California Projections

1. County rankings by severity
2. **Daily infection rates** (1/10 thou)



Natural Disasters

Historic Data

FEMA: Federal Emergency Management Agency

Data: Disaster Declaration Summary

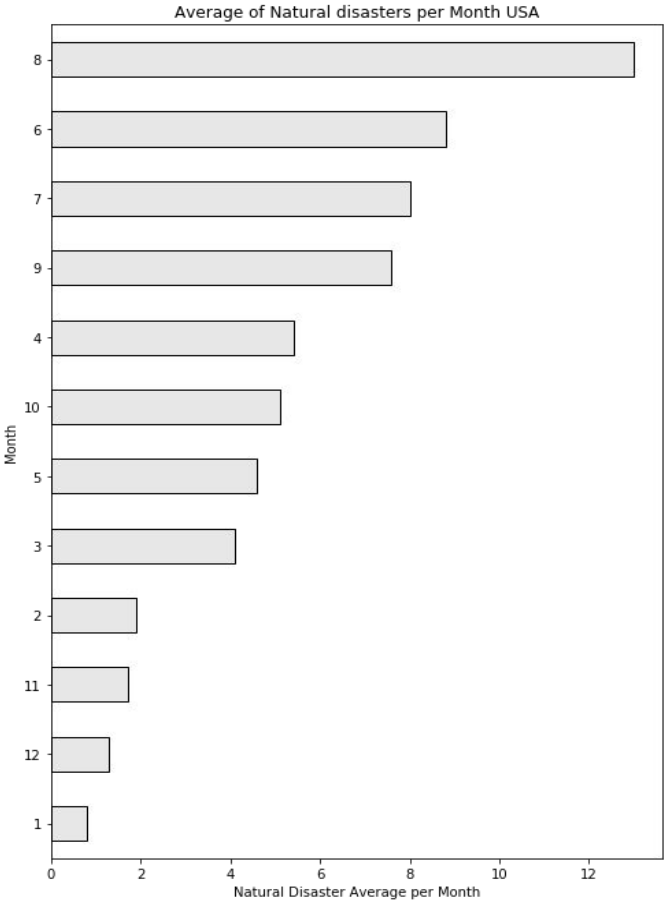
2010 - 2020

Nationwide: USA - states

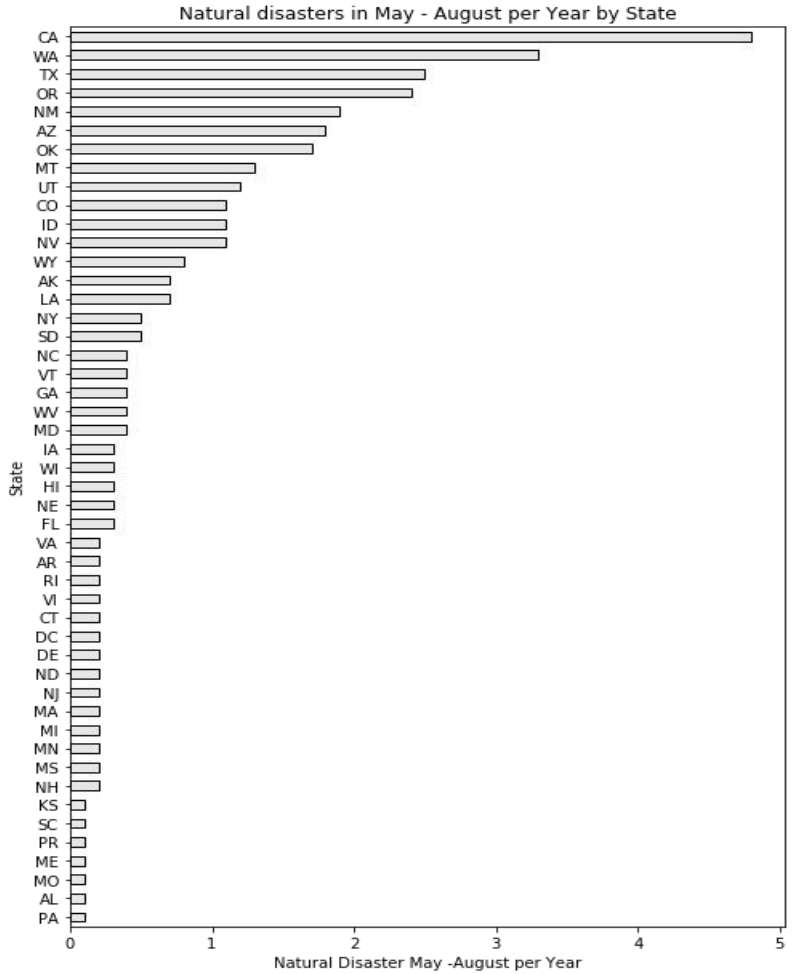
State: CA - counties

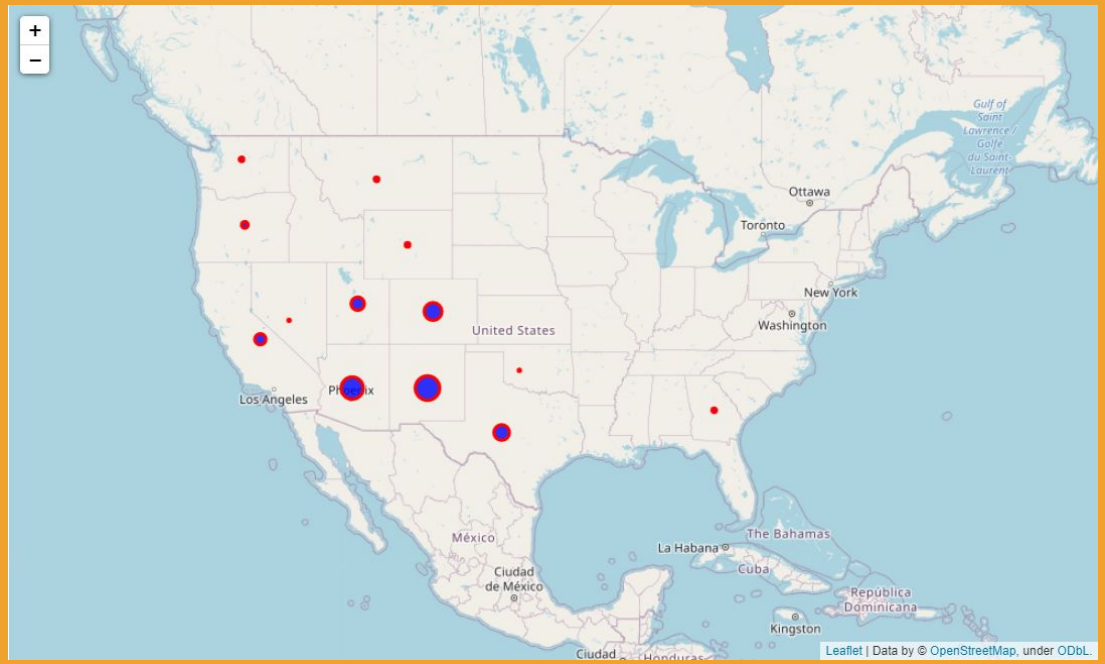
Evacuation in the Event of a Natural Disaster

Average of Natural Disasters per Month Nationwide



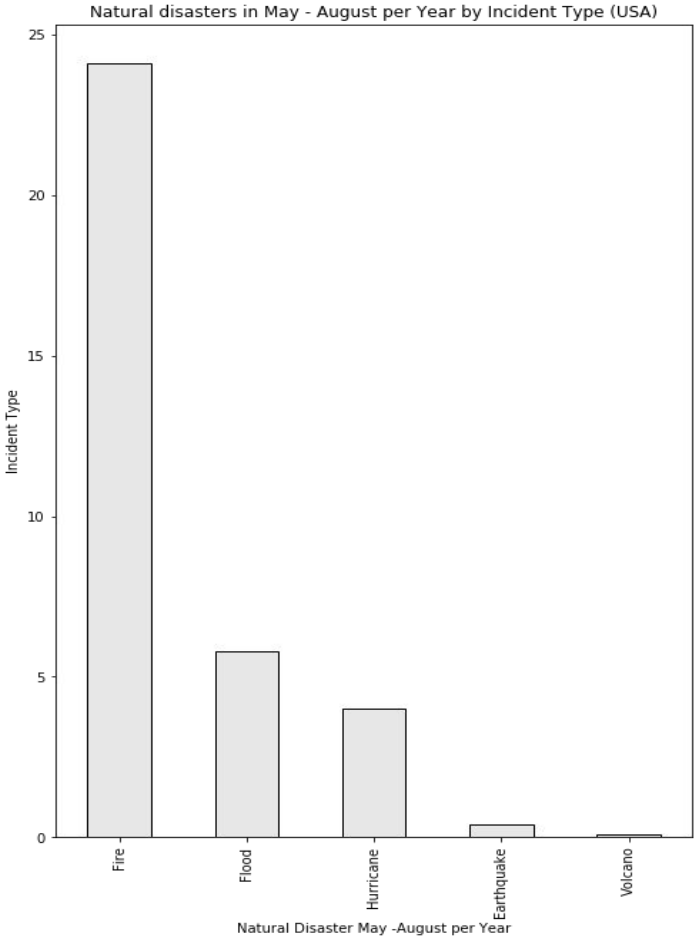
Natural Disasters from May through August per State





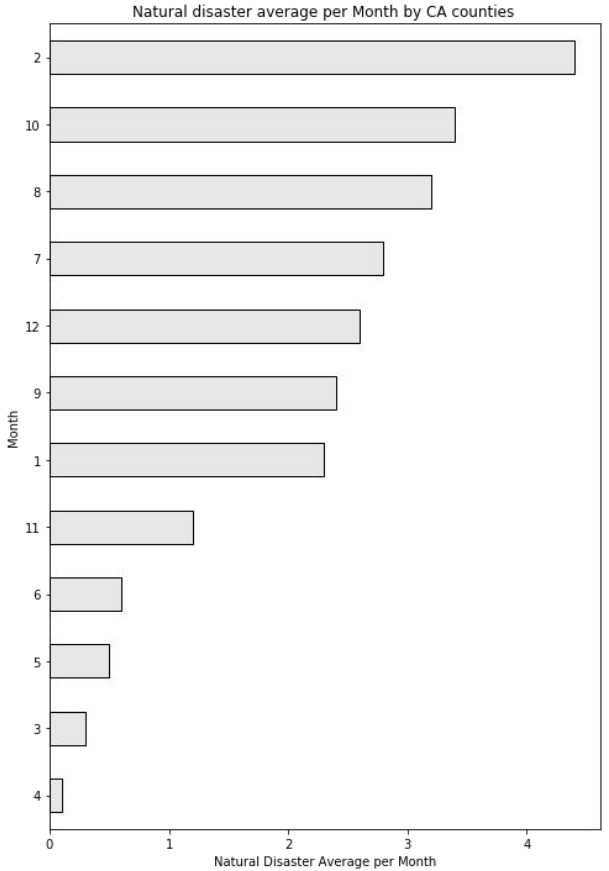
Natural Disasters in June by states

Natural Disasters from May to August per Year by Incident Type (USA)

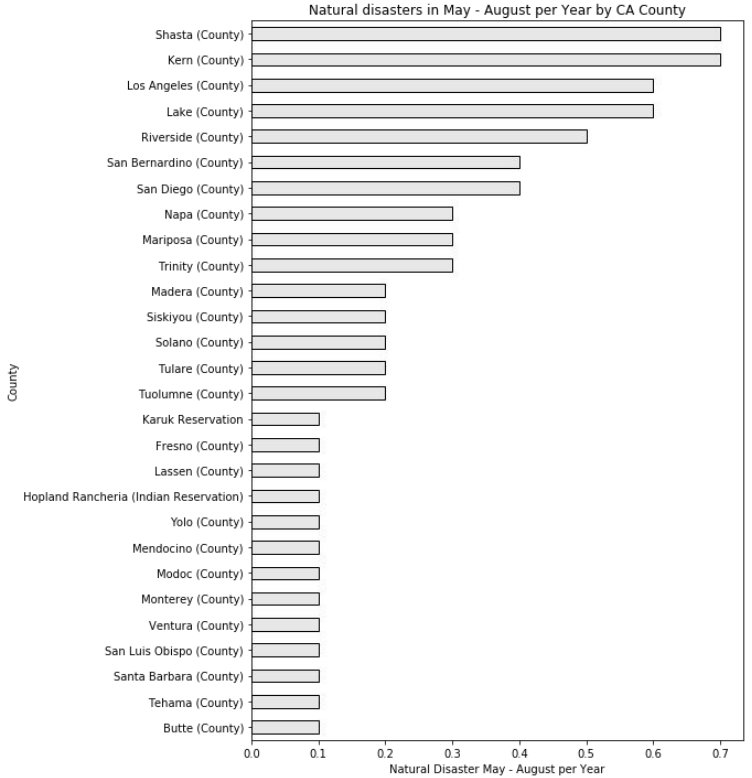


CALIFORNIA by Counties

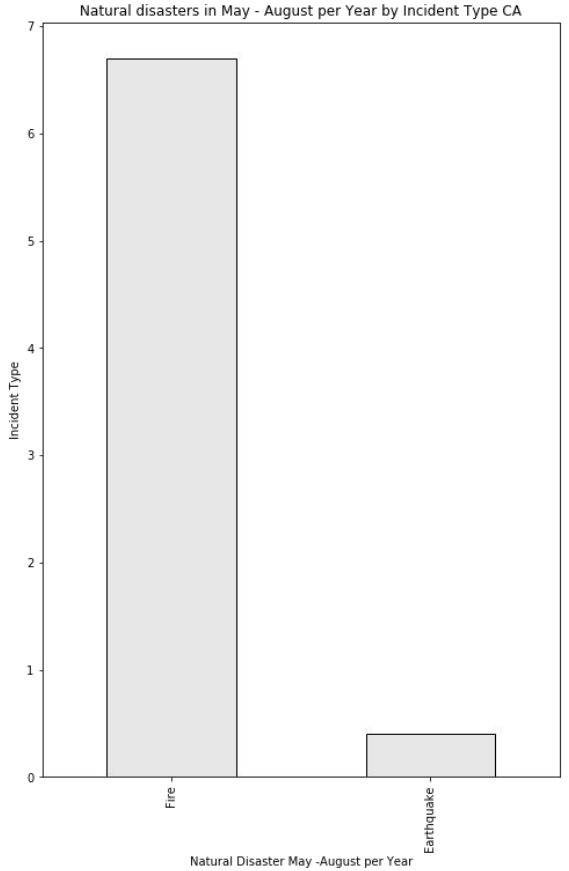
Natural Disaster Average per Month by CA Counties



Natural Disasters Average from May – August per Year by County



Natural Disasters from May to August per Year by Incident Type



Summary

-Nationwide June through August are the months with more of these 6 natural disasters.

-Fires and floods are the main natural disasters that can affect with the spread of COVID-19 nationwide.

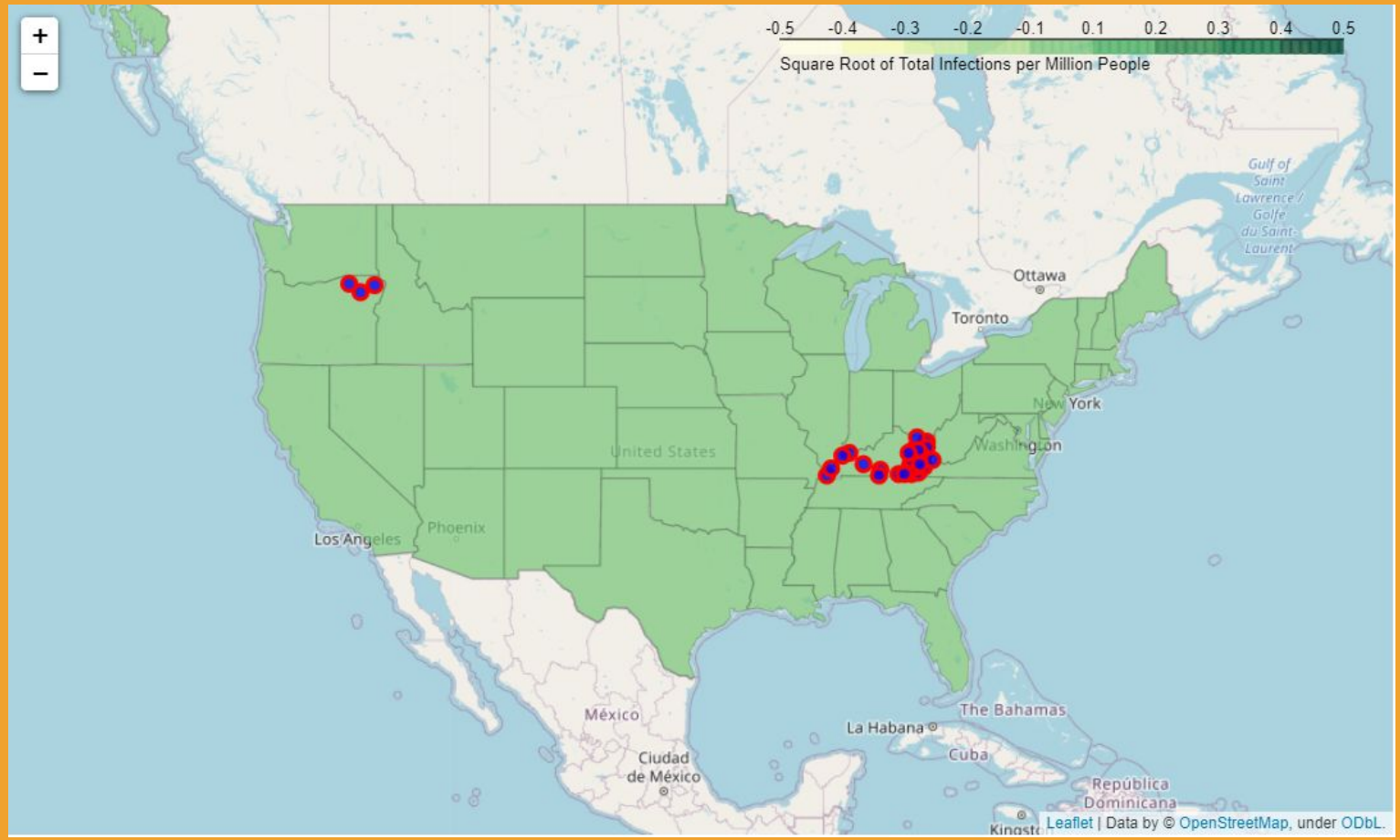
-California is the state with more of these 6 natural disasters per year.

-In Feb, October and August are the months with more of these 6 natural disasters in CA.

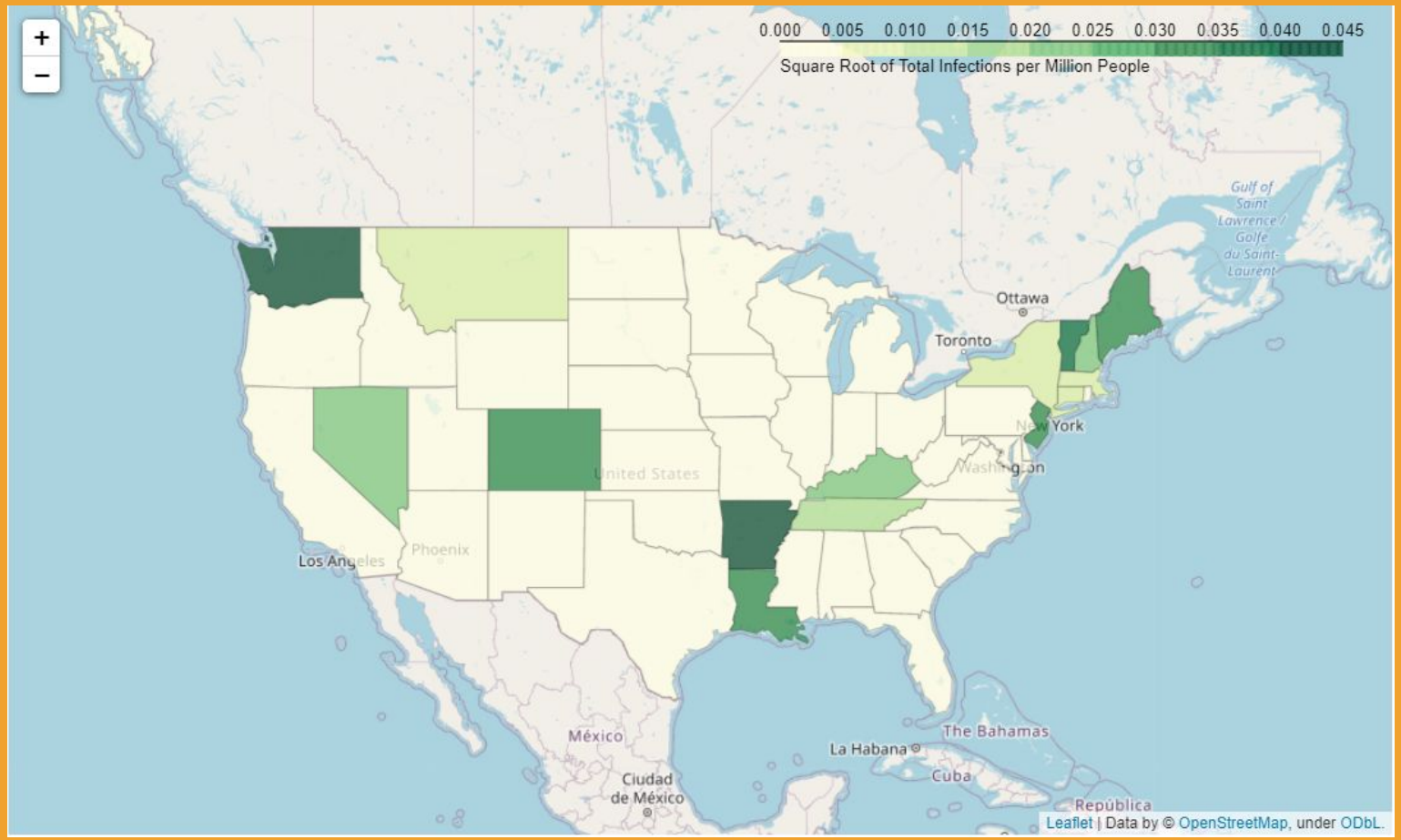
-Fires and earthquakes are the main natural disasters that might affect the spread of COVID-19 in the state of CA.

COVID / Natural Disaster 2020 History

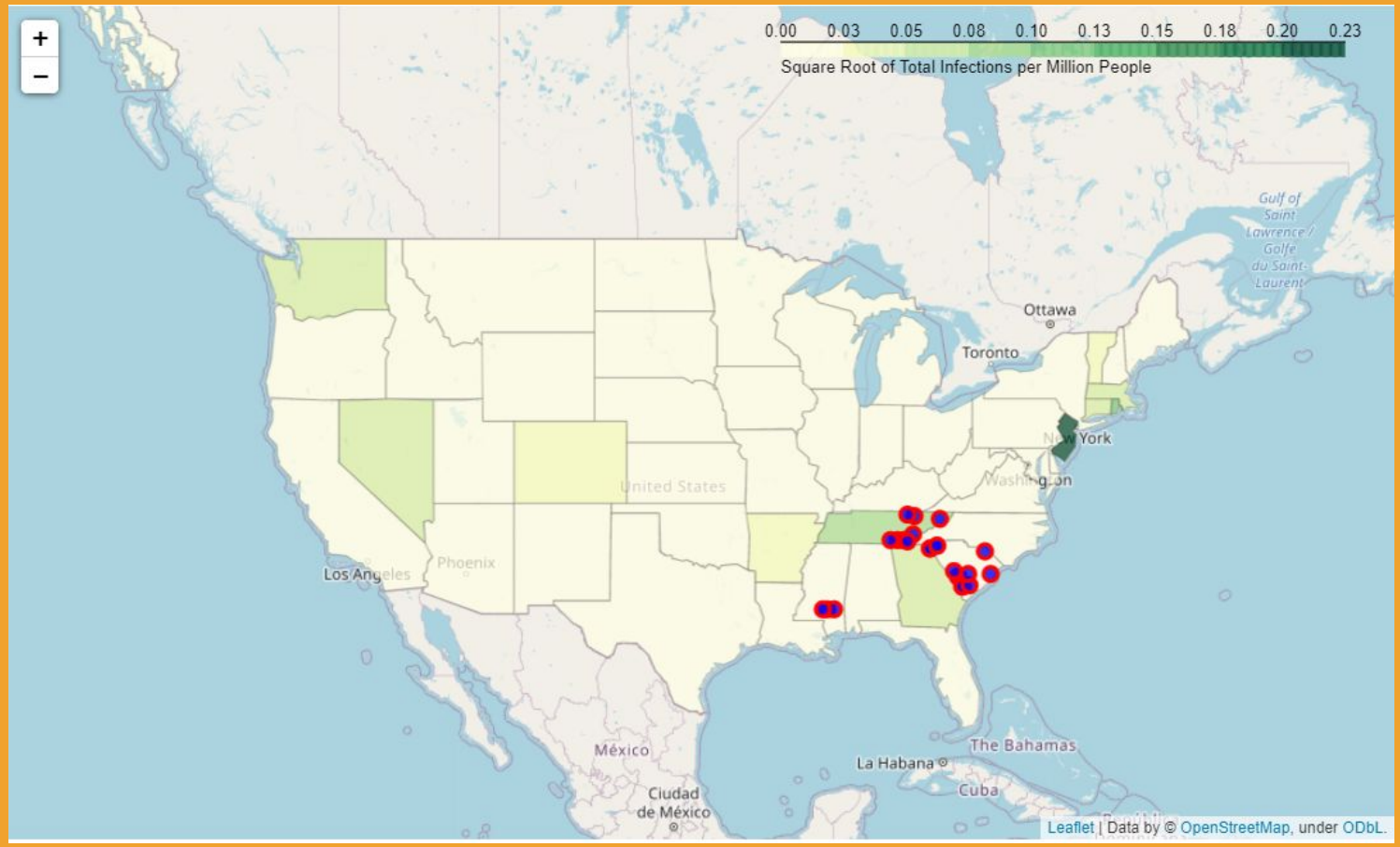
Jan 23- Beginning of COVID data



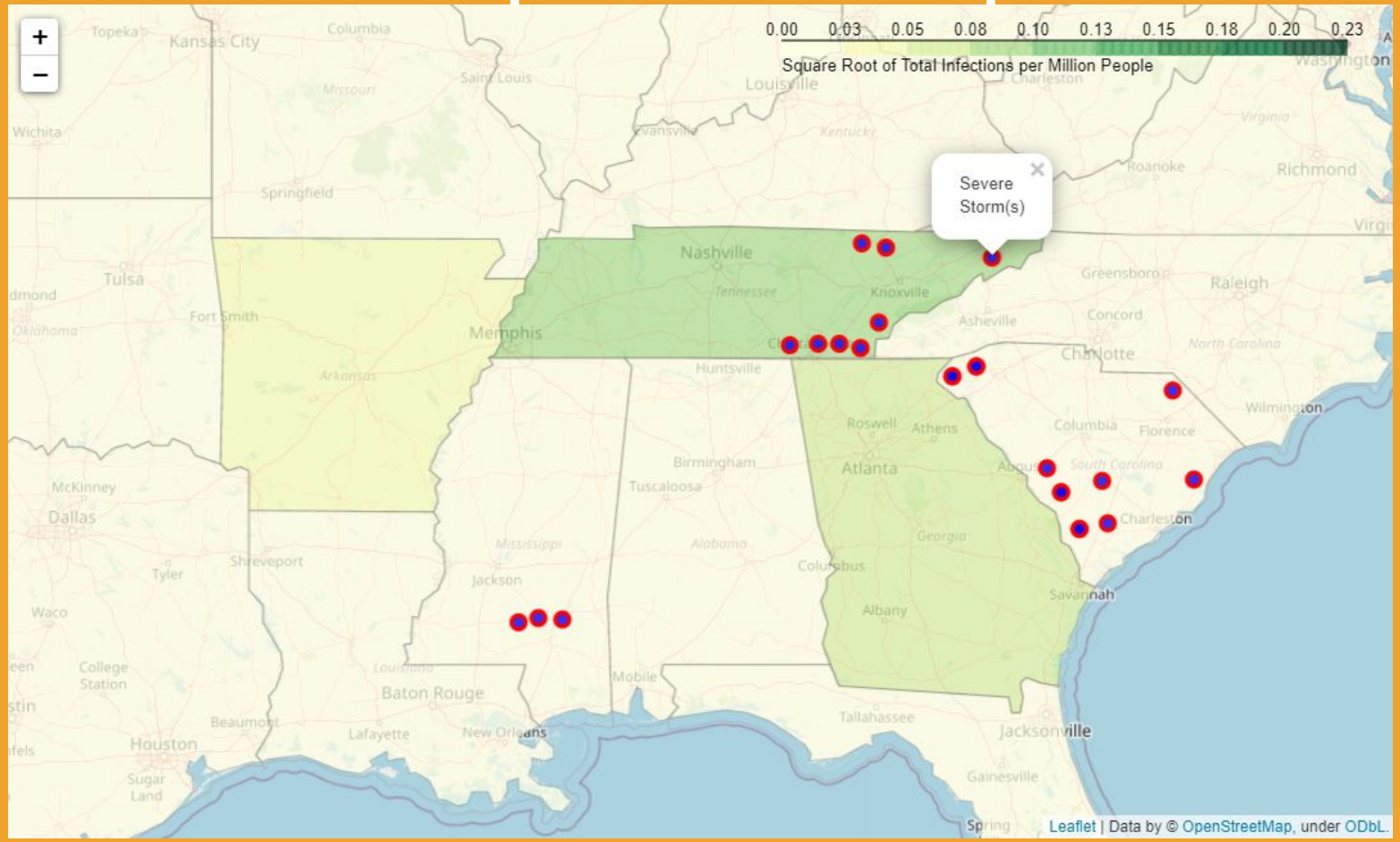
March 19 - signs of COVID emerging



April 1 - Major storms hit the southern states



April 1 - Closeup of storm path



SE Storm: Mid-April

"In Alabama, people seeking shelter from tornadoes huddled in community shelters, protective masks covering their faces to guard against the new coronavirus."

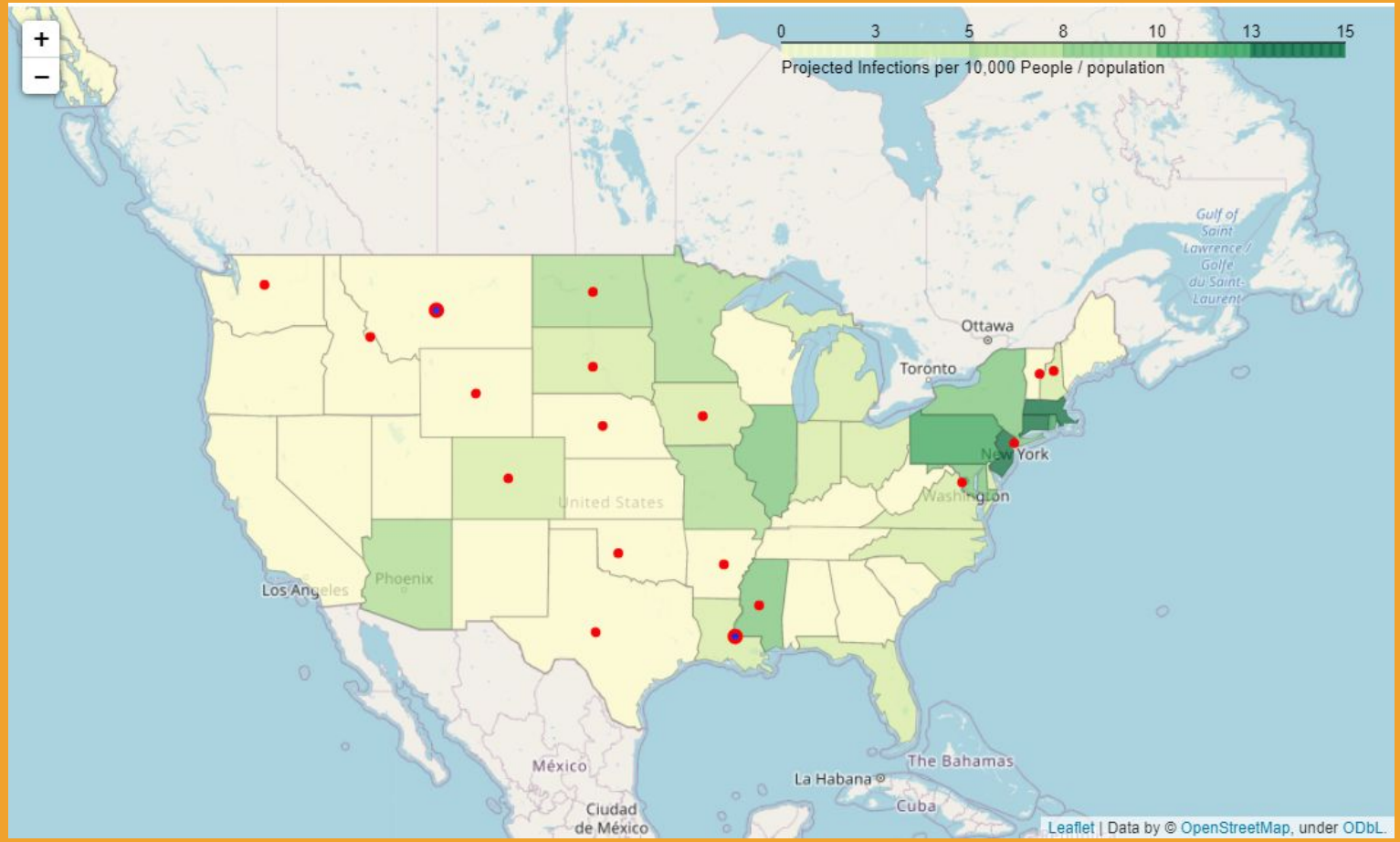
“Mississippi Gov. Tate Reeves said the storms were ‘as bad or worse than anything we’ve seen in a decade.’”

“Georgia Gov. Brian Kemp said some storm victims already were out of work because of shutdowns caused by COVID-19. ‘Now they have lost literally everything they own’”

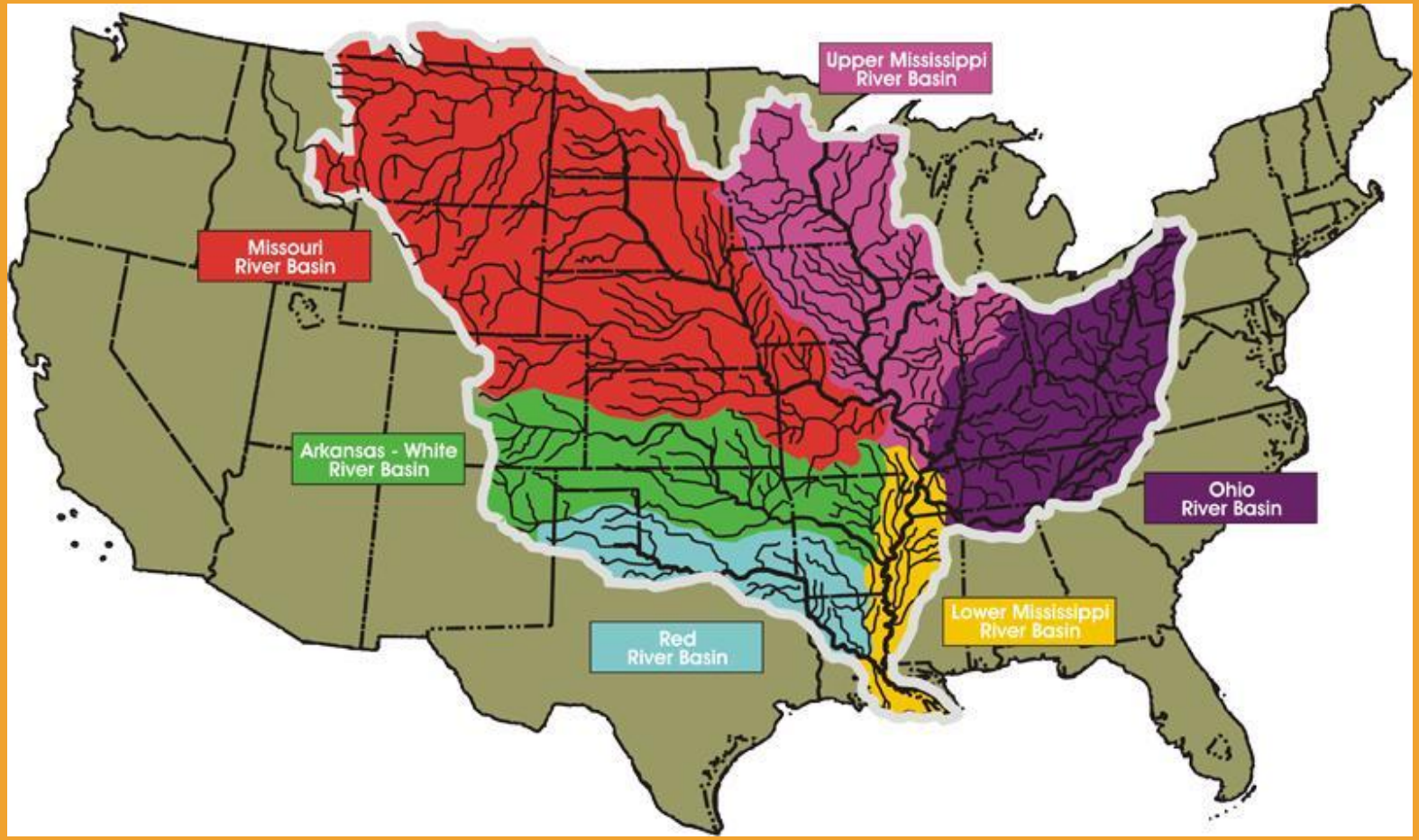
Source: www.usnews.com/news/us/articles/2020-04-13/easter-storms-sweep-south-killing-at-least-12-peoplehttps://www.usnews.com/news/us/articles/2020-04-13/easter-storms-sweep-south-killing-at-least-12-people

COVID / Natural Disaster 2020 Projected

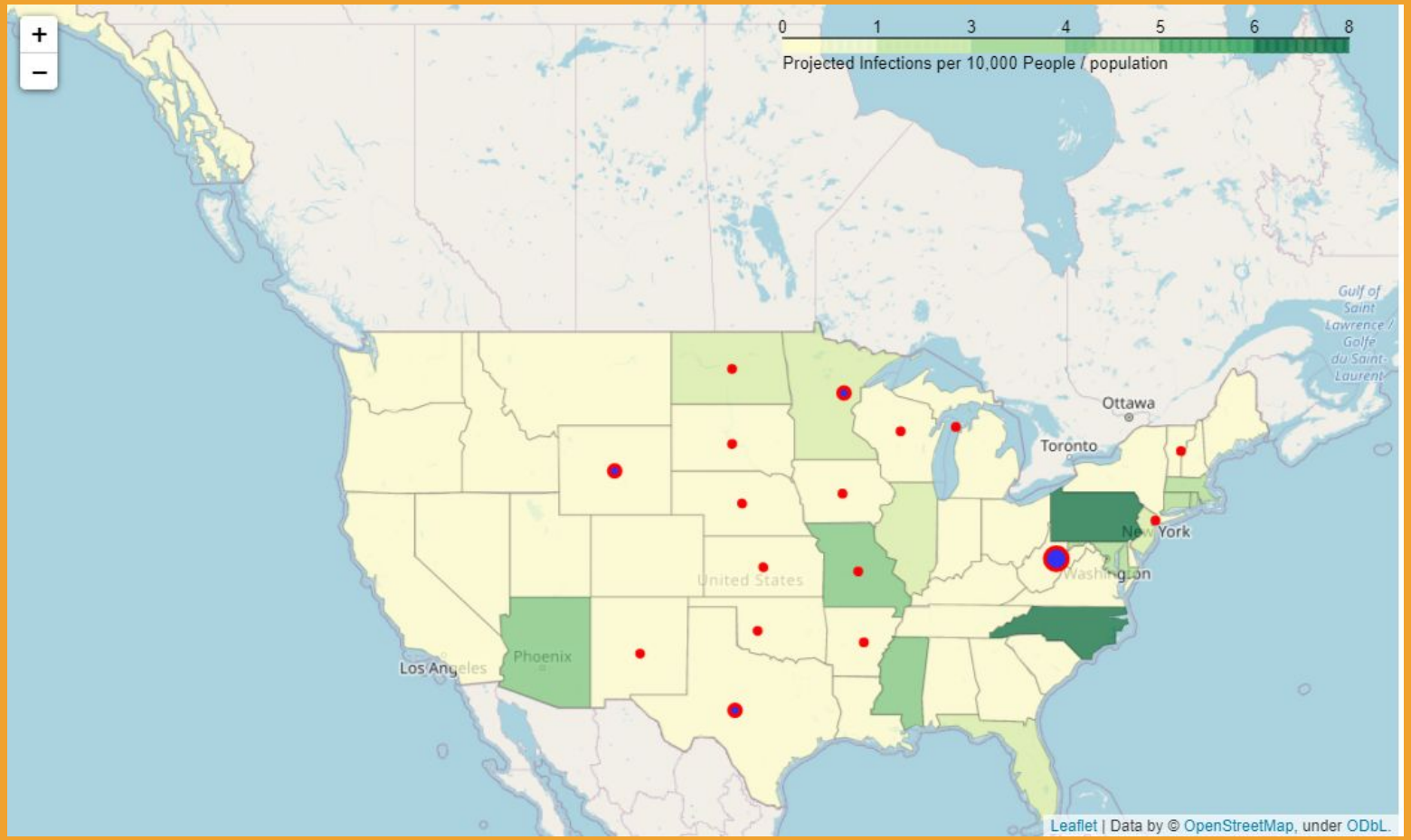
Flooding - May 15



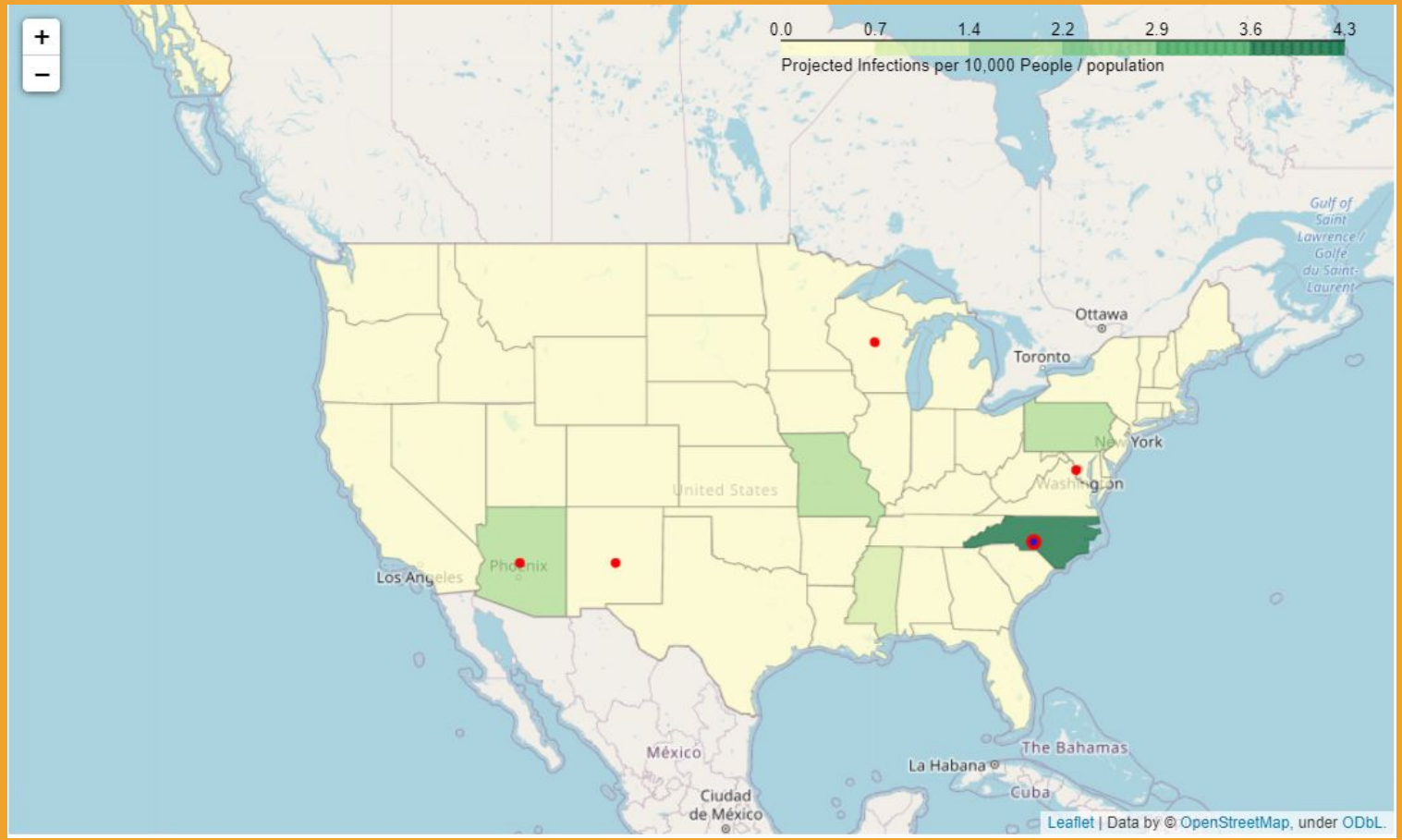
Flood Basins



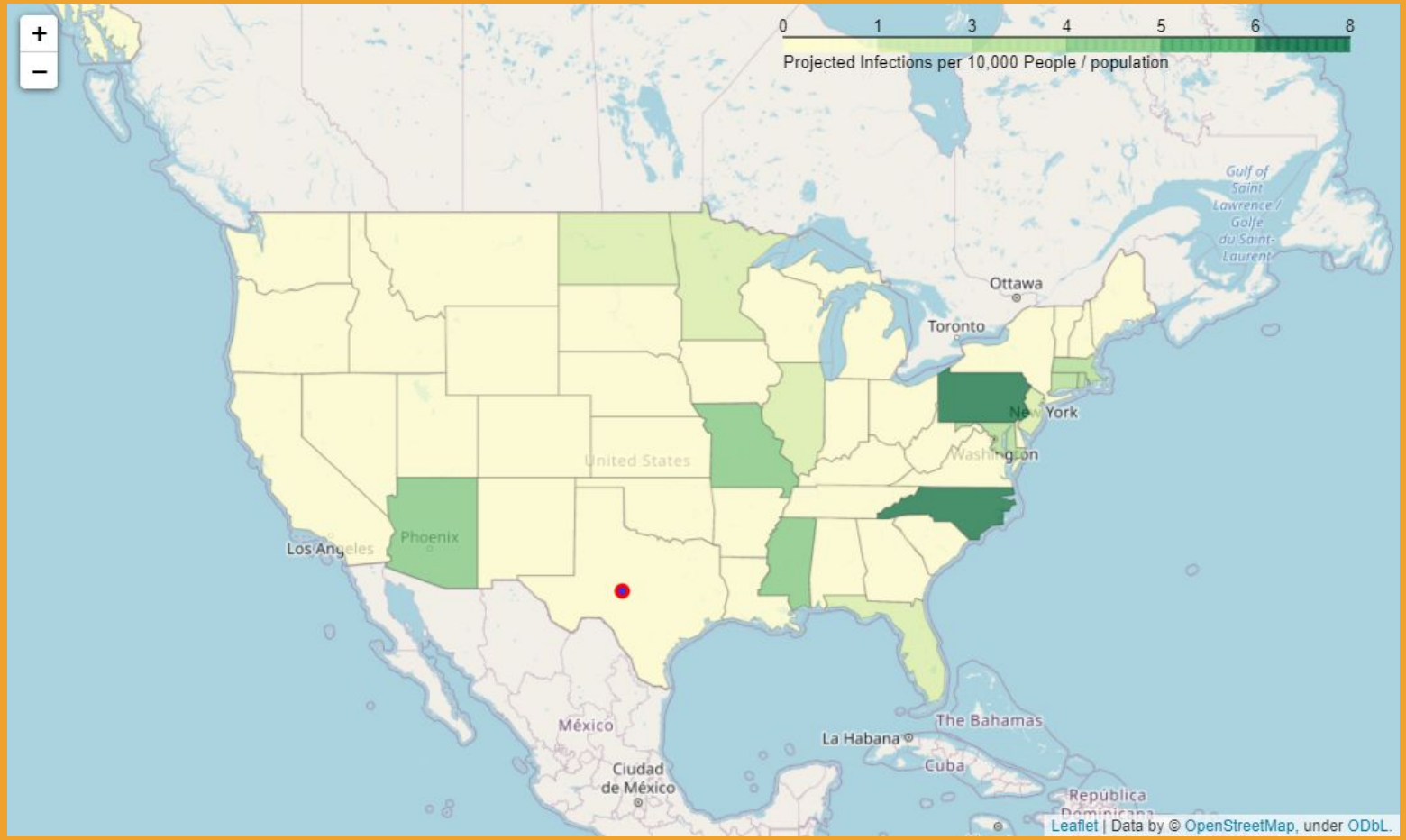
Flooding - June 15



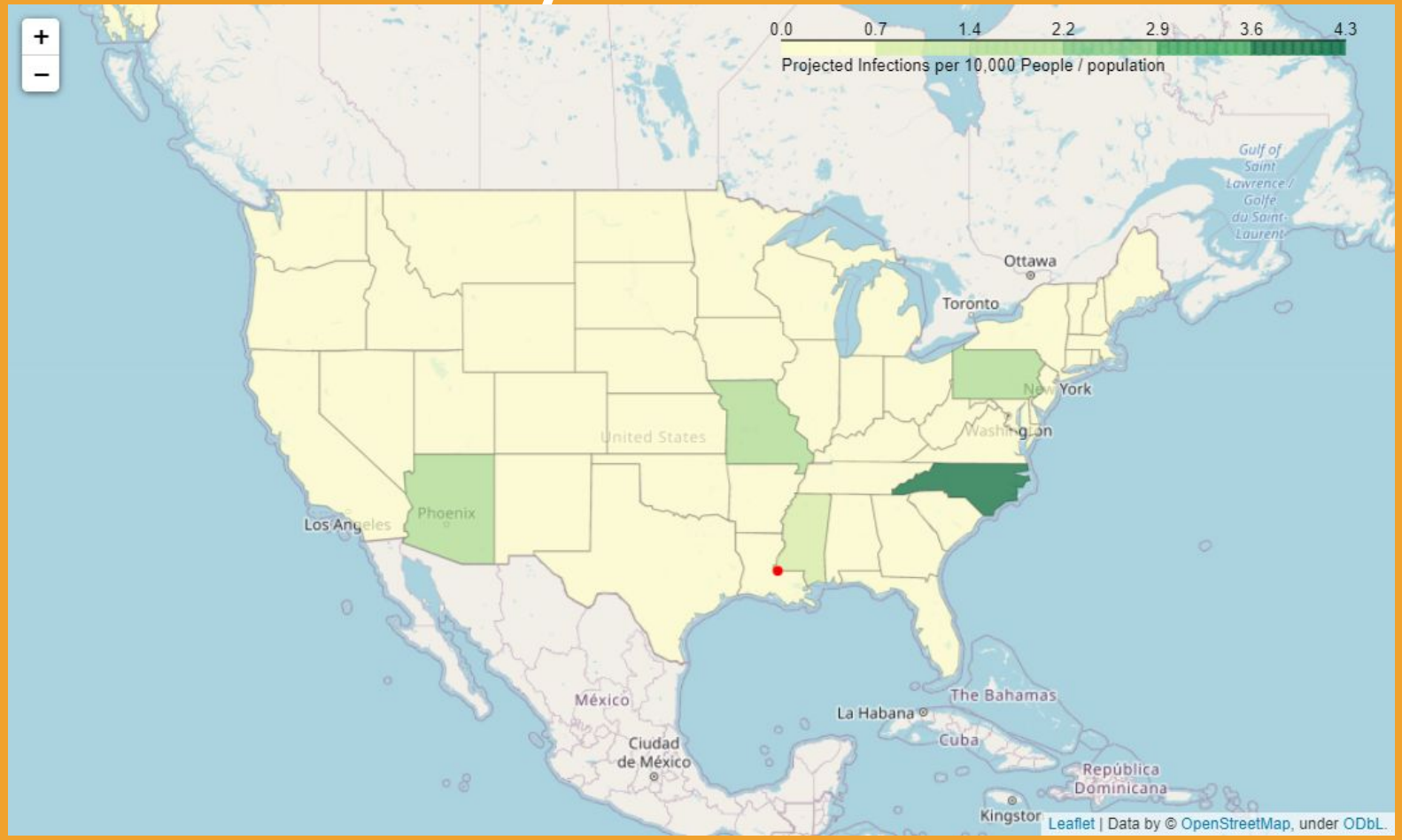
Flooding - July 15



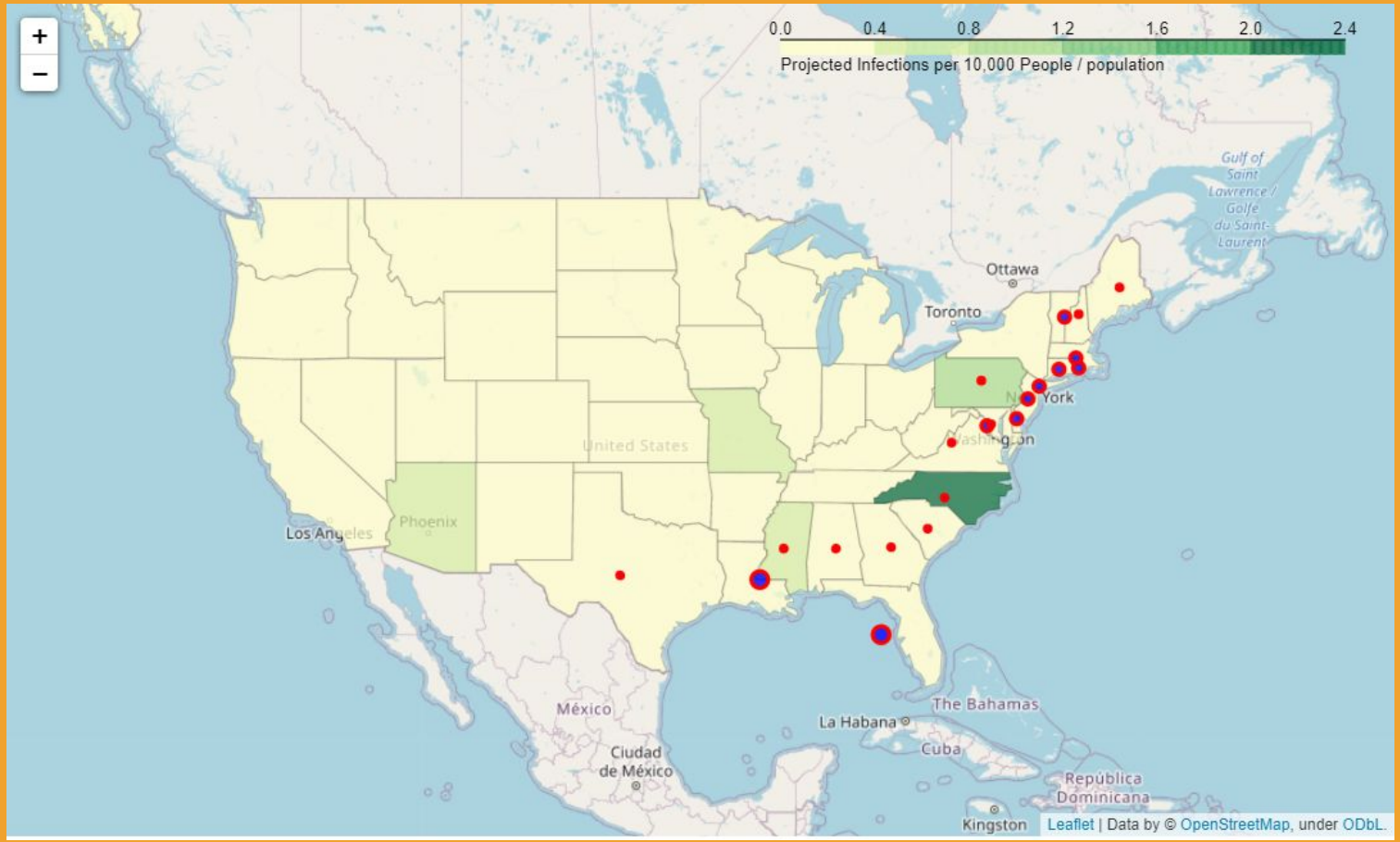
Hurricanes - June 15



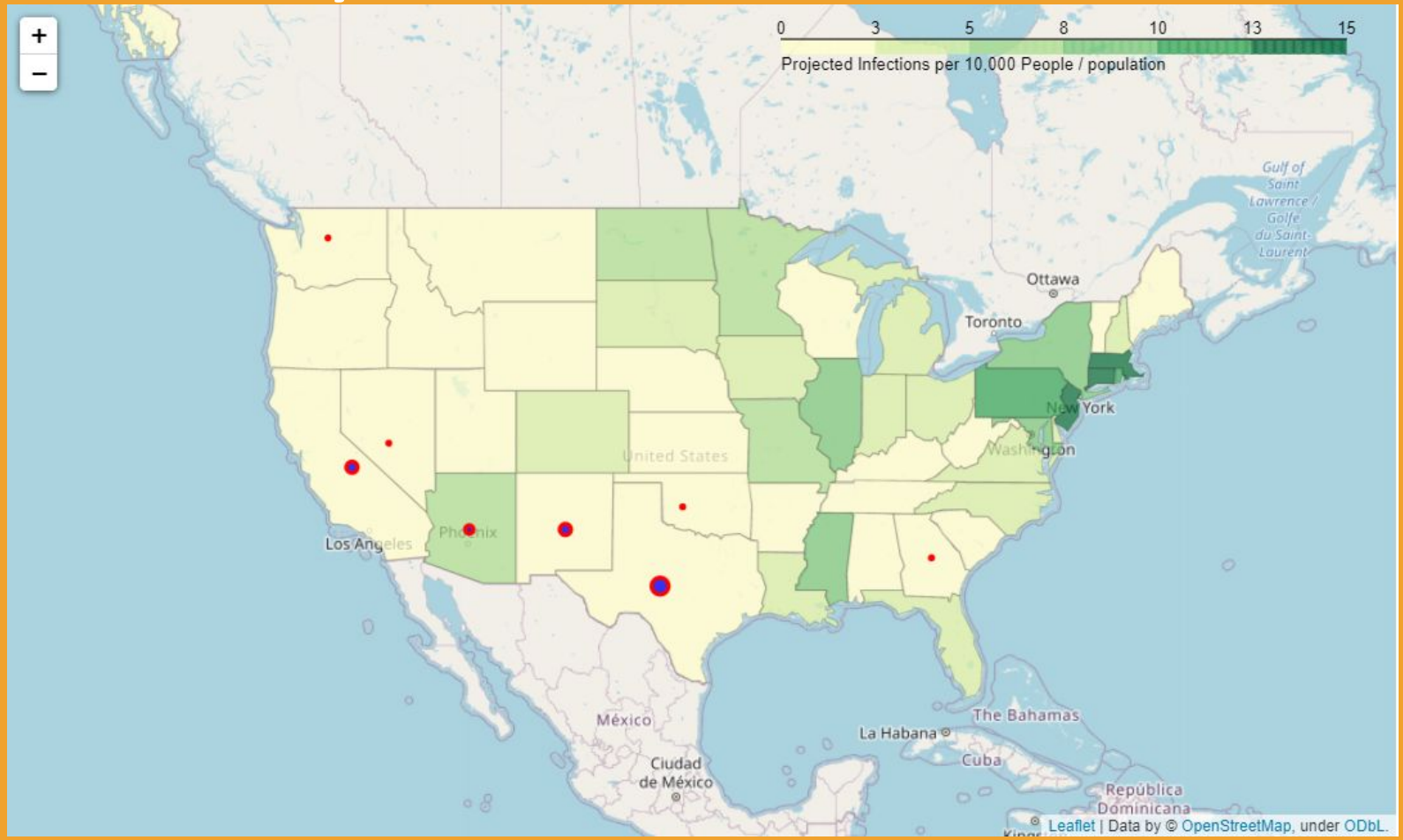
Hurricanes - July 15



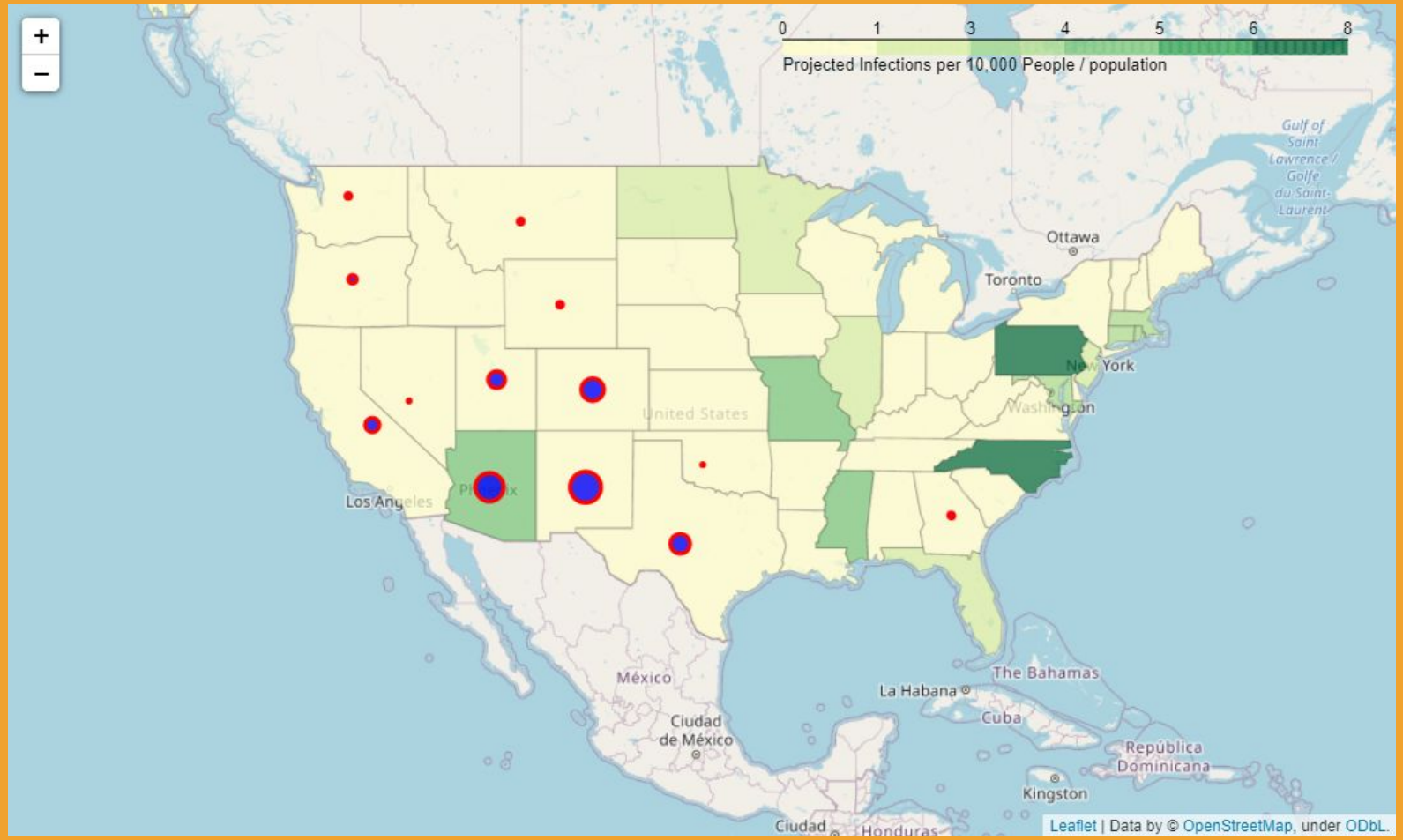
Hurricanes - August 12



Fires - May 15

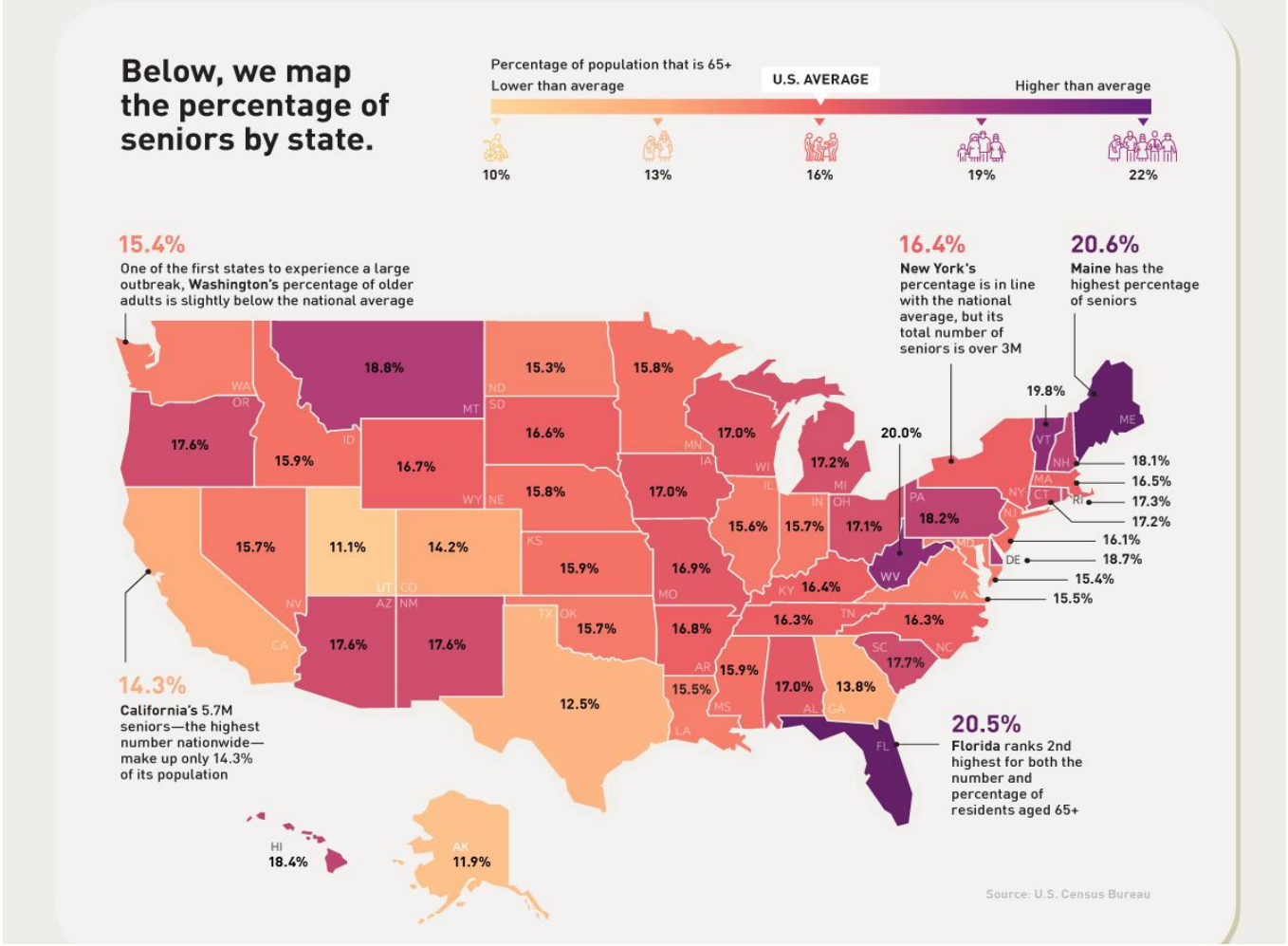

0 3 5 8 10 13 15
Projected Infections per 10,000 People / population
Gulf of Saint Lawrence / Golfe du Saint-Laurent
Ottawa
Toronto
New York
United States
Los Angeles
Phoenix
México
Ciudad de México
La Habana
The Bahamas
Cuba
República Dominicana
Leaflet | Data by © OpenStreetMap, under ODbL.

Fires - June 15

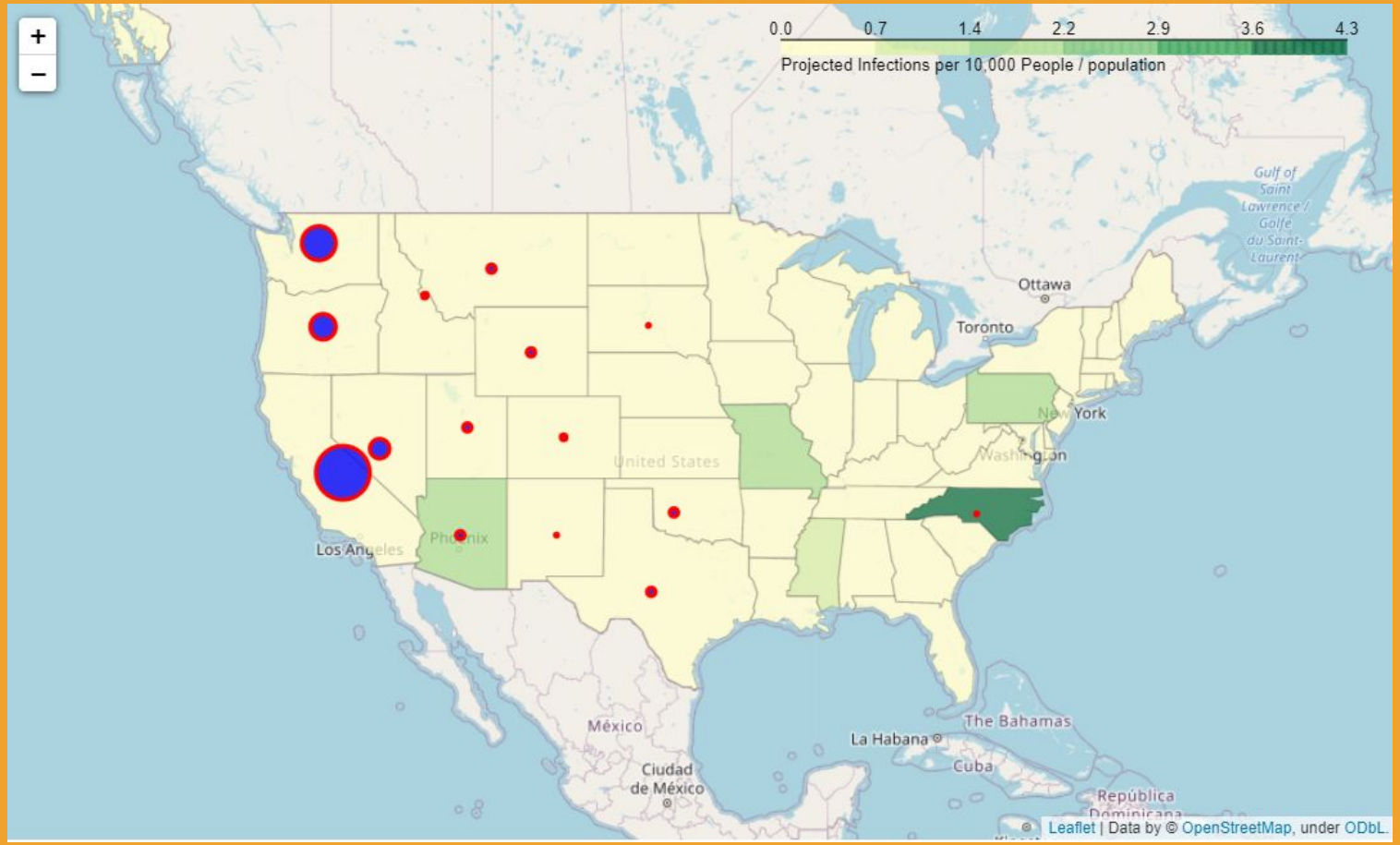


Source: https://www.abc15.com/news/state/navajo-nation-continues-to-be-hit-hard-by-coronavirus-as-arizona-reopens

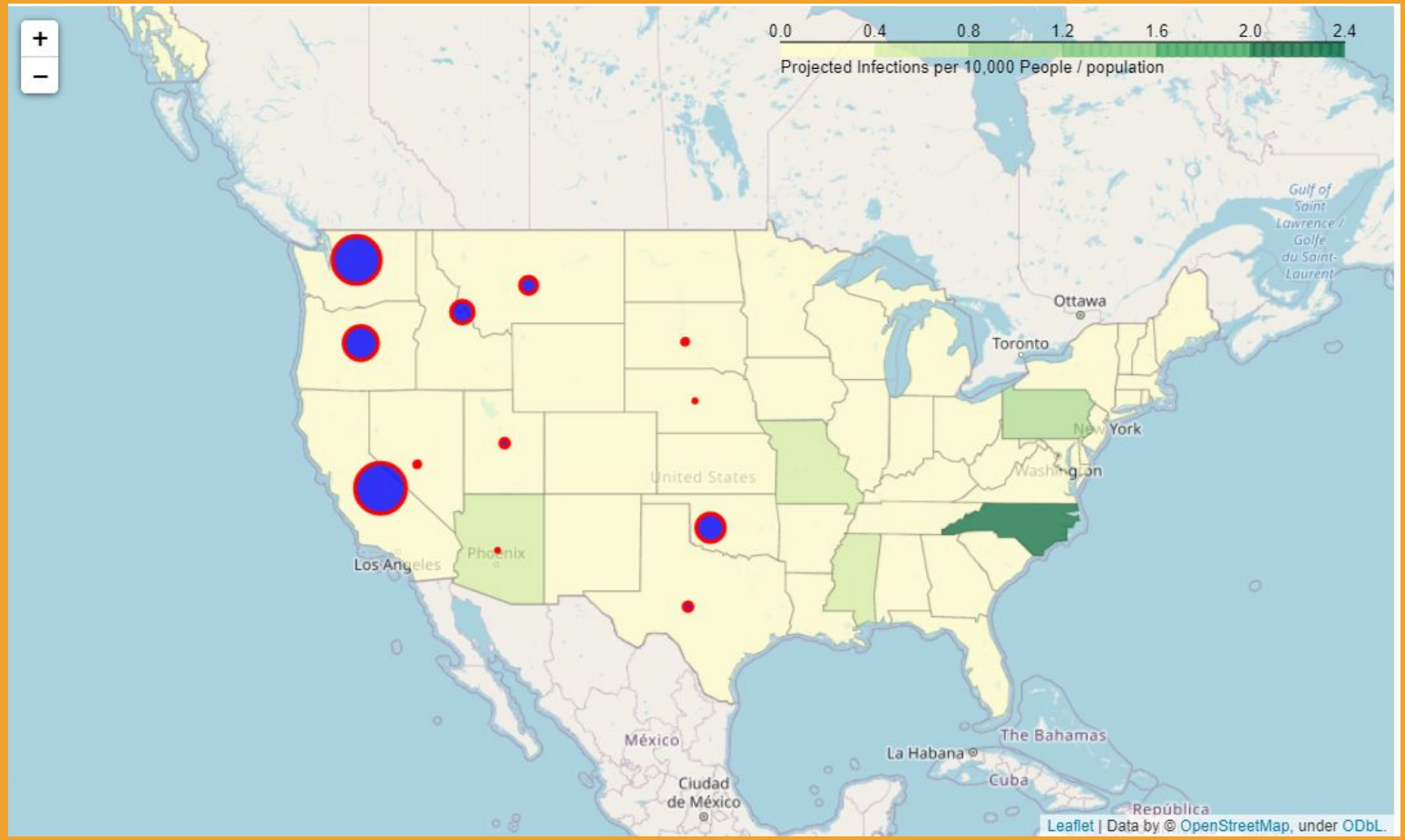
Source :
https://www.visualcapitalist.com/mapped-us-senior-population-covid-19/



Fires - July 15



Fires - August 12



Conclusions

Imminent threats:

- Flooding overlaps with COVID rise in basin states during May/June

Rapidly approaching threats:

- West coast fires and gulf storms

Long- term threats:

- Hurricane season overlapping with latest projected recovery states

Next Steps

- Explore financial impacts that are caused by natural disasters and how that could complicate the COVID response

- Update models with more recent natural disasters and see if there is a correlation between certain types of disasters and an increase in infection rates

Thank you

Questions?