A Web-based Geographic Framework to Detect and Visualize Space-time Clusters of COVID-19

Yu Lan & Eric Delmelle

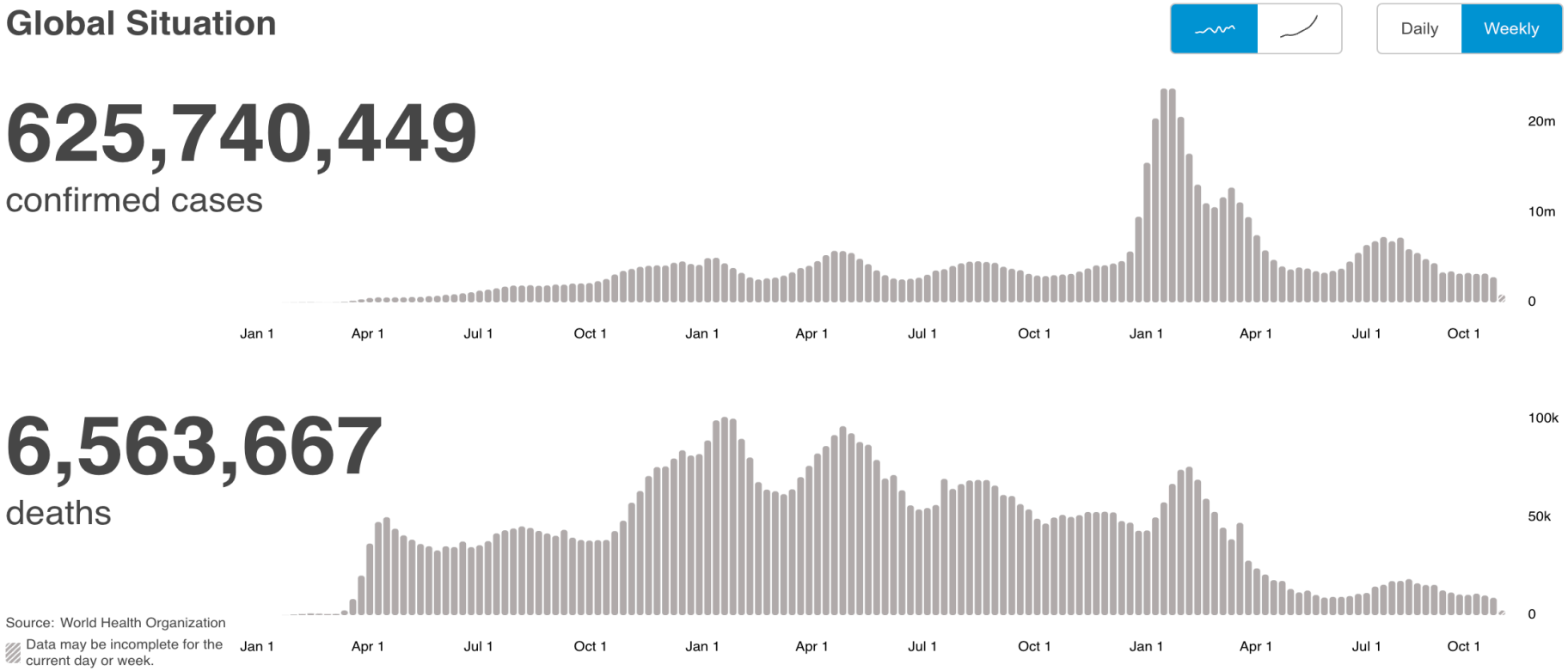
University of North Carolina at Charlotte

Department of Geography and Earth Science

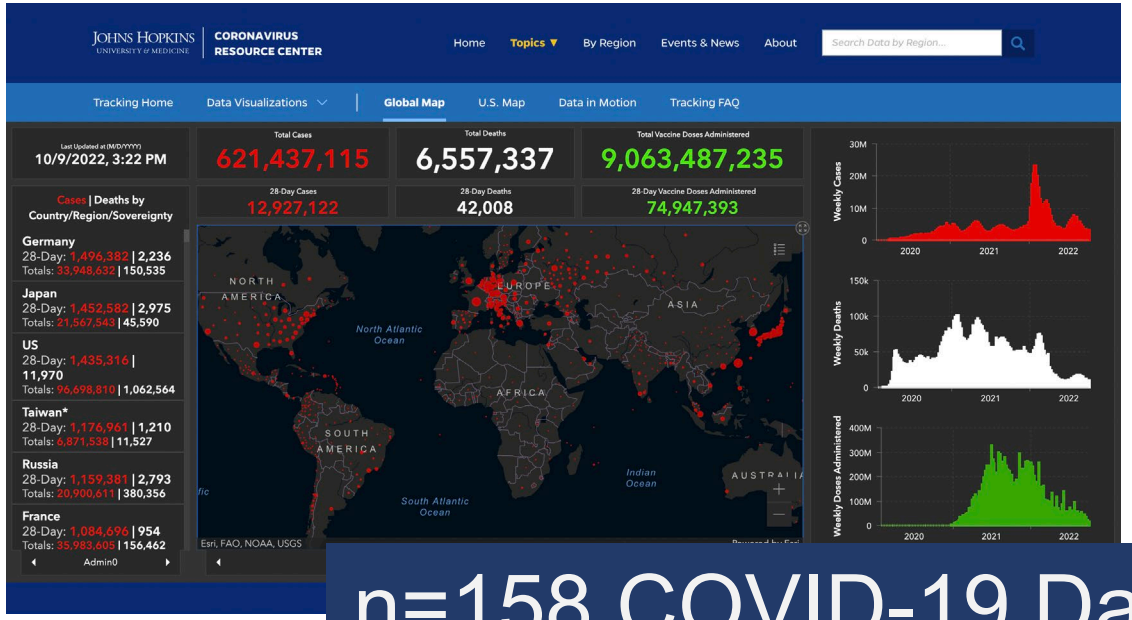
UNIVERSITY OF NORTH CAROLINA CHARLOTTE

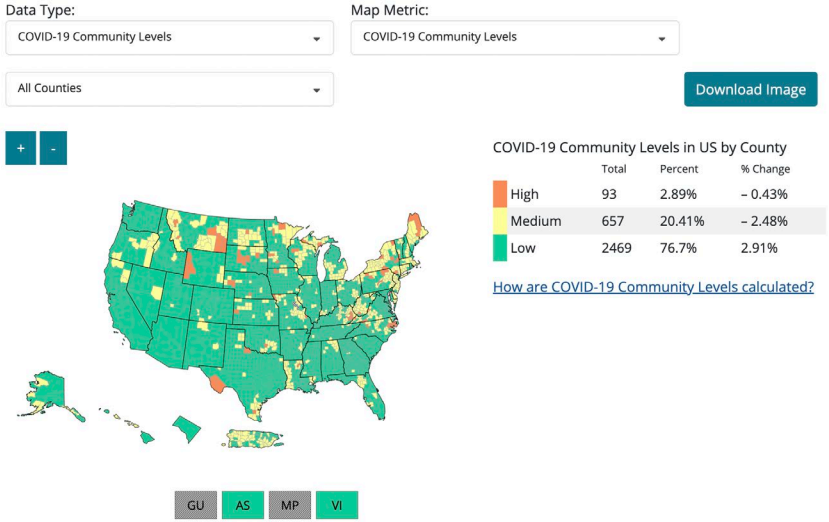
World Health Organization

Global Situation

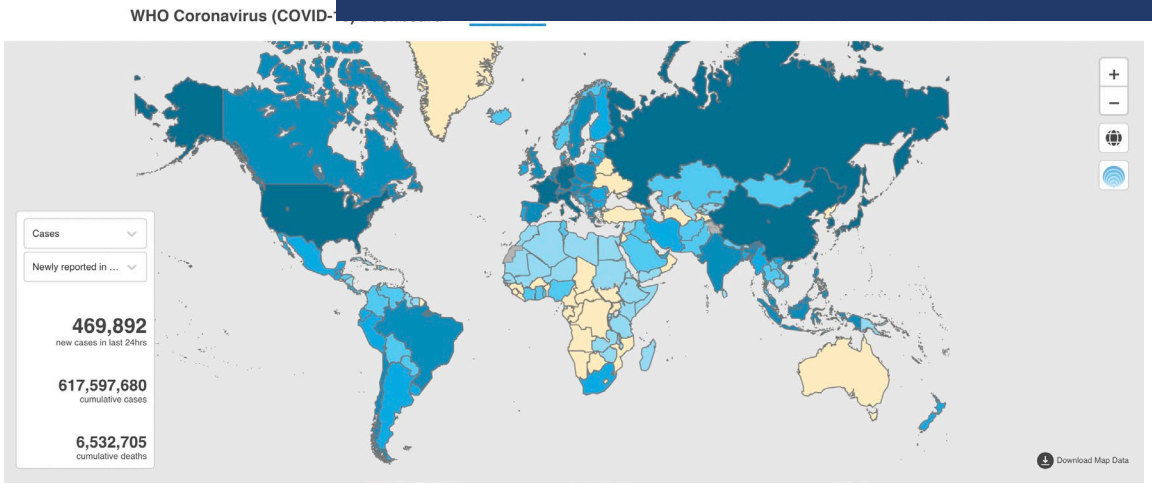


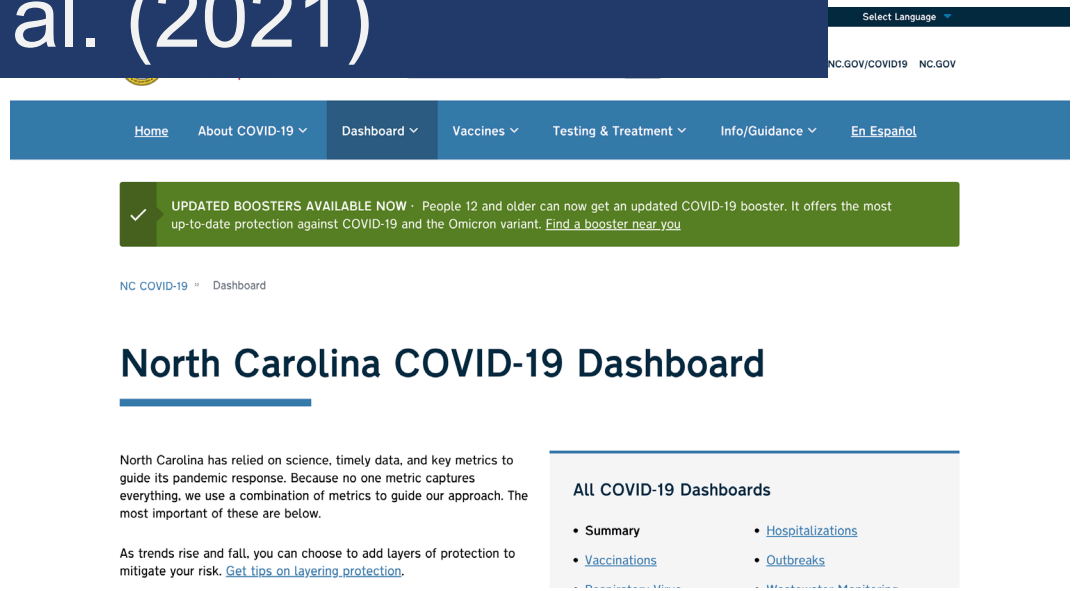
Globally, as of **5:52pm CEST, 26 October 2022**, there have been **625,740,449 confirmed cases** of COVID-19, including **6,563,667 deaths**, reported to WHO. As of **26 October 2022**, a total of **12,830,378,906 vaccine doses** have been administered.





n=158 COVID-19 Dashboards by July 2020, Ivanković et al. (2021)





Space-time Cluster Detection Methods in COVID-19 studies

Space-time scan statistic

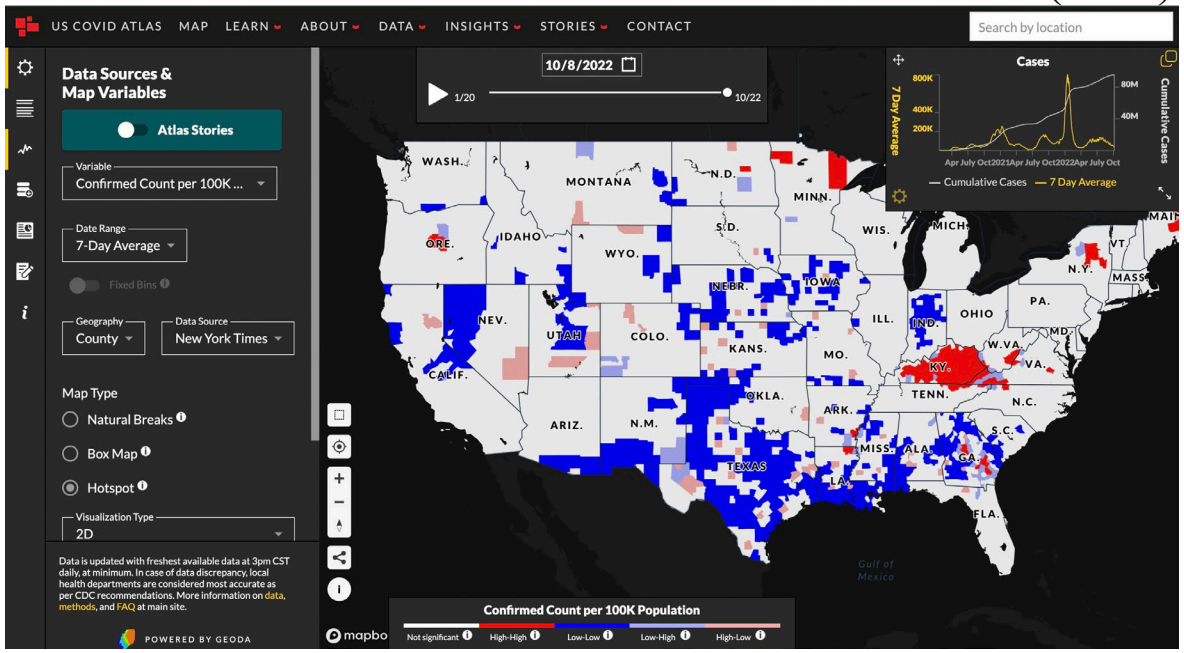
- US (Desjardins et al. 2020), Brazil (Gomes et al., 2020), Hong Kong (Kan et al., 2021), Bangladesh (Masrur et al., 2020), Spain (Rosillo et al., 2021), and so on.

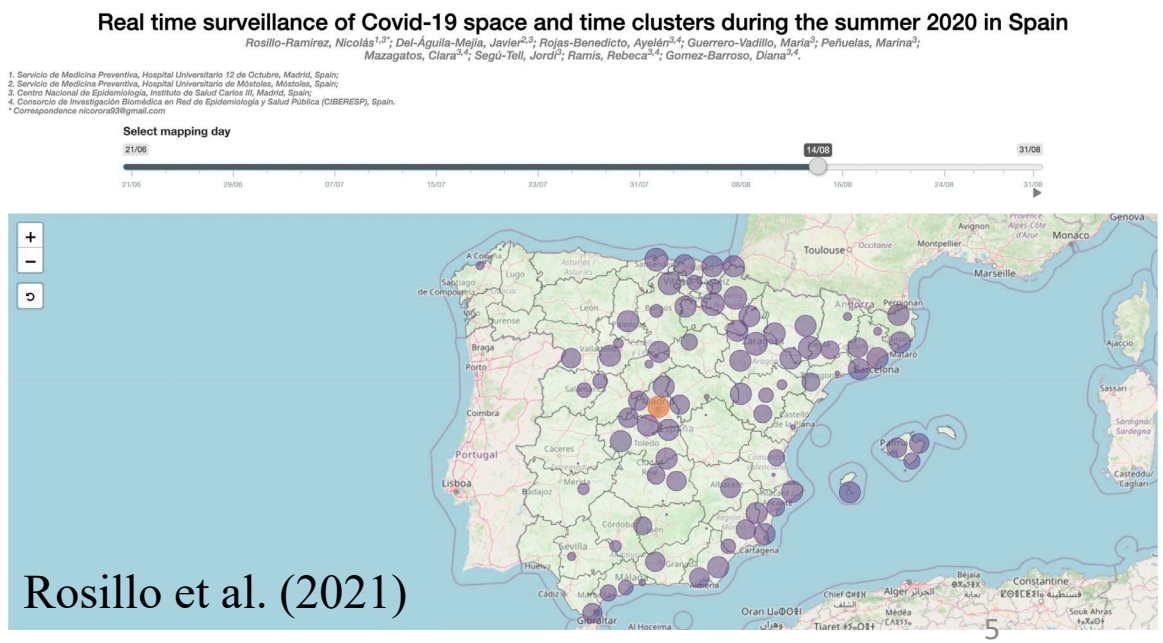
Local Indicators of Spatial Association (LISA) & Getis-Ord Gi* statistic

- Finland(Siljander et al. 2022)

Space-time cluster detection: web-based solutions

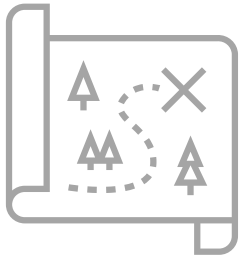
Kolak et al. (2021)





Gaps

Limited use of space-time detection methods



Limited visualization

Limited time representation

Our solution: an automatically, daily updated COVID-19 surveillance system (US COVID-19 YuTu)

use the prospective space-time scan statistics and LISA

at the county level in the continental US

objectives of this system

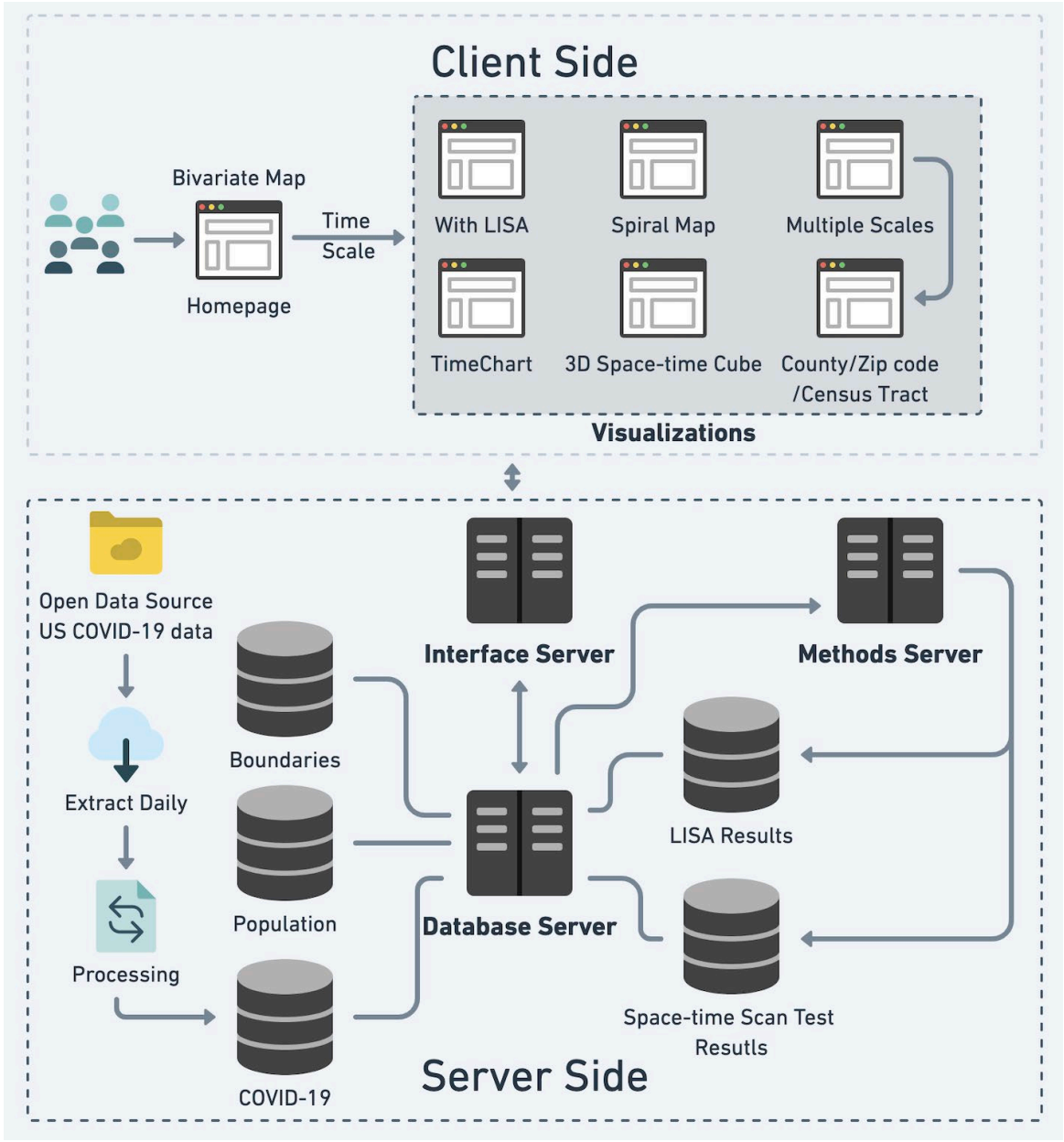
- to implement automatic space-time clustering detection
- to generate novel visual features of space-time clustering
- a tight coupling system that incorporates daily data updated and components for the objective one and two

Data

- JHU CSSE COVID-19 Data
- January 22nd, 2020 ~ current
- at the county level
- 7-day moving average of daily new cases
- 2019 Community Survey 5-year estimates
- the 2020 TIGER/Line shapefiles for boundaries

	A	B	C	D	E	F	G	H	I	J	K	L	M	XU	XV	XW
1	UID	iso2	iso3	code3	FIPS	Admin2	Province_	Country_F	Lat	Long_	Combinec	1/22/2020	1/23/2020	...	10/16/2021	10/17/2021
2	84001001	US	USA	840	1001	Autauga	Alabama	US	32.53953	-86.6441	Autauga, A	0	0	...	9893	9901
3	84001003	US	USA	840	1003	Baldwin	Alabama	US	30.72775	-87.7221	Baldwin, A	0	0	...	37069	37087
4	84001005	US	USA	840	1005	Barbour	Alabama	US	31.86826	-85.3871	Barbour, A	0	0	...	3554	3556
5	84001007	US	USA	840	1007	Bibb	Alabama	US	32.99642	-87.1251	Bibb, Alab	0	0	...	4216	4217
6	84001009	US	USA	840	1009	Blount	Alabama	US	33.98211	-86.5679	Blount, Al	0	0	...	10094	10102
7	84001011	US	USA	840	1011	Bullock	Alabama	US	32.10031	-85.7127	Bullock, A	0	0	...	1517	1517
8	84001013	US	USA	840	1013	Butler	Alabama	US	31.753	-86.6806	Butler, Ala	0	0	...	3247	3248
9	84001015	US	USA	840	1015	Calhoun	Alabama	US	33.77484	-85.8263	Calhoun, A	0	0	...	22142	22159
10	84001017	US	USA	840	1017	Chambers	Alabama	US	32.9136	-85.3907	Chambers	0	0	...	5559	5560
11	84001019	US	USA	840	1019	Cherokee	Alabama	US	34.17806	-85.6064	Cherokee	0	0	...	2999	3000

Framework

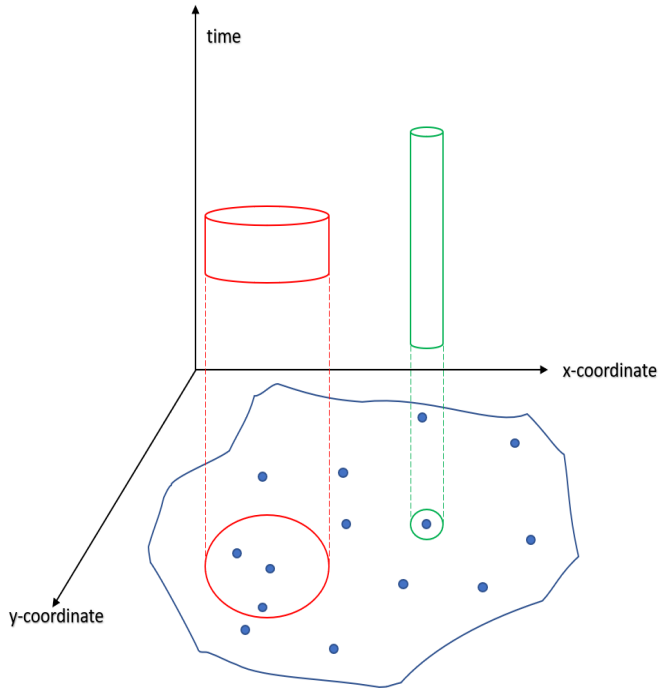

Client Side
Bivariate Map
Time
Scale
Homepage
With LISA
Spiral Map
Multiple Scales
TimeChart
3D Space-time Cube
County/Zip code
/Census Tract
Visualizations
Open Data Source
US COVID-19 data
Extract Daily
Processing
Boundaries
Population
COVID-19
Interface Server
Methods Server
Database Server
LISA Results
Space-time Scan Test
Resutls
Server Side

Analysis module- *Space-time scan statistics*

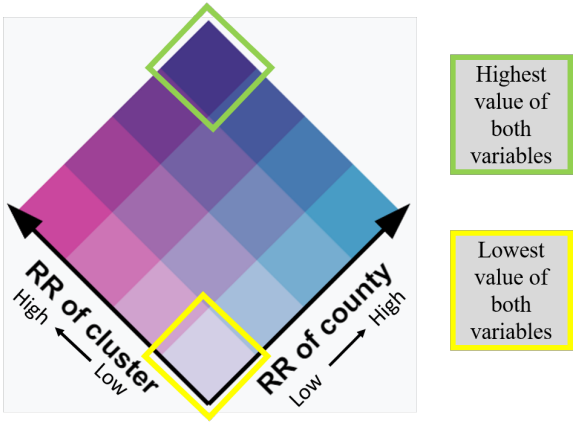
Hypothesis

H0: the probability of being a case within a cylinder Z (a) is same to the probability of being a case outside this cylinder(b).

Ha: the alternative hypothesis is that the probability of being a case within a cylinder Z is larger than the probability of being a case outside this cylinder (a > b).



Space-time scan statistics cont.



Formulas

- The expected number of cases (μ) $\mu = p \times \frac{N}{P}$

with p the population in the cylinder, N the total number of cases within the cylinder Z, and P the total population within the study area.

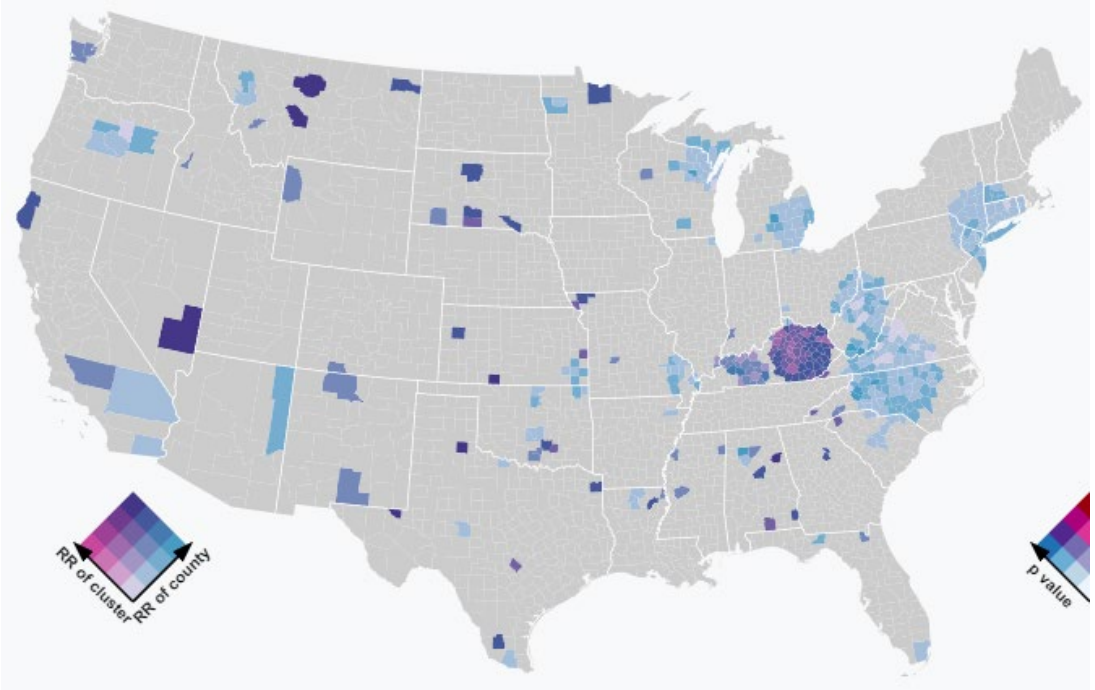
- The maximum likelihood ration to identify space-time clusters is defined as:

$$\frac{L(Z)}{L_0} = \frac{\left(\frac{n_Z}{\mu(Z)}\right)^{n_Z}\left(\frac{N-n_Z}{N-\mu(Z)}\right)^{N-n_Z}}{\left(\frac{N}{\mu(T)}\right)^{N}}$$

where $L(Z)$ is the likelihood function for the cylinder Z, and L_0 is the likelihood for the null hypothesis H0, n_Z is the number of cases in the cylinder Z, $\mu(Z)$ is the number of expected cases in cylinder Z, and $\mu(T)$ is the total number of expected cases within all time periods in the study area.

October 10th, 2022

The bivariate map



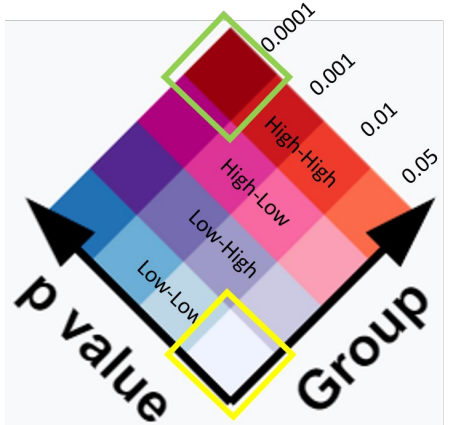
Analysis module- LISA

Local indicators of spatial association- LISA

- For region *i*, the local indicators of spatial association I_i defined as:

$$I_i = \frac{(n-1)(x_i - \bar{x})}{\sum_{j=1}^{n}(x_j - \bar{x})^2} \sum_{j=1}^{n} w_{ij}(x_j - \bar{x})$$

- where x_i is the attribute of the variable of interest (here, the disease rate) in region *i*, $\bar{x}$ is the mean of x_i (*i*=1…, n), w_{ij} is the spatial weight between regions *i* and *j*.

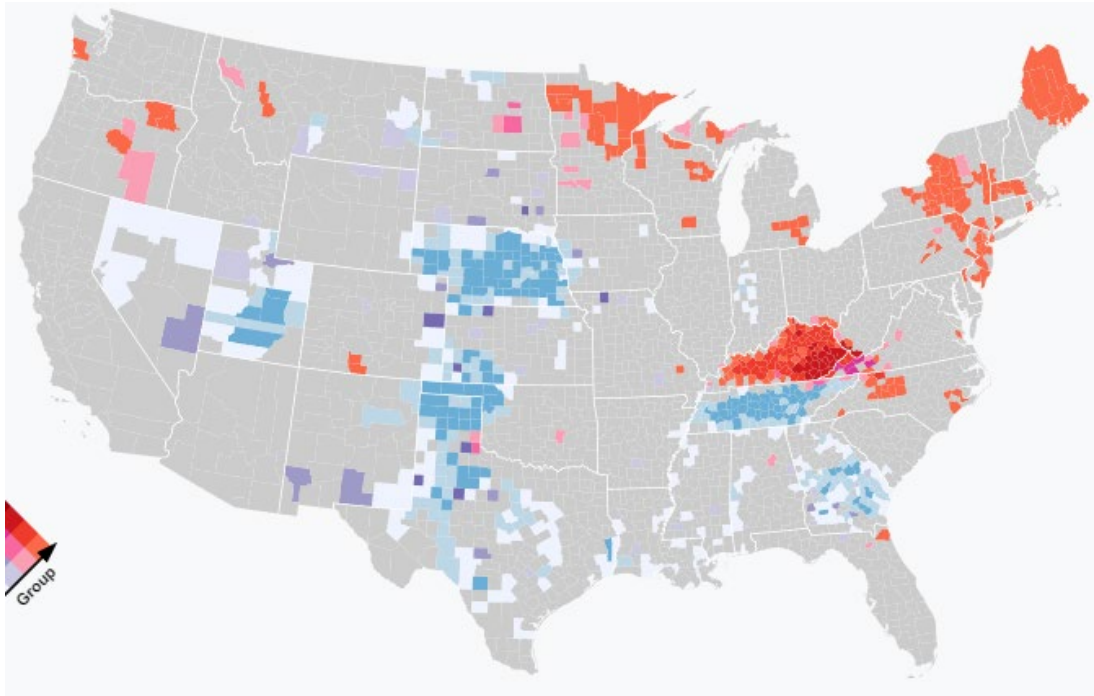


High-High
0.0001

Low-Low
0.05

October 10th, 2022

The bivariate map



Analysis module- Analysis procedure

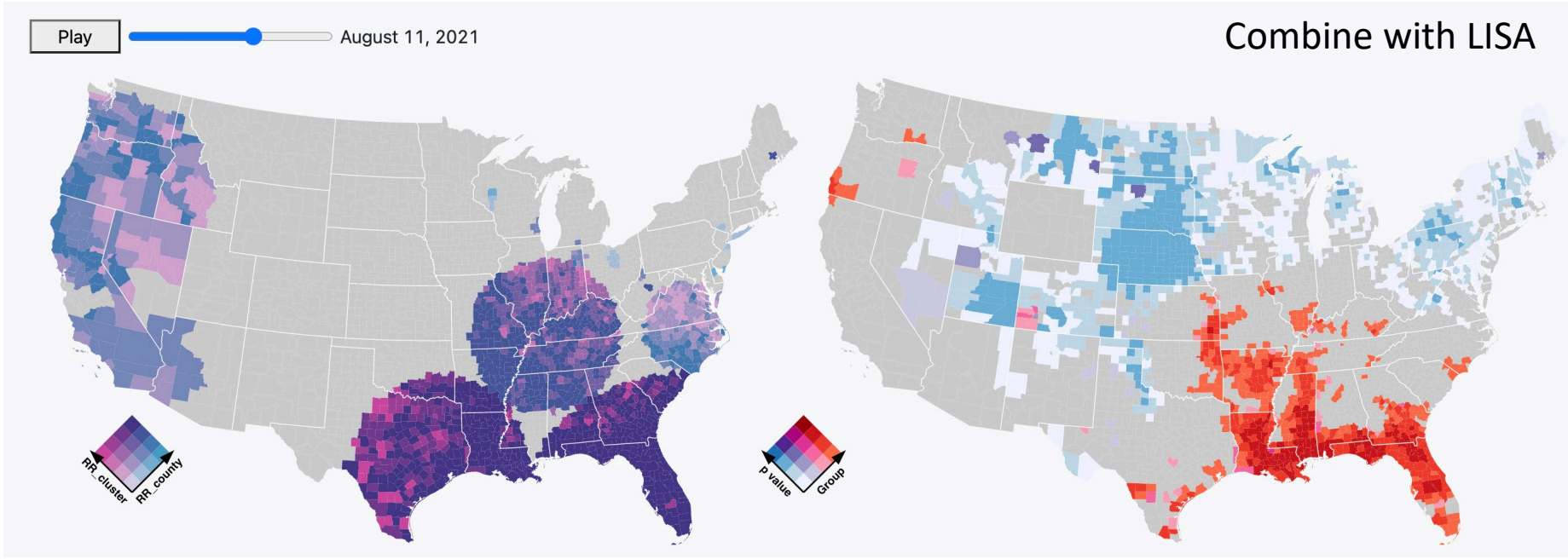
Space-time scan statistics

- prospective space-time Poisson discrete scan statistics
- the maximum spatial cluster size as 50% of the population at risk
- the maximum temporal cluster size as 50 days
- SaTScan
- true space-time clusters

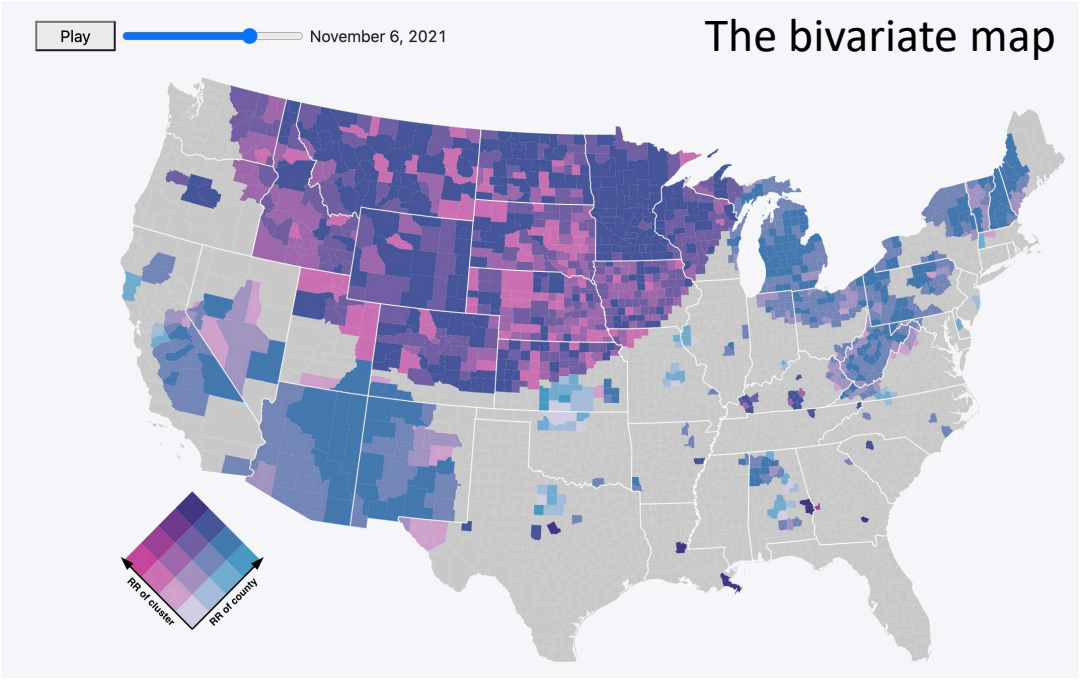
LISA

- cases per day
- incidence rate (7-day average cases divided by population)
- Queen contiguity weights
- repeat every day
- pygeoda
- not-true space-time clusters

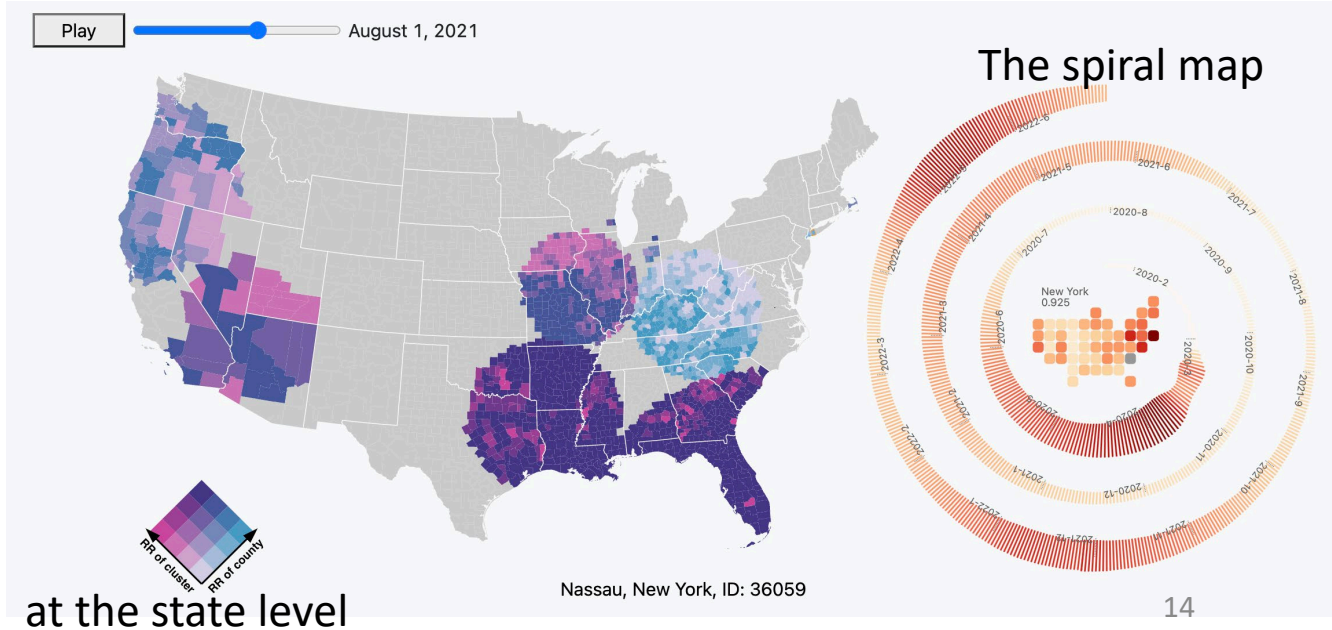
Visualization module



Combine with LISA



The bivariate map

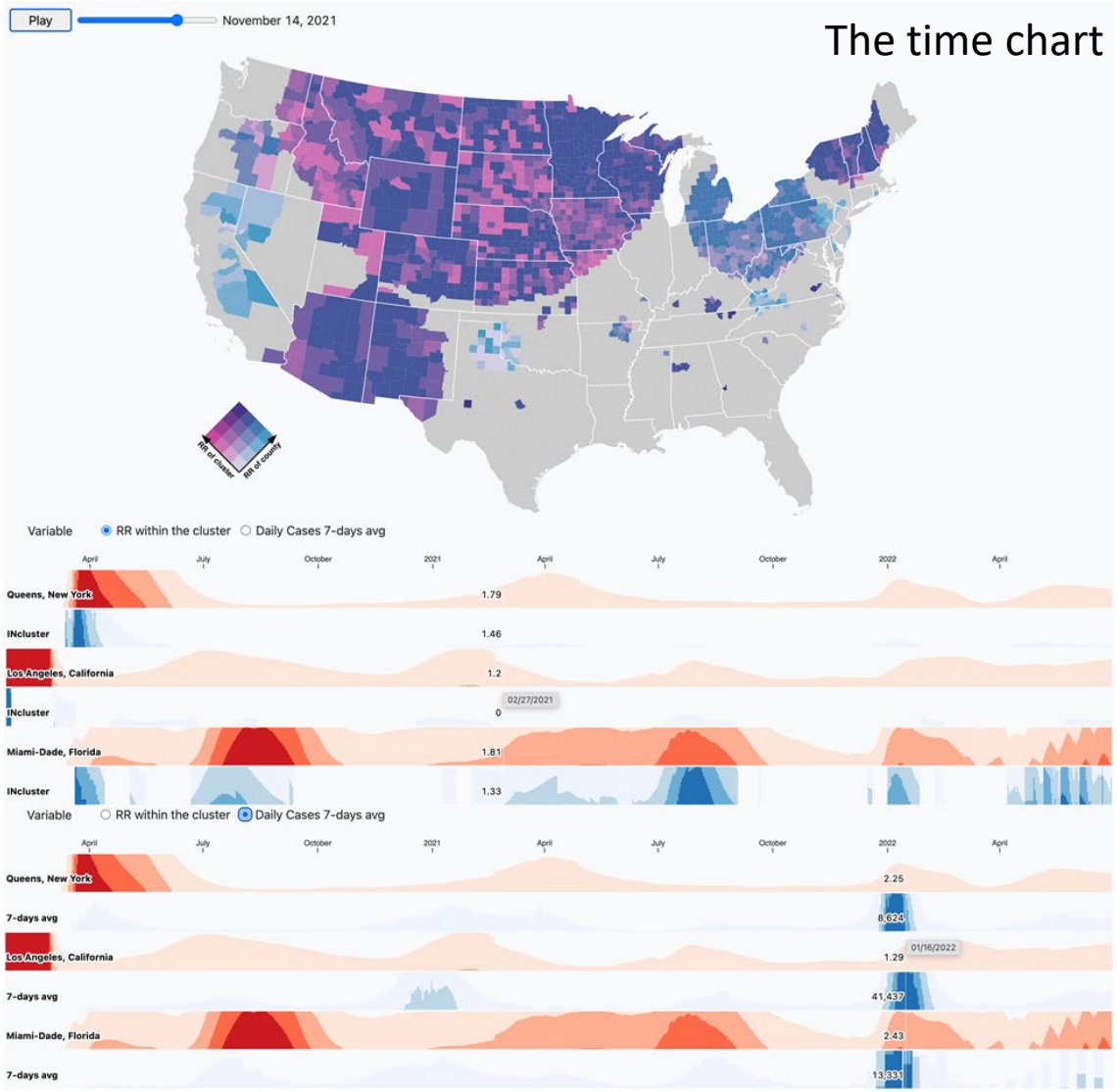


The spiral map

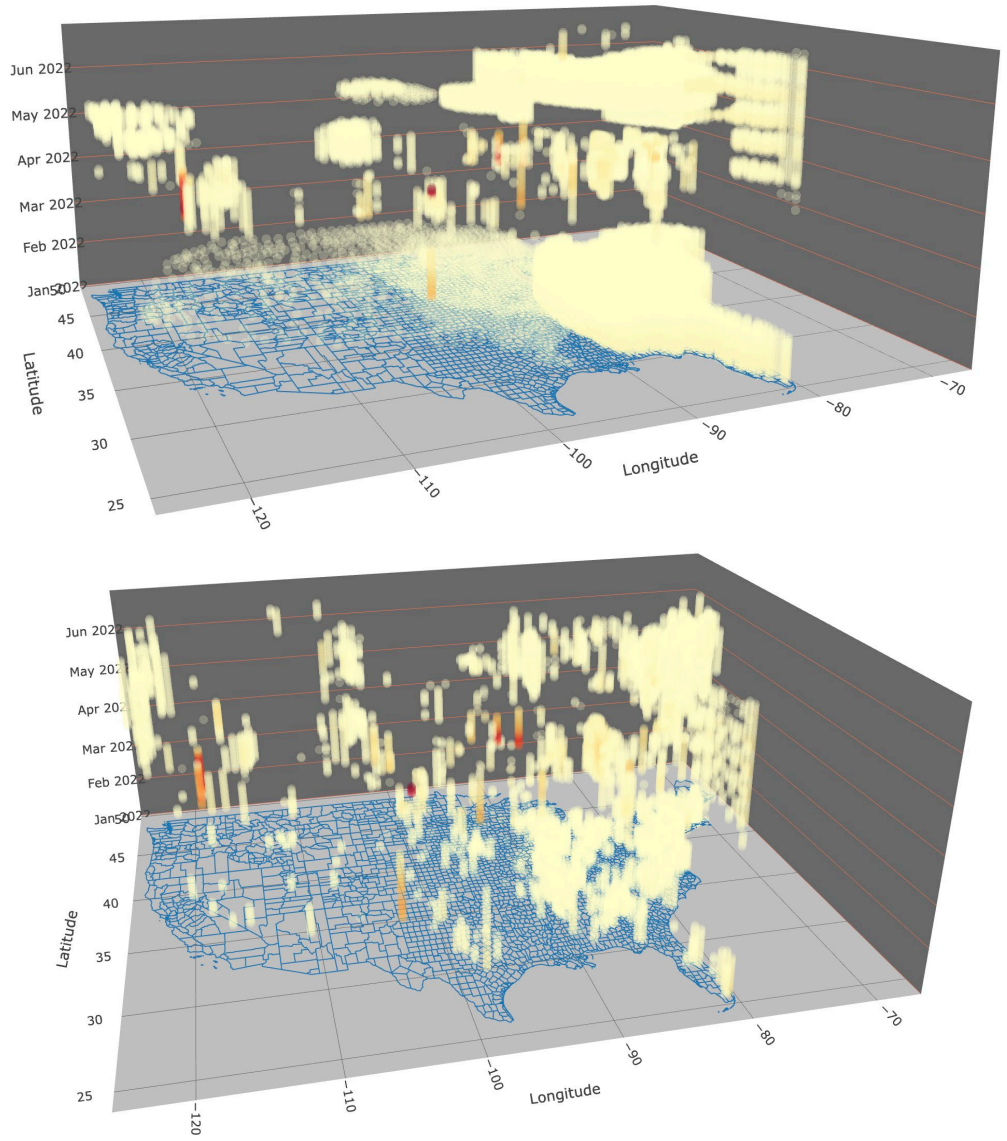
at the state level

Visualization module cont.

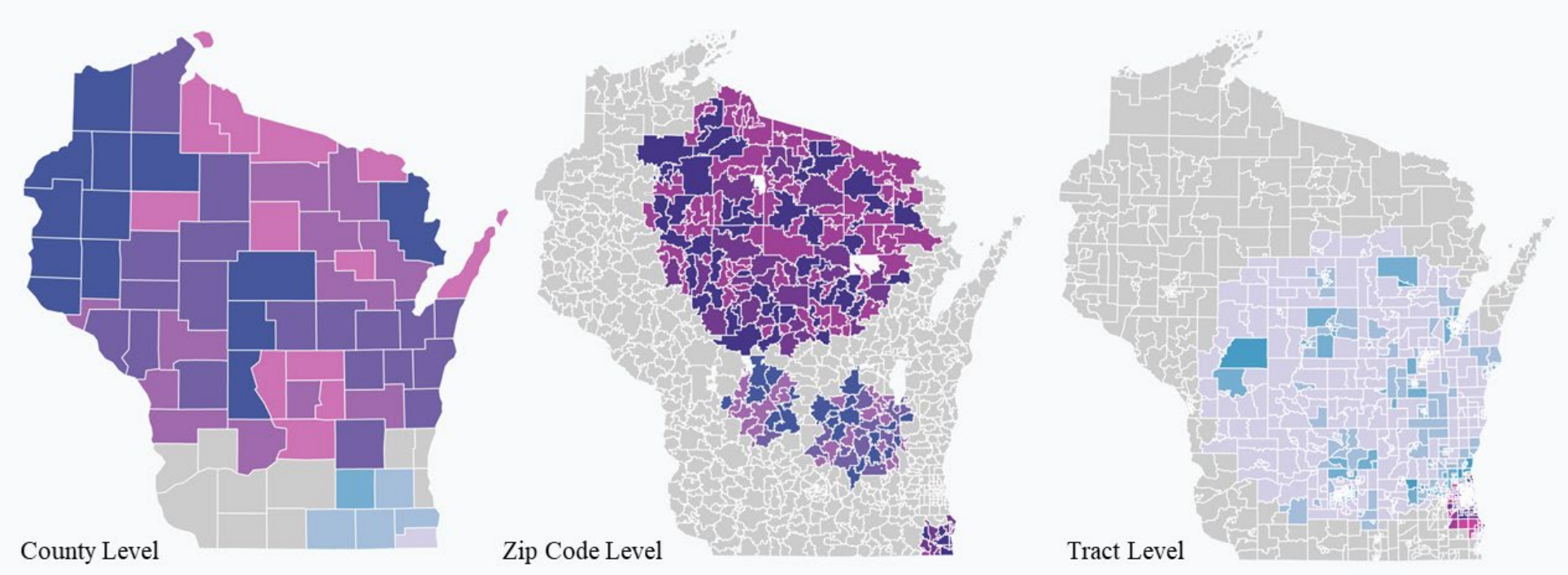
The time chart



The 3D space-time cube

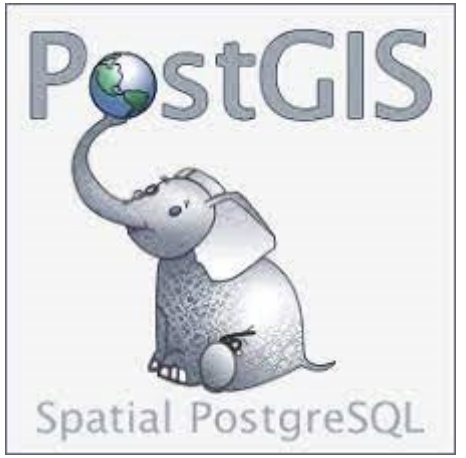


Visualization module cont.
-different scales



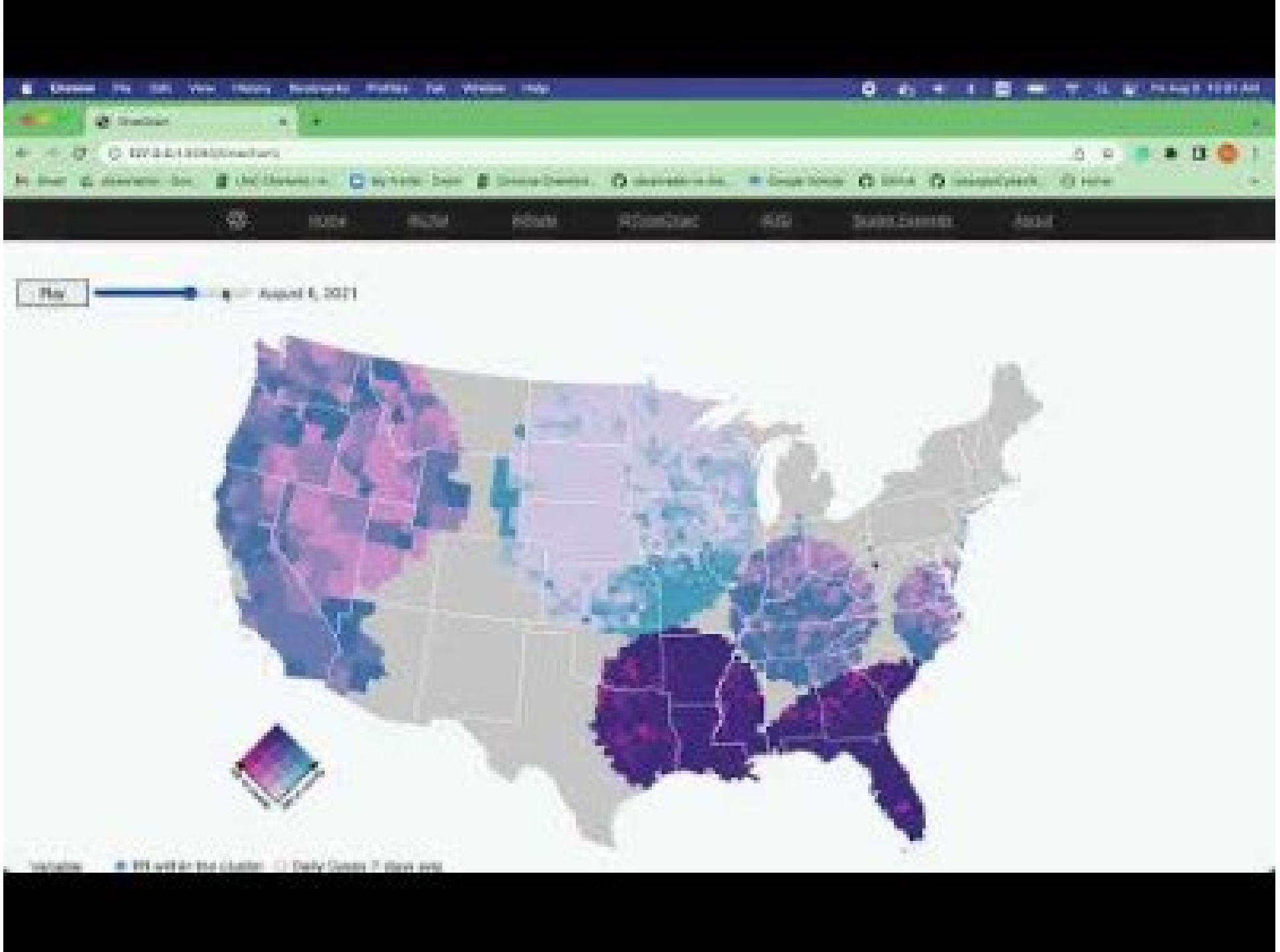
Data processing module

- PostGIS
- Python
- Shell
- D3
- Observable
- Docker
- DigitalOcean

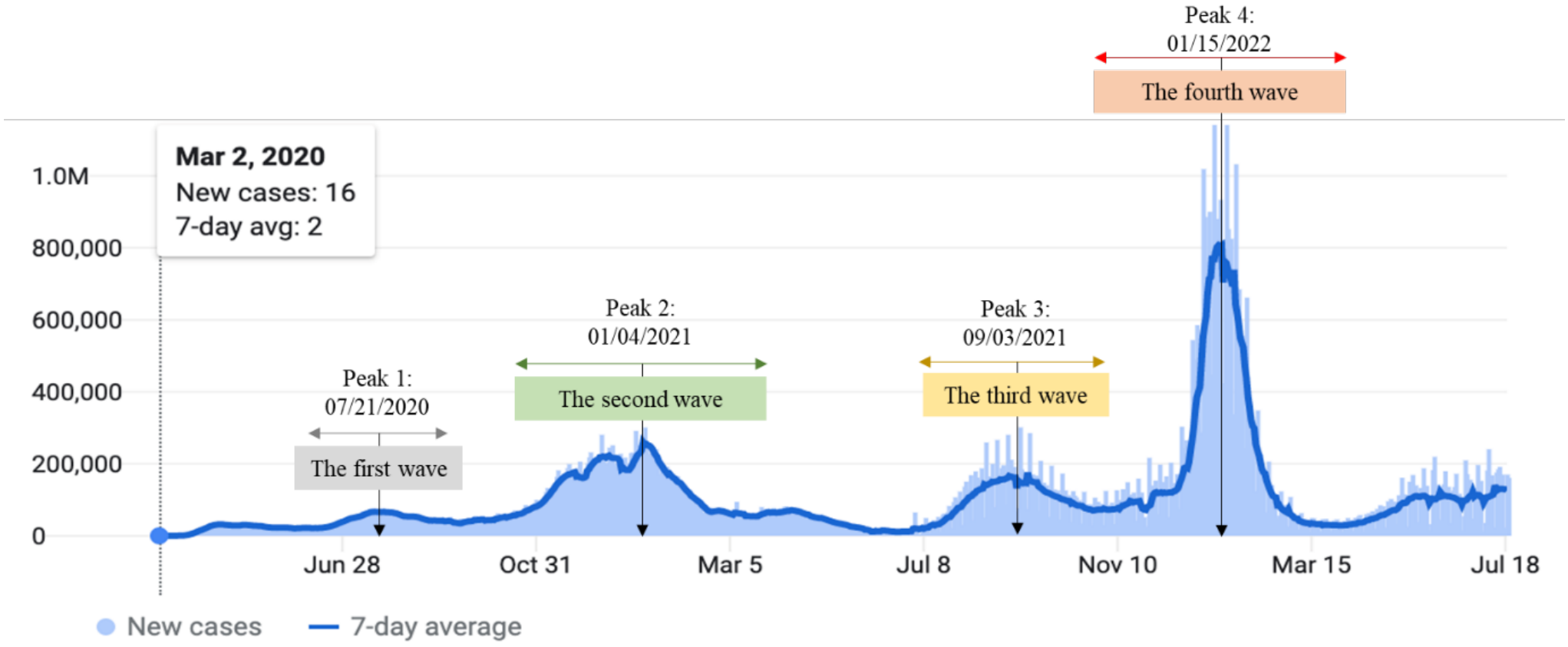


US COVID-19 YuTu: A Daily Surveillance System

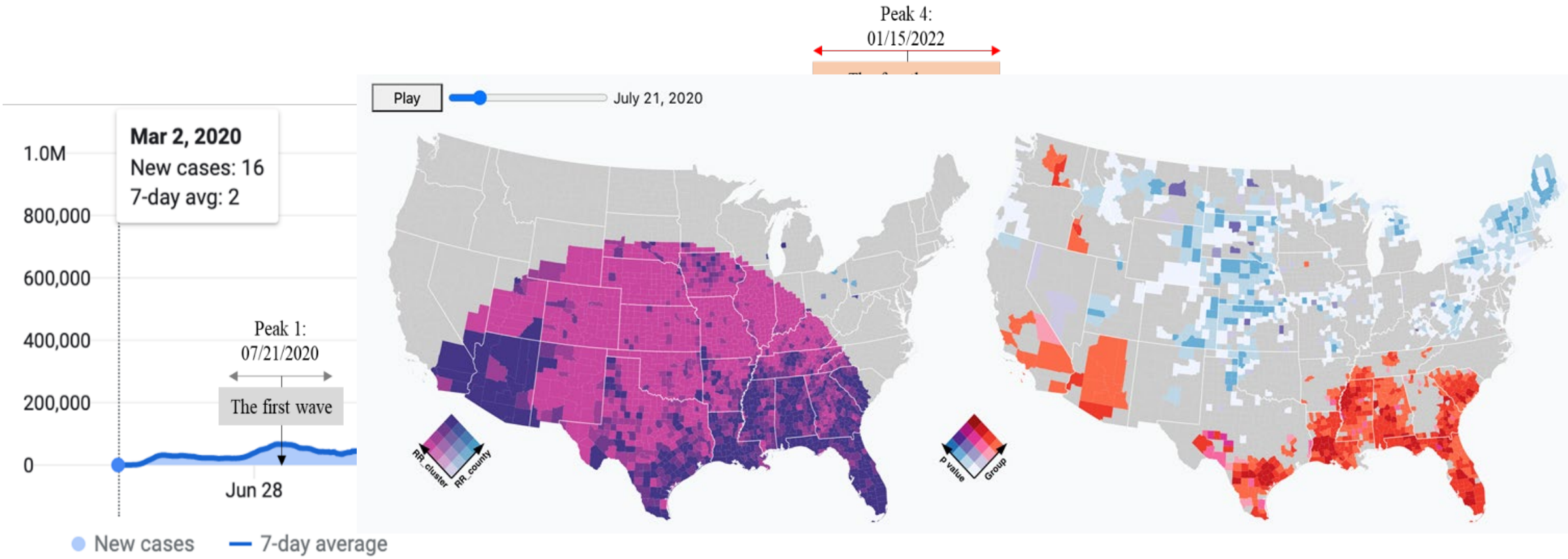
http://159.223.164.41/app/


Play
August 1, 2021

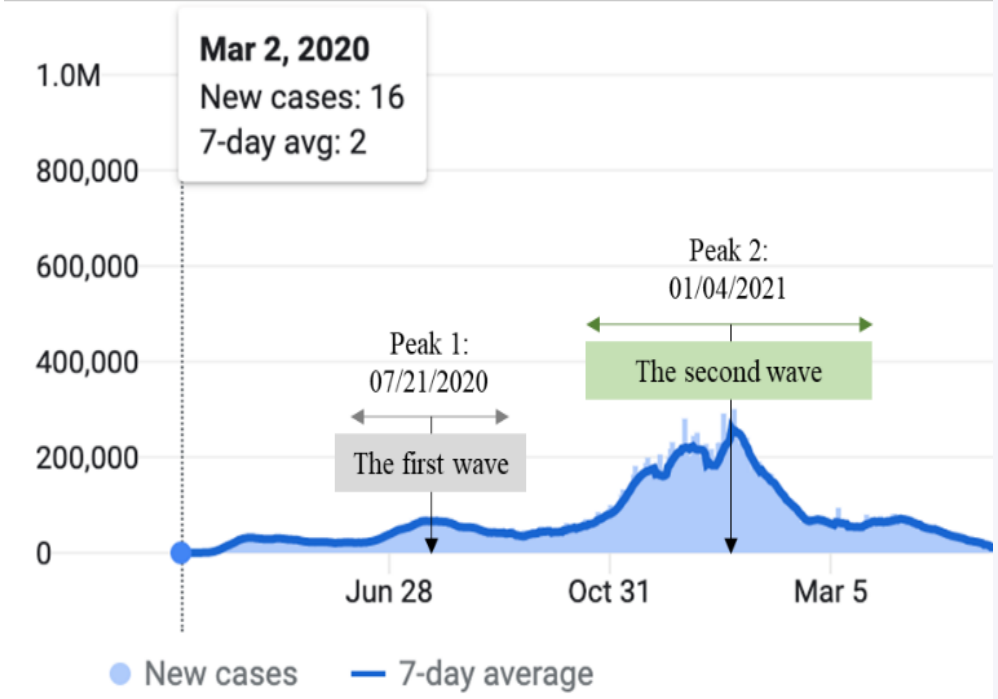
Case study- compare peaks

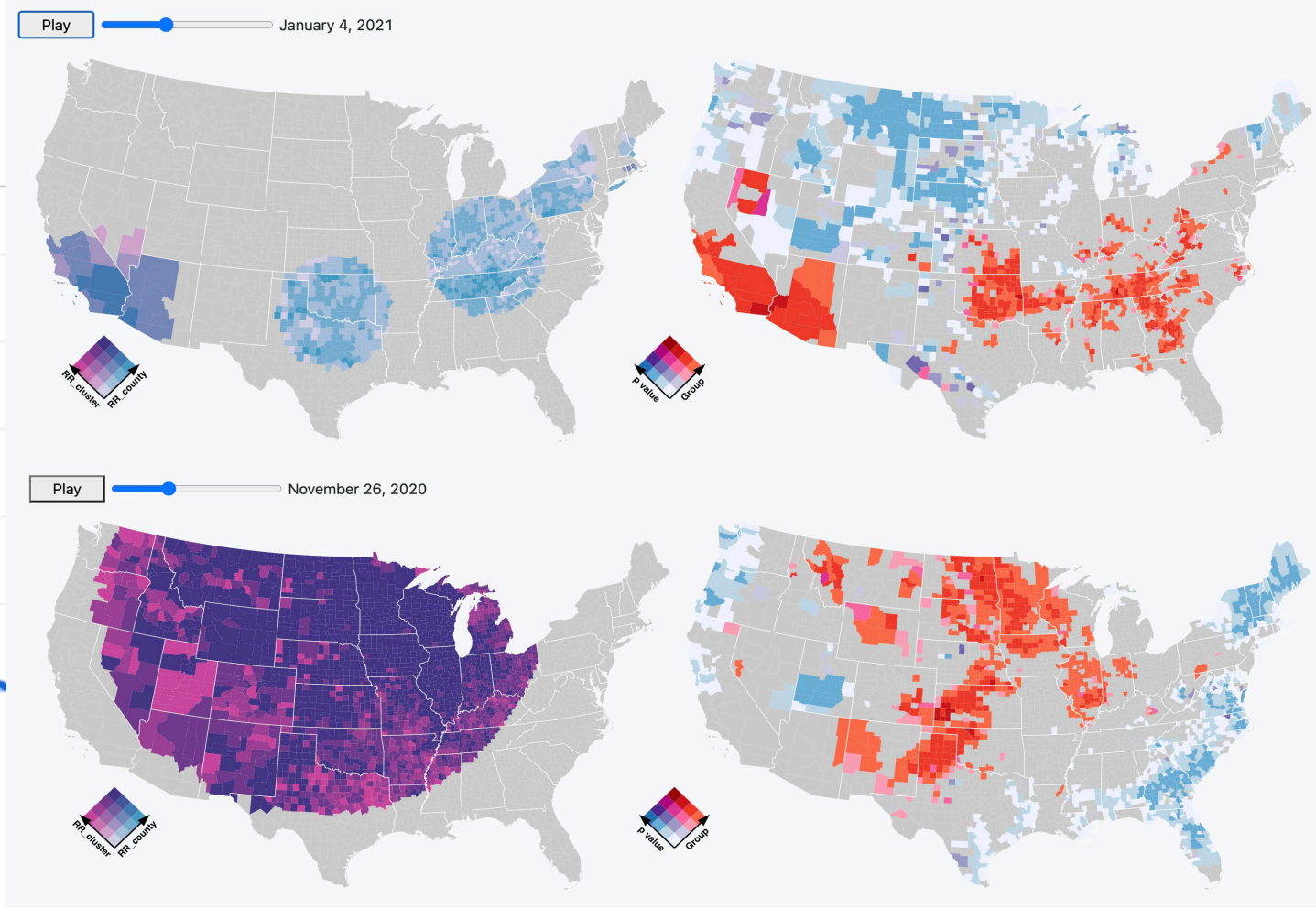


Peak 1

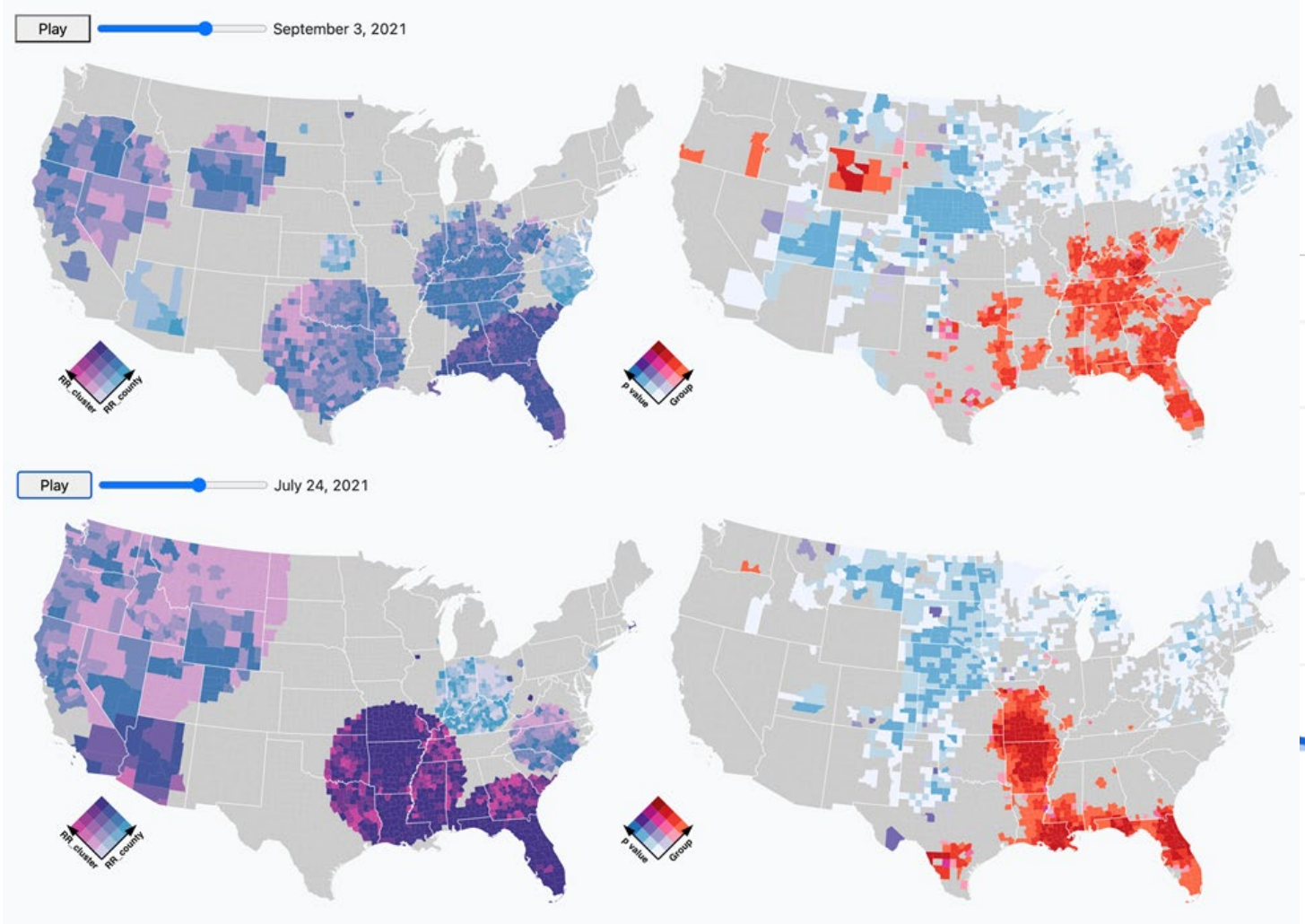

Peak 4:
01/15/2022
Play
July 21, 2020
Mar 2, 2020
New cases: 16
7-day avg: 2
1.0M
800,000
600,000
400,000
200,000
0
Peak 1:
07/21/2020
The first wave
Jun 28
RR_cluster
RR_county
p value
Group
New cases
7-day average

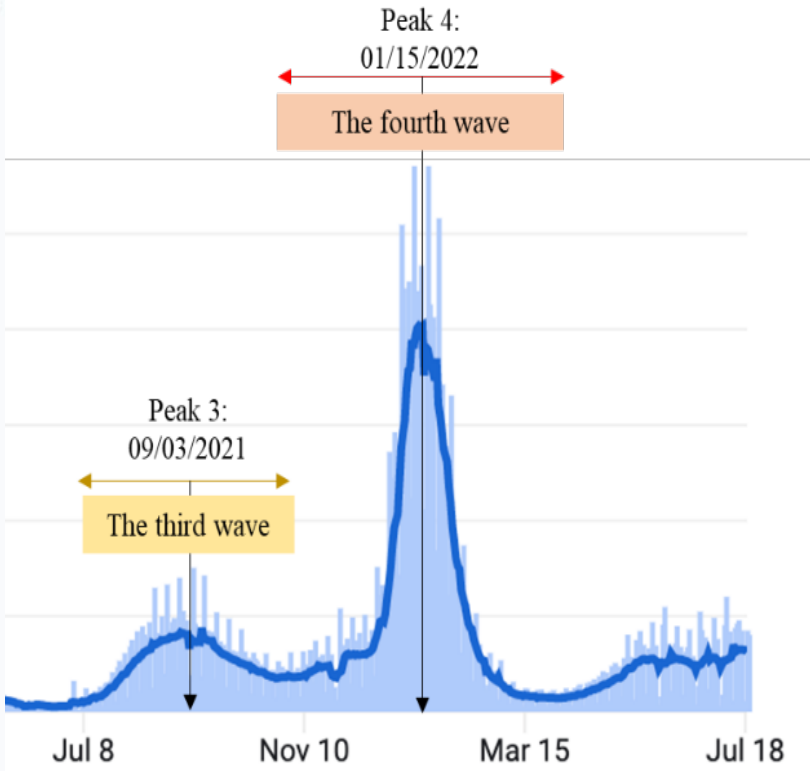
Peak 2


Mar 2, 2020
New cases: 16
7-day avg: 2
1.0M
800,000
600,000
400,000
200,000
0
Peak 1:
07/21/2020
The first wave
Peak 2:
01/04/2021
The second wave
Jun 28
Oct 31
Mar 5
New cases
7-day average

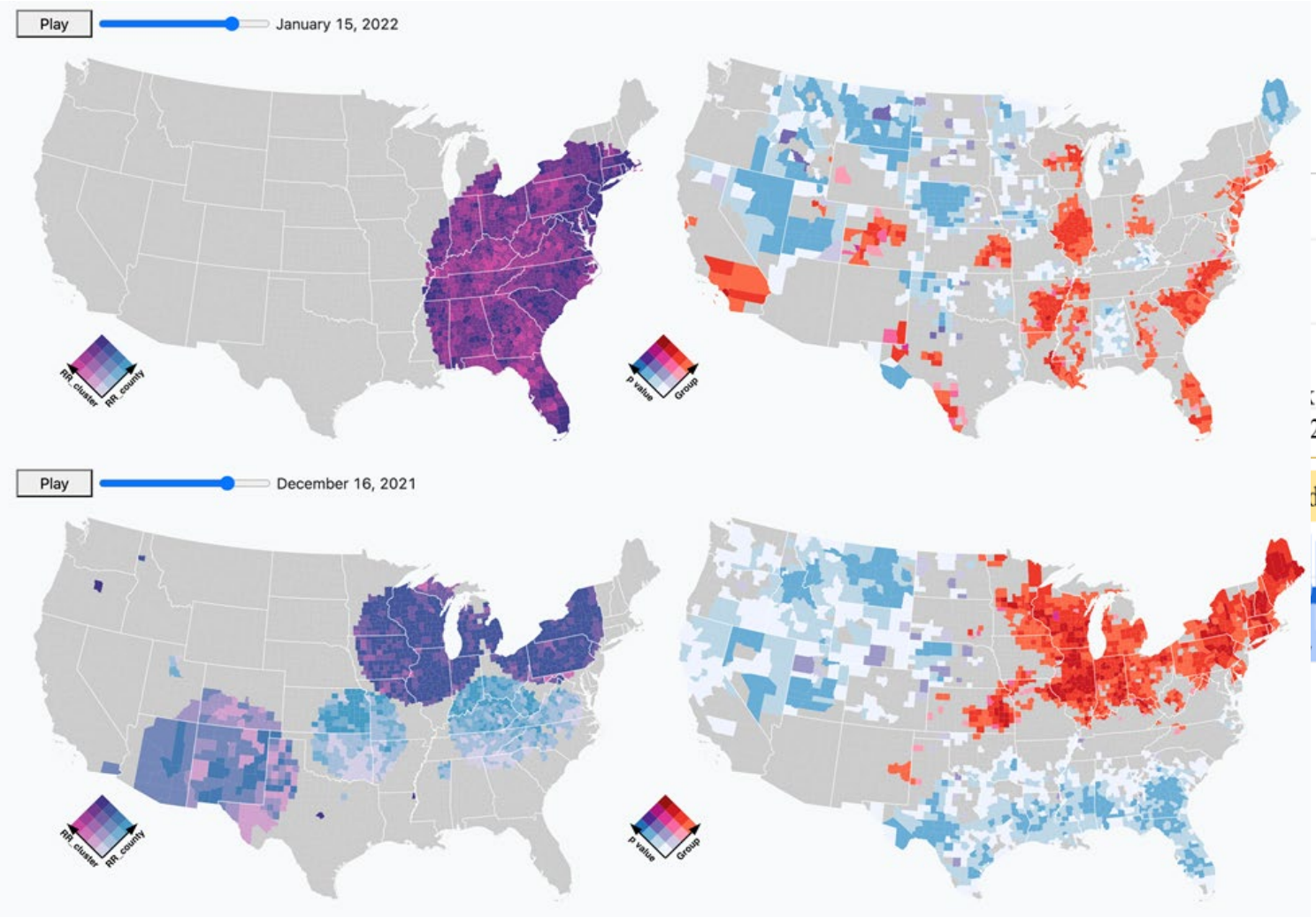

Play
January 4, 2021
RR_cluster
RR_county
p value
Group
Play
November 26, 2020
RR_cluster
RR_county
p value
Group

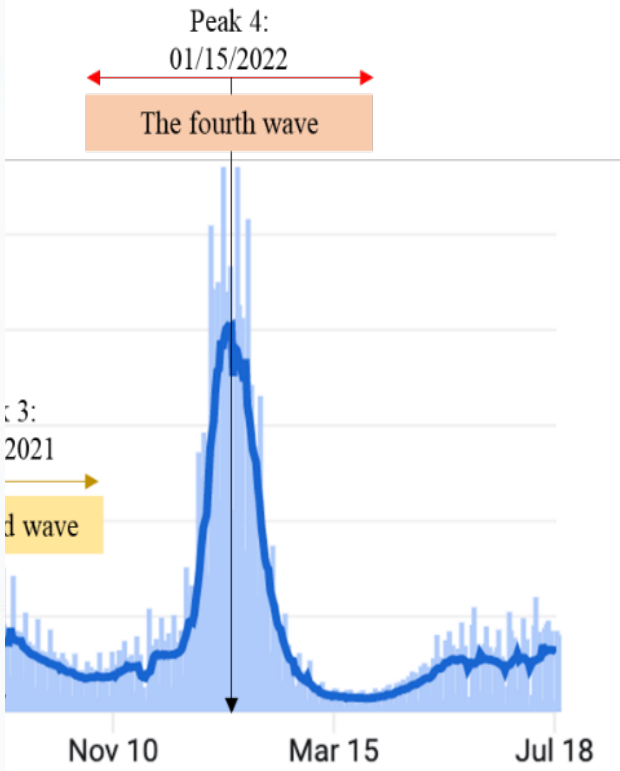
Peak 3


Play
September 3, 2021
RR_cluster
RR_county
p value
Group
Play
July 24, 2021
RR_cluster
RR_county
p value
Group

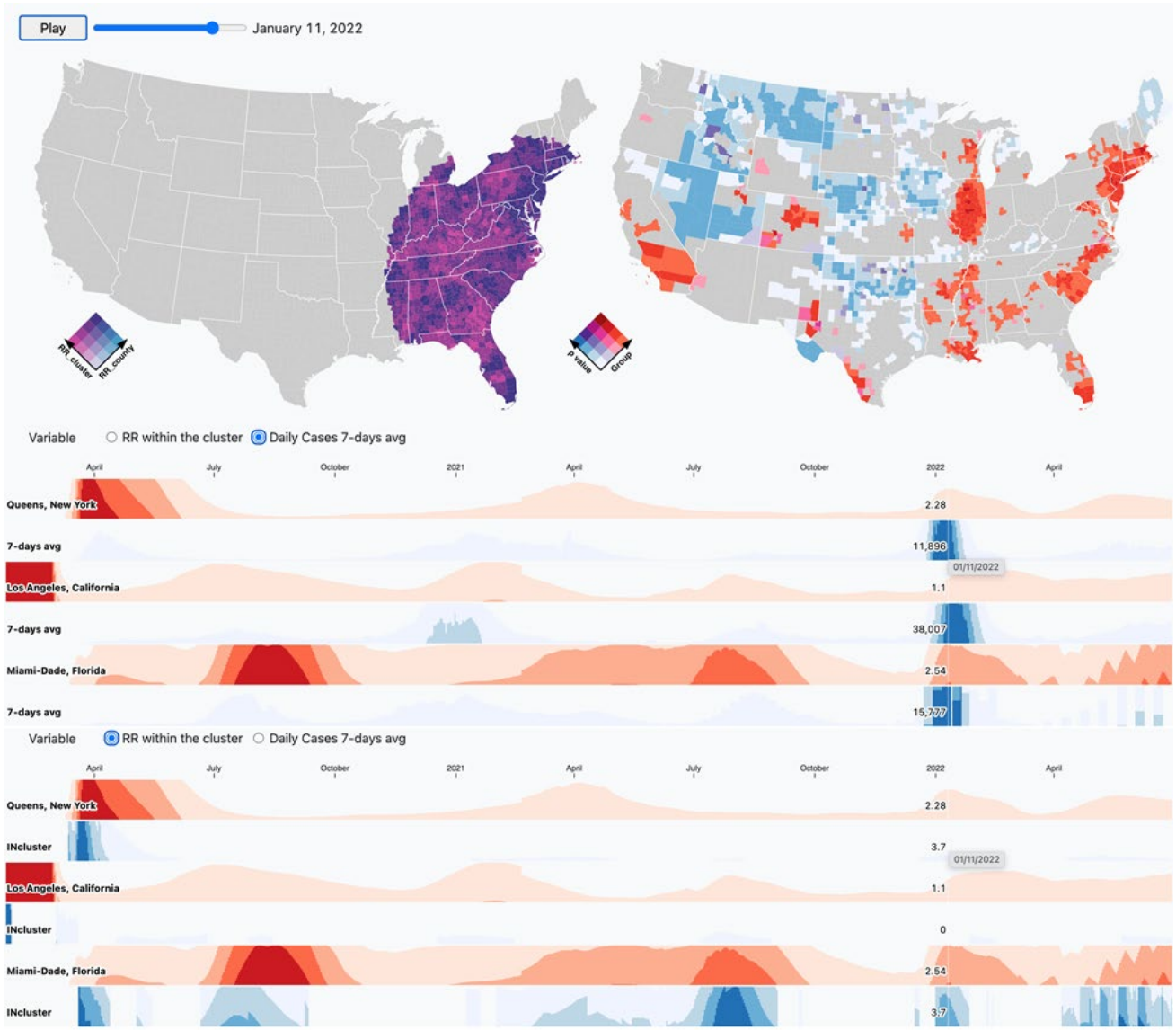

Peak 4:
01/15/2022
The fourth wave
Peak 3:
09/03/2021
The third wave
Jul 8
Nov 10
Mar 15
Jul 18

Peak 4

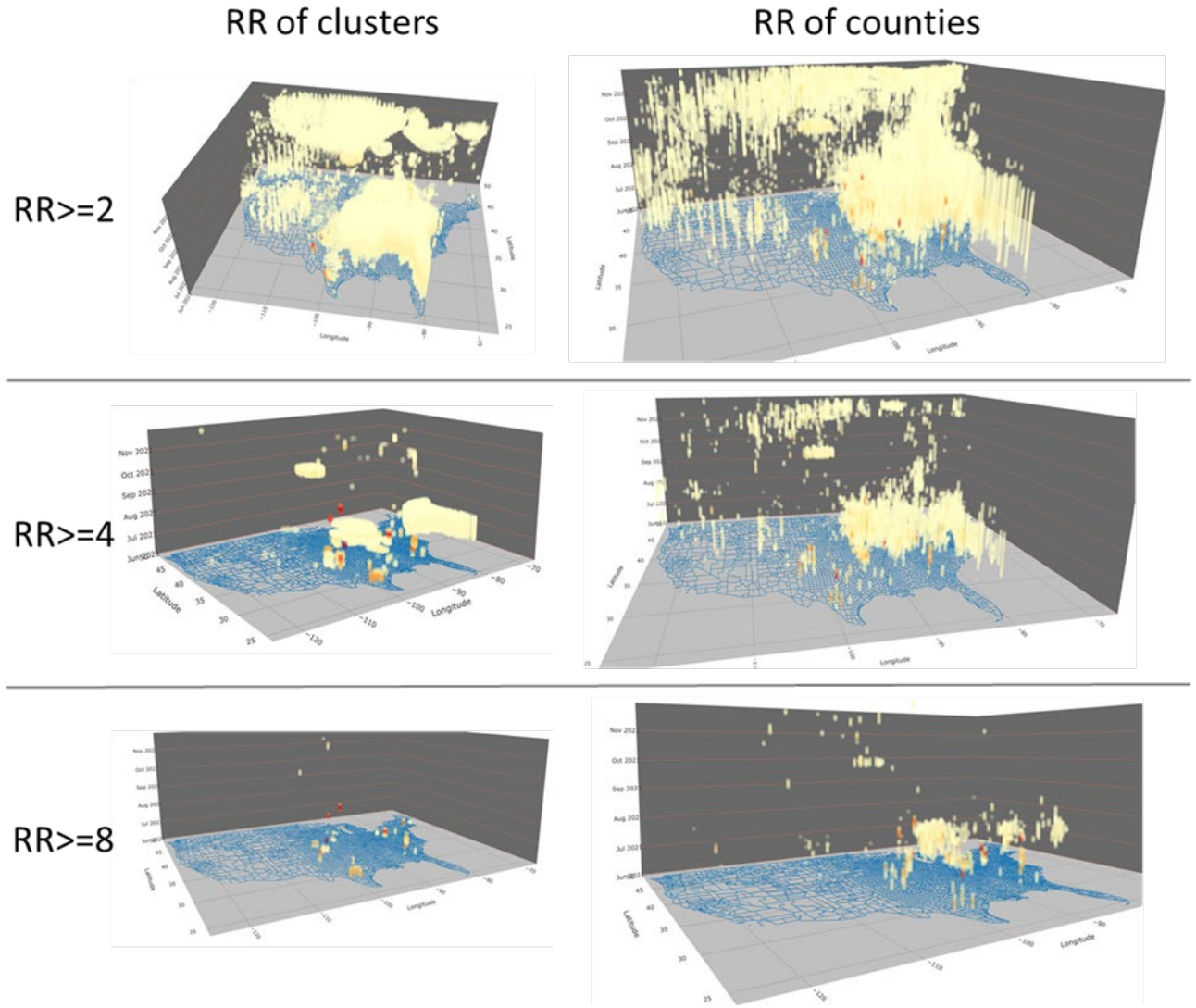

Play
January 15, 2022
RR_cluster
RR_county
p value
Group
Play
December 16, 2021
RR_cluster
RR_county
p value
Group


Peak 4:
01/15/2022
The fourth wave
Nov 10
Mar 15
Jul 18

Case study- compare counties



Case study- look at one wave (the third wave)



Discussion and conclusion

- space-time cluster detection analysis and visualization
- near real-time monitoring
- multiple scales
- novel visual features
- portable to other methods and regions

Future works

need to be evaluated (working on)

setting parameters of analysis methods requires further collaboration with epidemiologists

use other information, like hospitalization rate

References

- Desjardins, M. R., Hohl, A., & Delmelle, E. M. (2020). Rapid surveillance of COVID-19 in the United States using a prospective space-time scan statistic: Detecting and evaluating emerging clusters. *Applied geography*, *118*, 102202.
- Gomes, D.-S., Andrade, L. A., Ribeiro, C. J. N., Peixoto, M., Lima, S., Duque, A., Cirilo, T. M., Góes, M., Lima, A., & Santos, M. (2020). Risk clusters of COVID-19 transmission in northeastern Brazil: prospective space–time modelling. *Epidemiology & Infection*, *148*.
- Hohl, A., Delmelle, E. M., Desjardins, M. R., & Lan, Y. (2020). Daily surveillance of COVID-19 using the prospective space-time scan statistic in the United States. *Spatial and spatio-temporal epidemiology*, *34*, 100354.
- Ivanković, D., Barbazza, E., Bos, V., Fernandes, Ó. B., Gilmore, K. J., Jansen, T., Kara, P., Larrain, N., Lu, S., & Meza-Torres, B. (2021). Features constituting actionable COVID-19 dashboards: descriptive assessment and expert appraisal of 158 public web-based COVID-19 dashboards. Journal of medical Internet research, 23(2), e25682.
- Kan, Z., Kwan, M. P., Huang, J., Wong, M. S., & Liu, D. (2021). Comparing the space-time patterns of high-risk areas in different waves of COVID-19 in Hong Kong. *Transactions in GIS*, *25*(6), 2982-3001.
- Kolak, M., Li, X., Lin, Q., Wang, R., Menghaney, M., Yang, S., & Anguiano Jr, V. (2021). The US COVID Atlas: A dynamic cyberinfrastructure surveillance system for interactive exploration of the pandemic. *Transactions in GIS*, *25*(4), 1741-1765.
- Masrur, A., Yu, M., Luo, W., & Dewan, A. (2020). Space-Time Patterns, Change, and Propagation of COVID-19 Risk Relative to the Intervention Scenarios in Bangladesh. *International Journal of Environmental Research and Public Health*, *17*(16), 5911. https://www.mdpi.com/1660-4601/17/16/5911
- Rosillo, N., Del-Águila-Mejía, J., Rojas-Benedicto, A., Guerrero-Vadillo, M., Peñuelas, M., Mazagatos, C., Segú-Tell, J., Ramis, R., & Gómez-Barroso, D. (2021). Real time surveillance of COVID-19 space and time clusters during the summer 2020 in Spain. *BMC Public Health*, *21*(1), 961. https://doi.org/10.1186/s12889-021-10961-z
- Siljander, M., Uusitalo, R., Pellikka, P., Isosomppi, S., & Vapalahti, O. (2022). Spatiotemporal clustering patterns and sociodemographic determinants of COVID-19 (SARS-CoV-2) infections in Helsinki, Finland. *Spatial and spatio-temporal epidemiology*, *41*, 100493.

Thank you!

ylan1@uncc.edu