

Space-time analysis of US COVID-19 data

Lee De Cola
LDECOLA@COMCAST.NET

2020-07-28



ISSUES

Coronavirus Stats & Data

Dive into the stats and data on all things COVID-19. Explore metrics including daily case counts, government spending to help people and businesses, health risk factors, and the virus's economic impact.



QUESTIONS TO ASK

Where has the virus spread to since the first confirmed case in the United States in January 2020?

Where can I get up-to-date data about coronavirus in my community?

COVID-19 Impact and Recovery

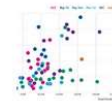
A daily resource of key factors affecting health and the economy

[View the Impact](#) →

RELATED ARTICLES

CORONAVIRUS
EDUCATION

Coronavirus and college football: what might this mean for colleges' bottom lines?



[usafactsstatic.blob.core.windows.net/
public/data/covid-19/
covid_confirmed_usafacts.csv](https://usafactsstatic.blob.core.windows.net/public/data/covid-19/covid_confirmed_usafacts.csv)

Keep up with the latest data and most popular content.

Enter your email address



VISUALIZATION

Coronavirus in the United States

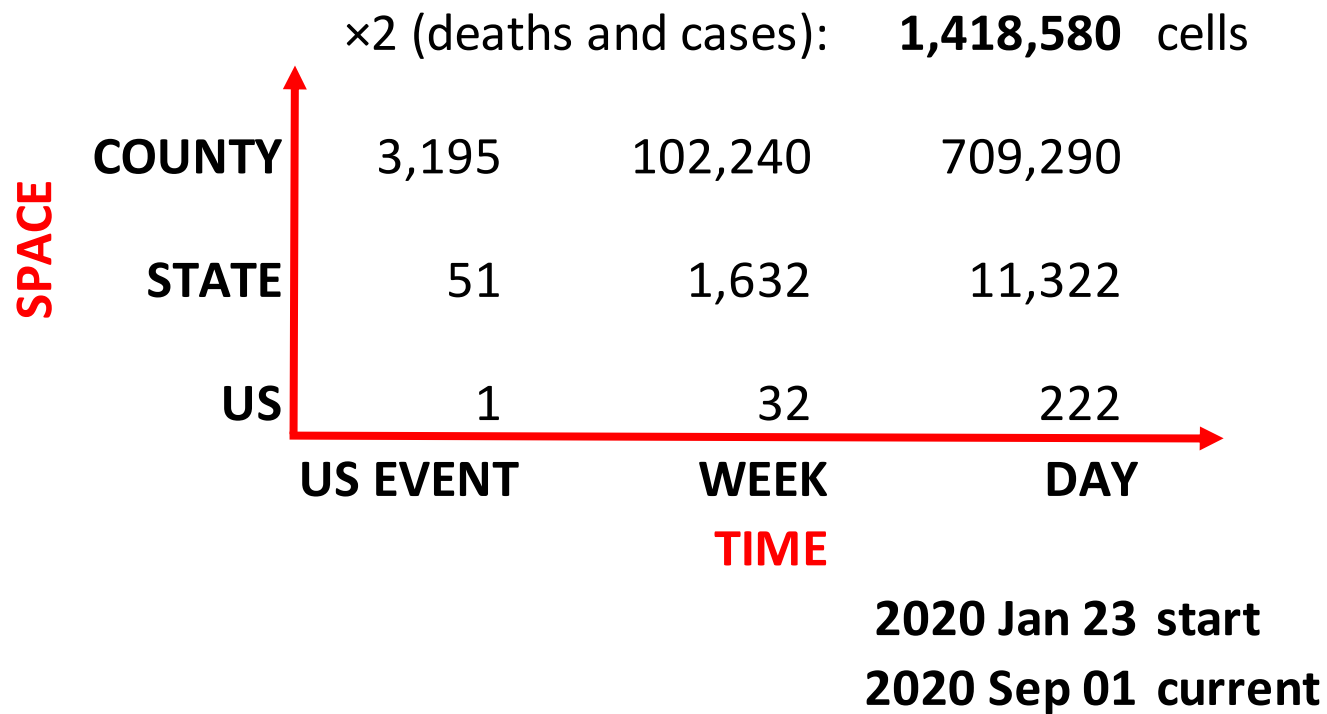
See how the outbreak has spread.

ELECTIONS

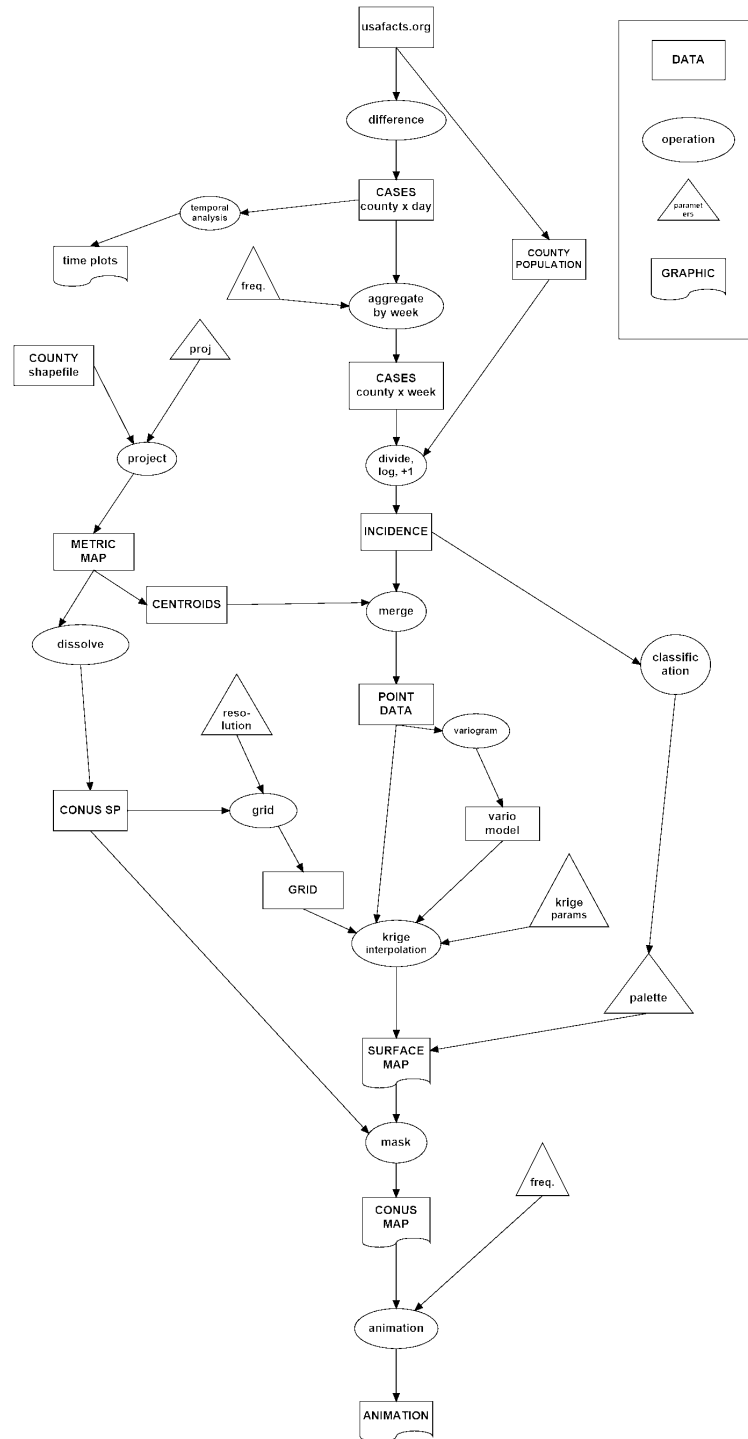
How common was by-mail voting before COVID-19 and which states will allow it in 2020?



DIMENSIONS & SCALE

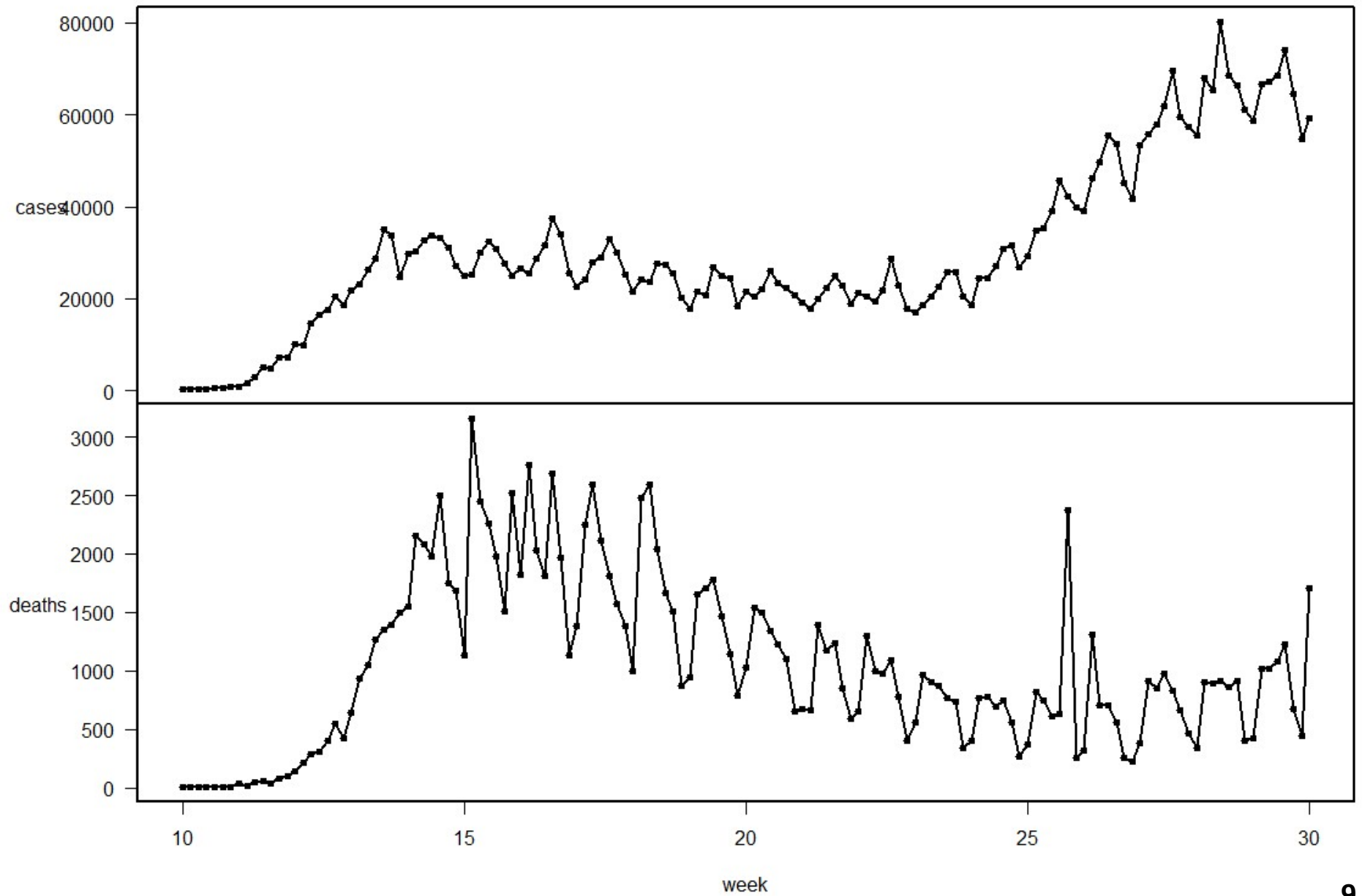


workflow

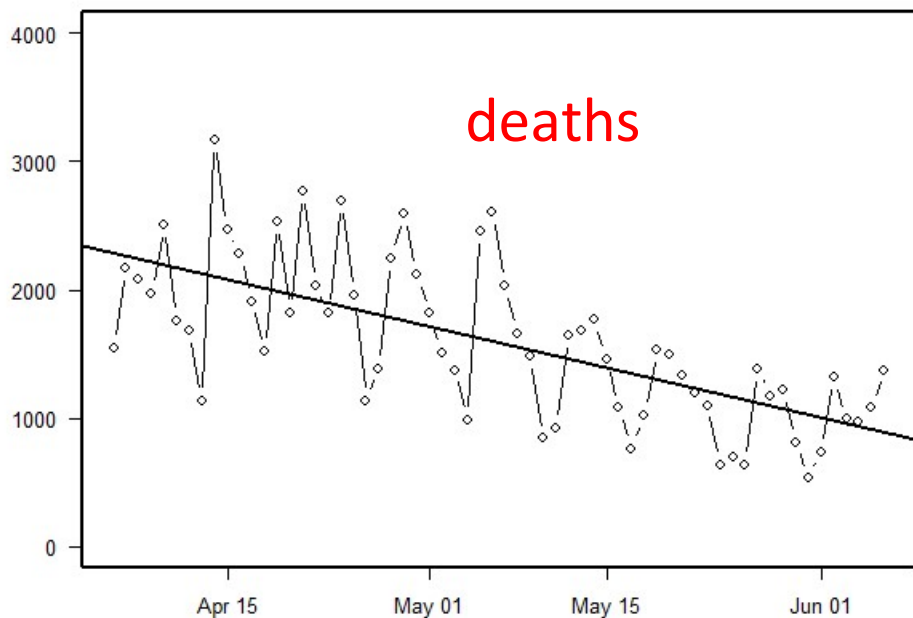


TEMPORAL ANALYSIS

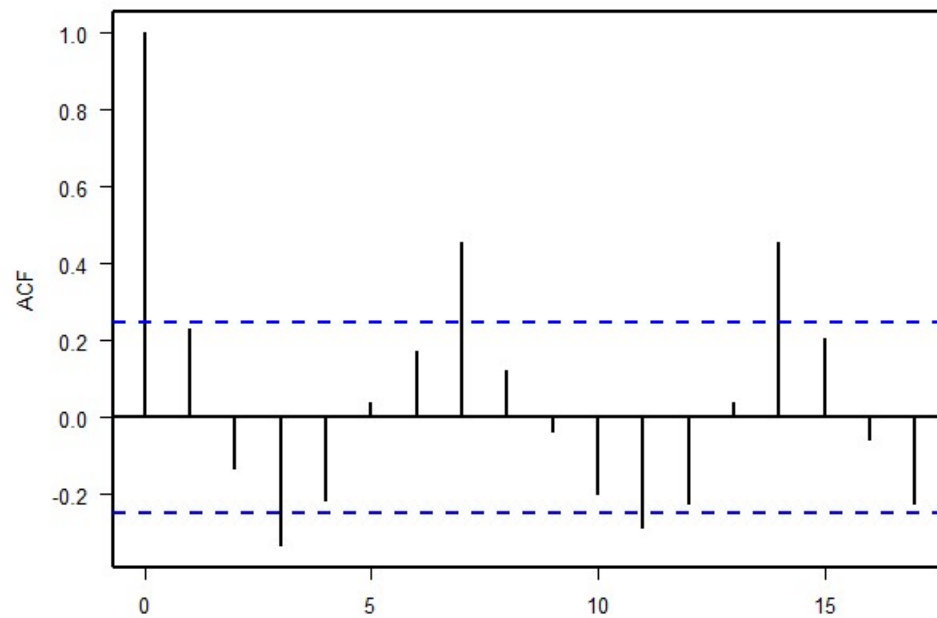
cases and deaths by day



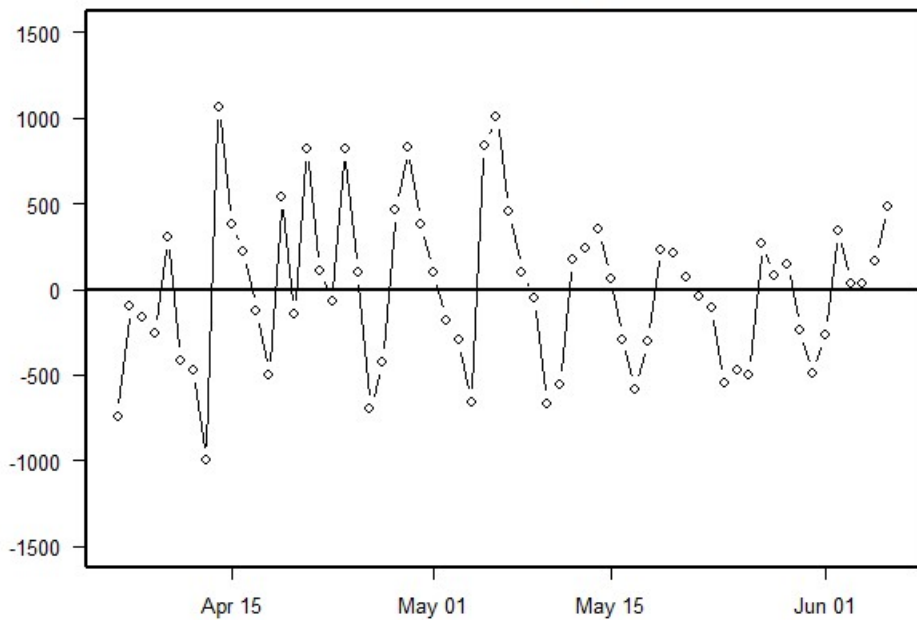
deaths



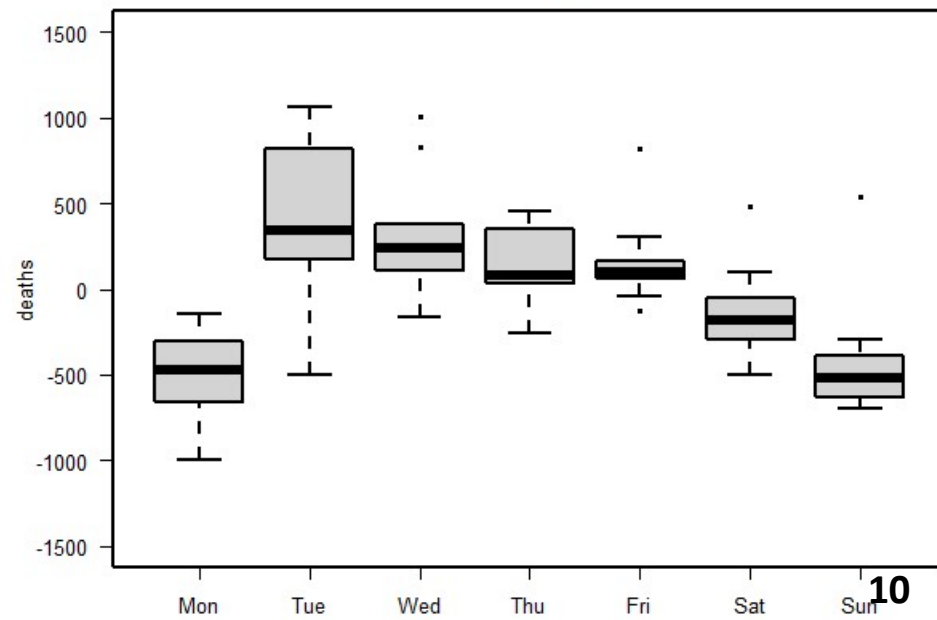
autocorrelation of residuals



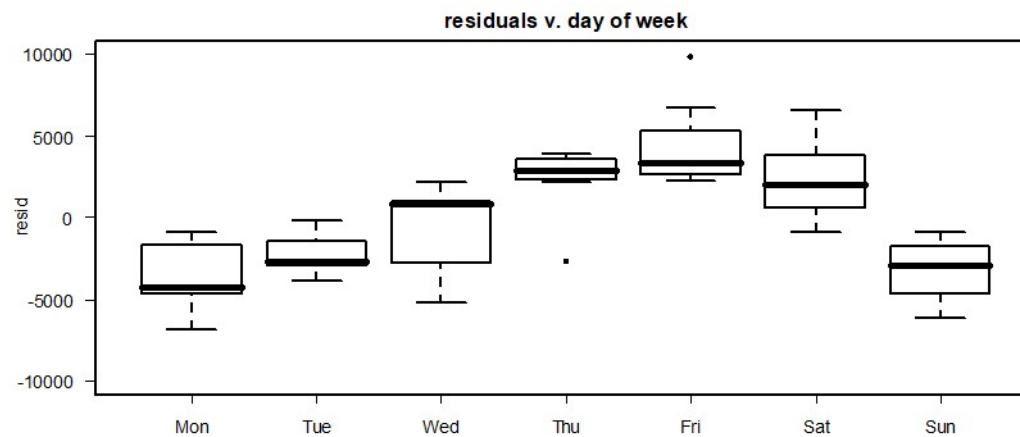
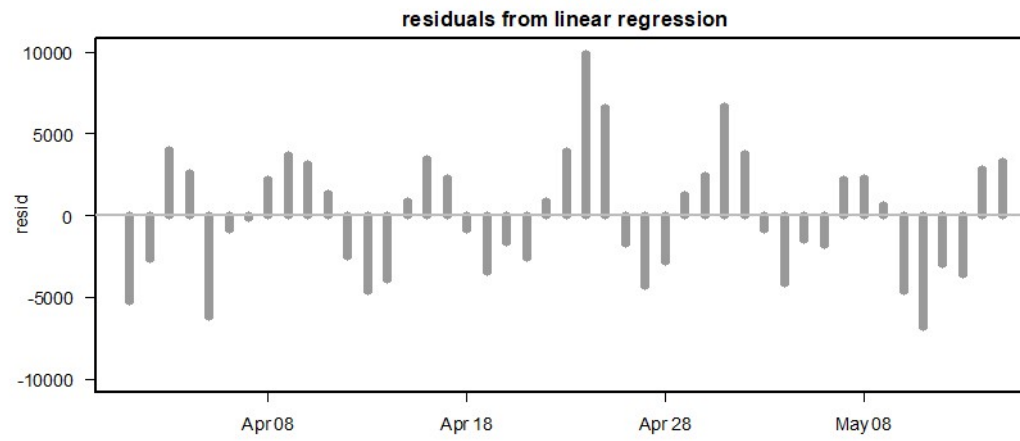
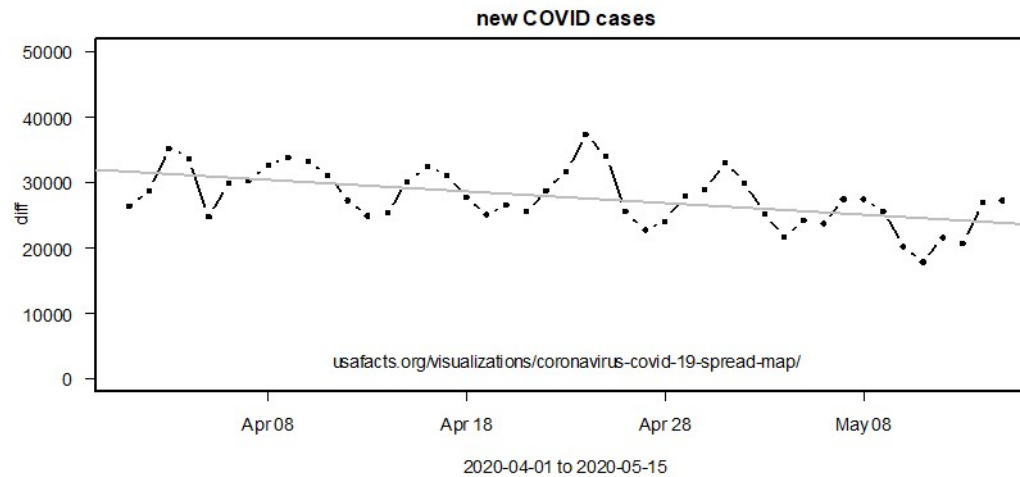
residuals from linear regression

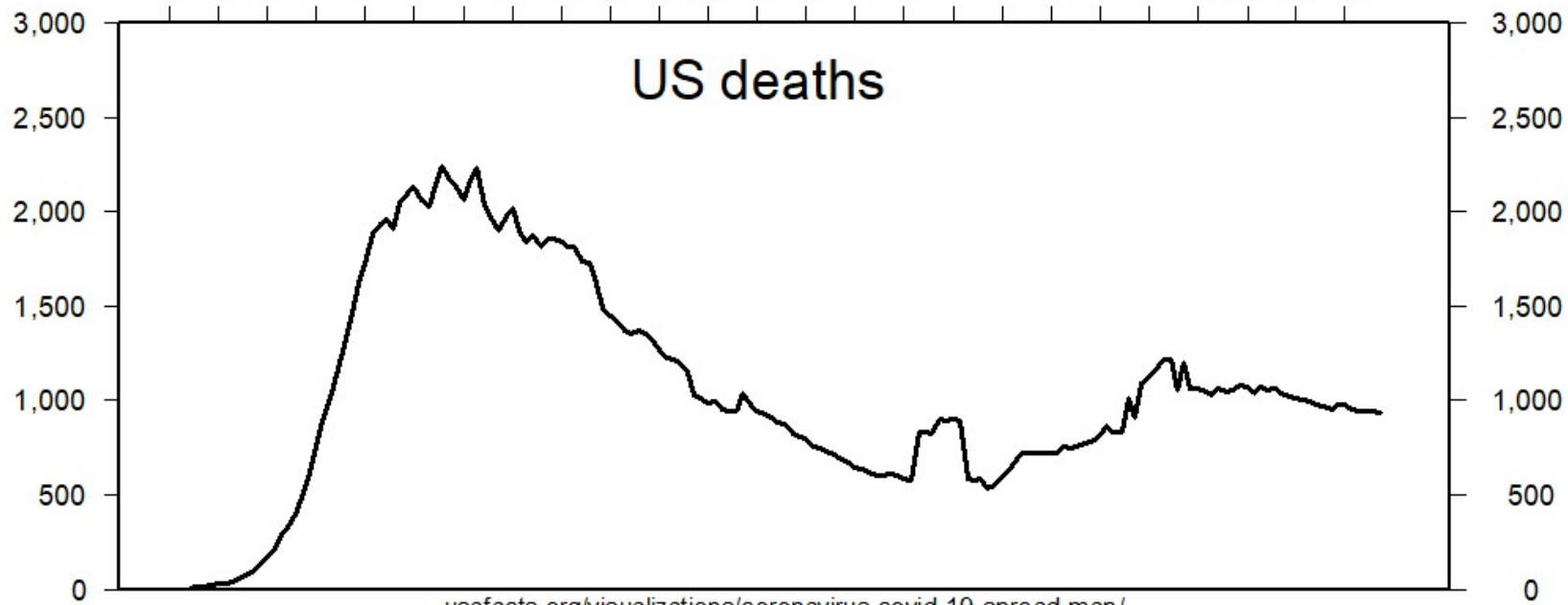
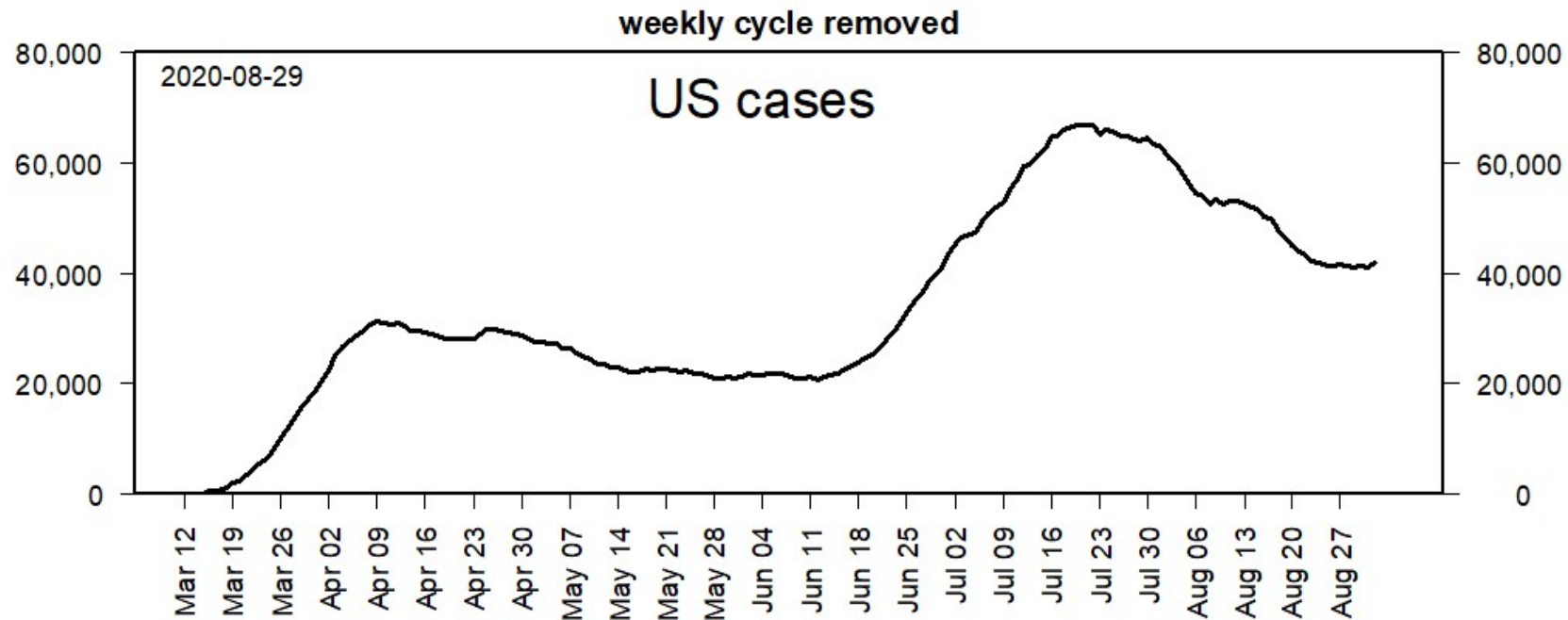


residuals v. day of week



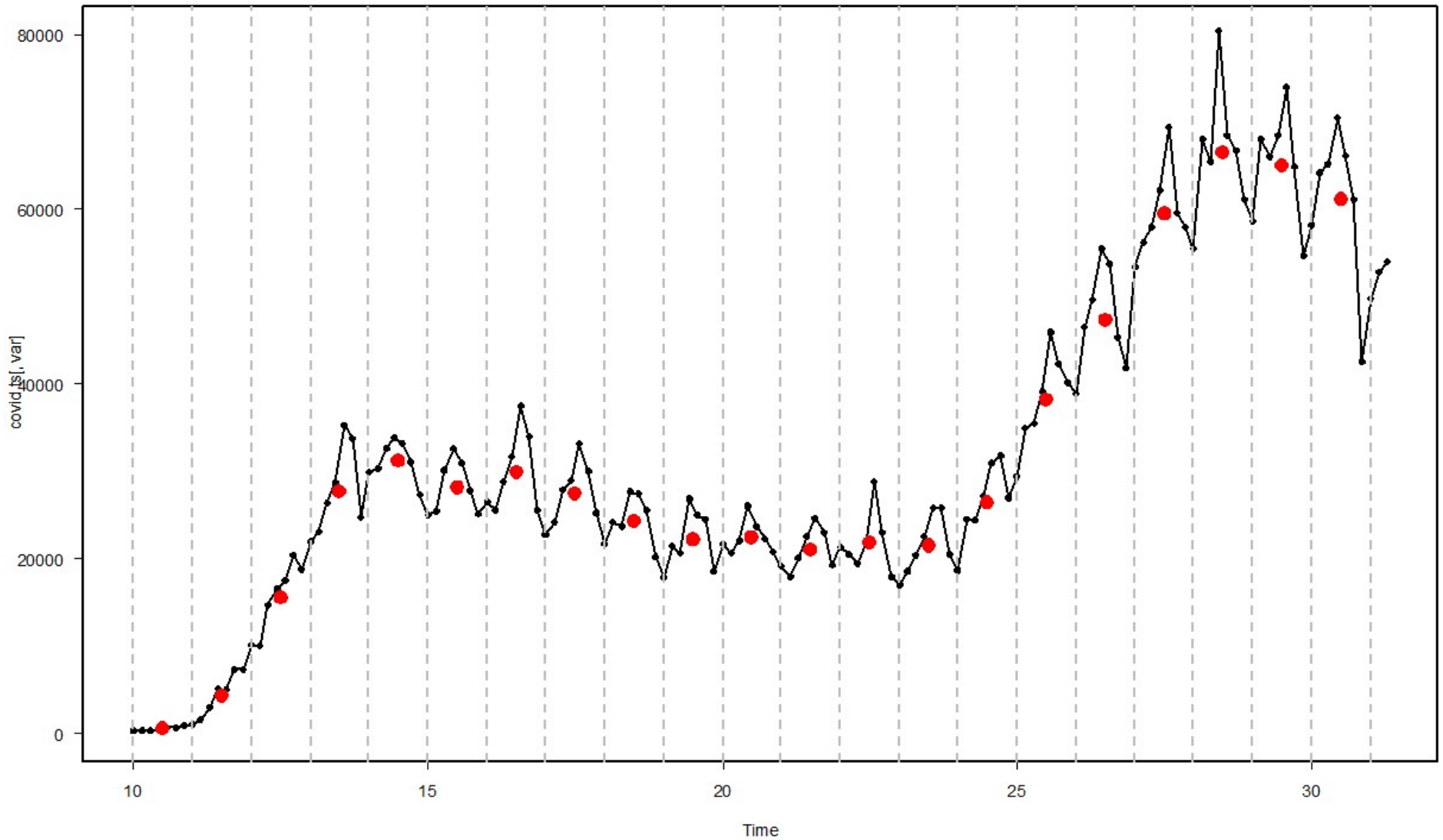
new cases





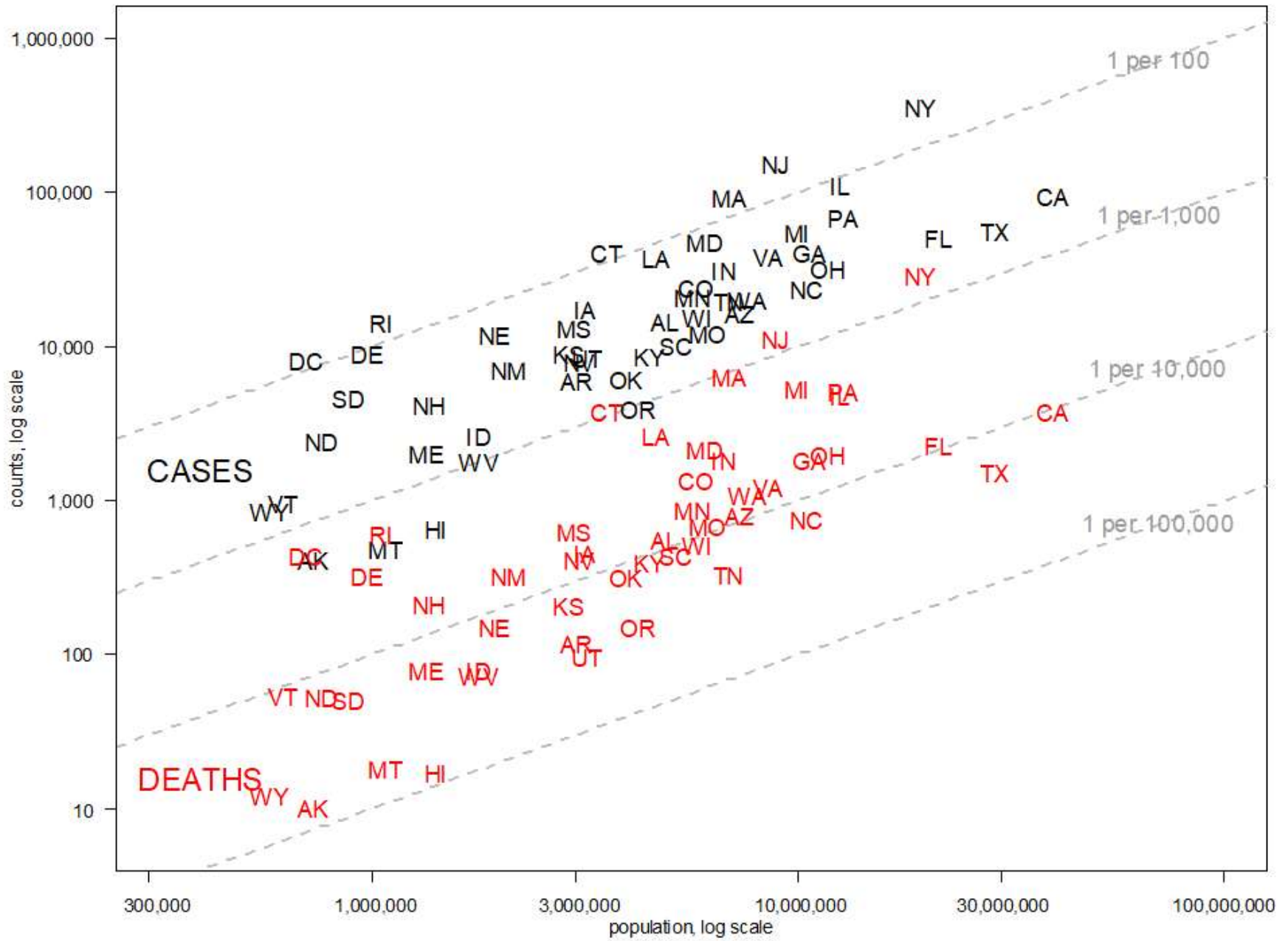
daily count, weekly mean

sum mean weekly cases

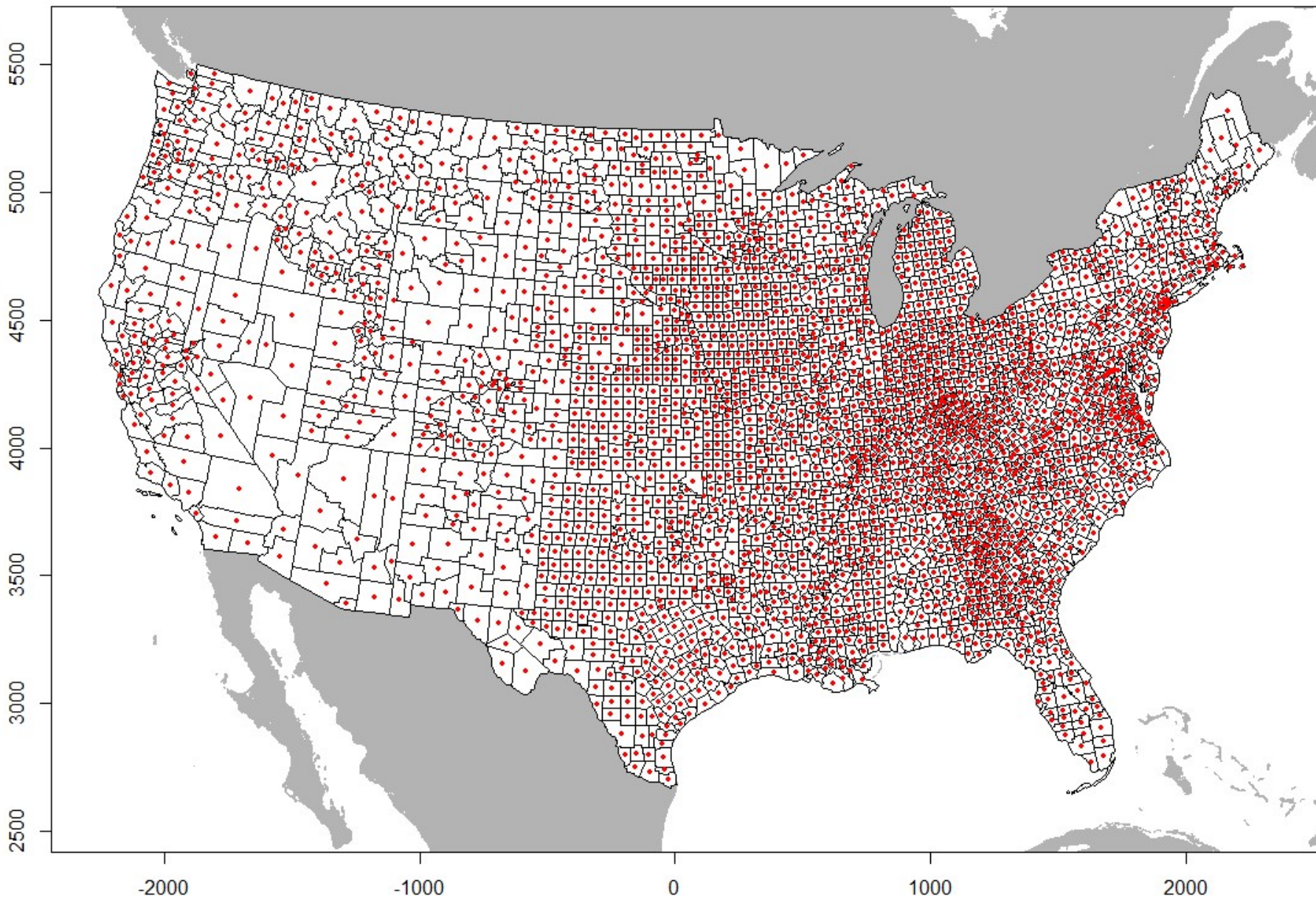


SPATIAL ANALYSIS

2020 May 24

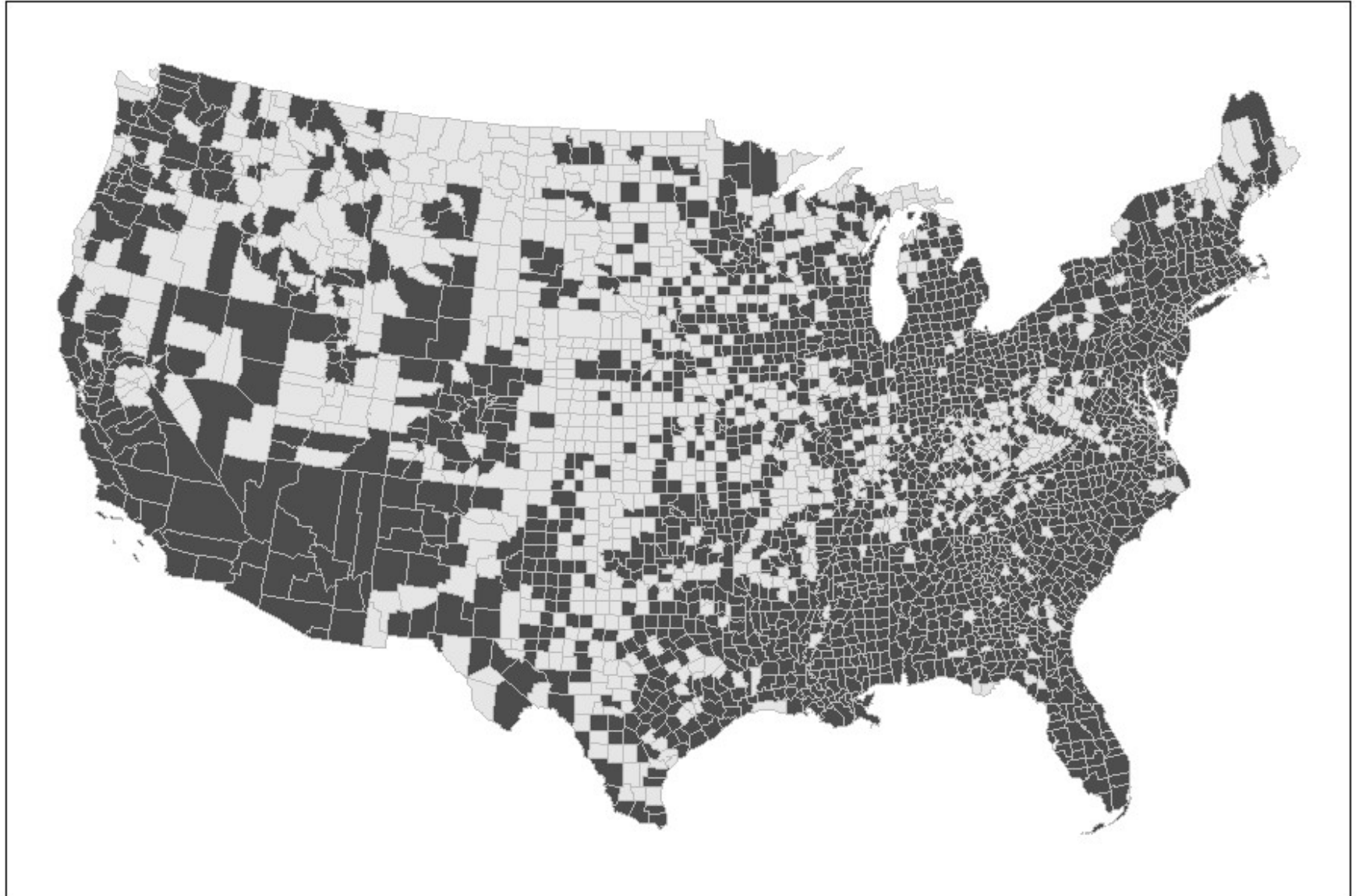


continental US centroids

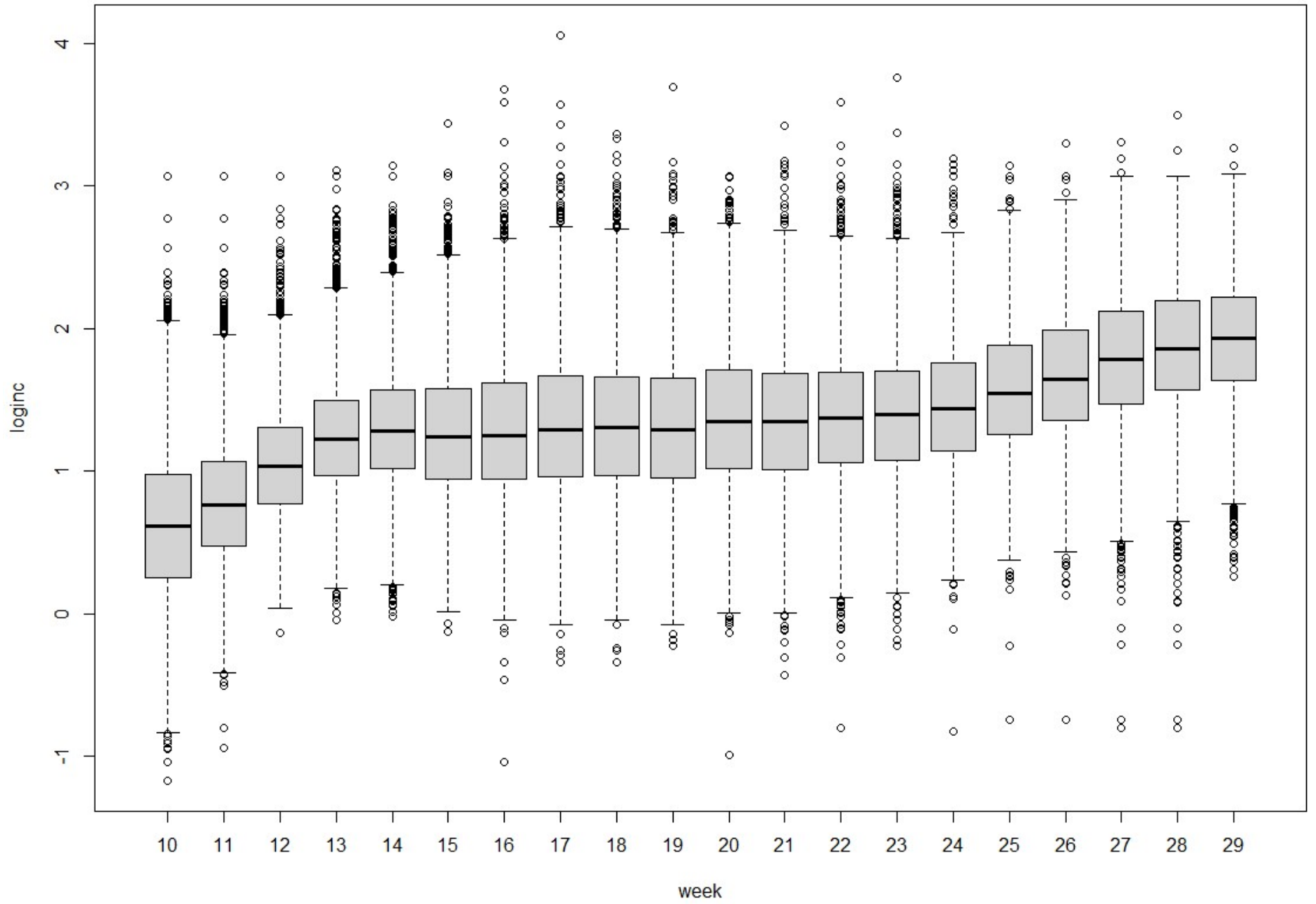


+proj=aea +lat_0=0 +lon_0=-97.2 +lat_1=29.5 +lat_2=45.5 +x_0=0 +y_0=0 +datum=WGS84 +units=km +no_defs

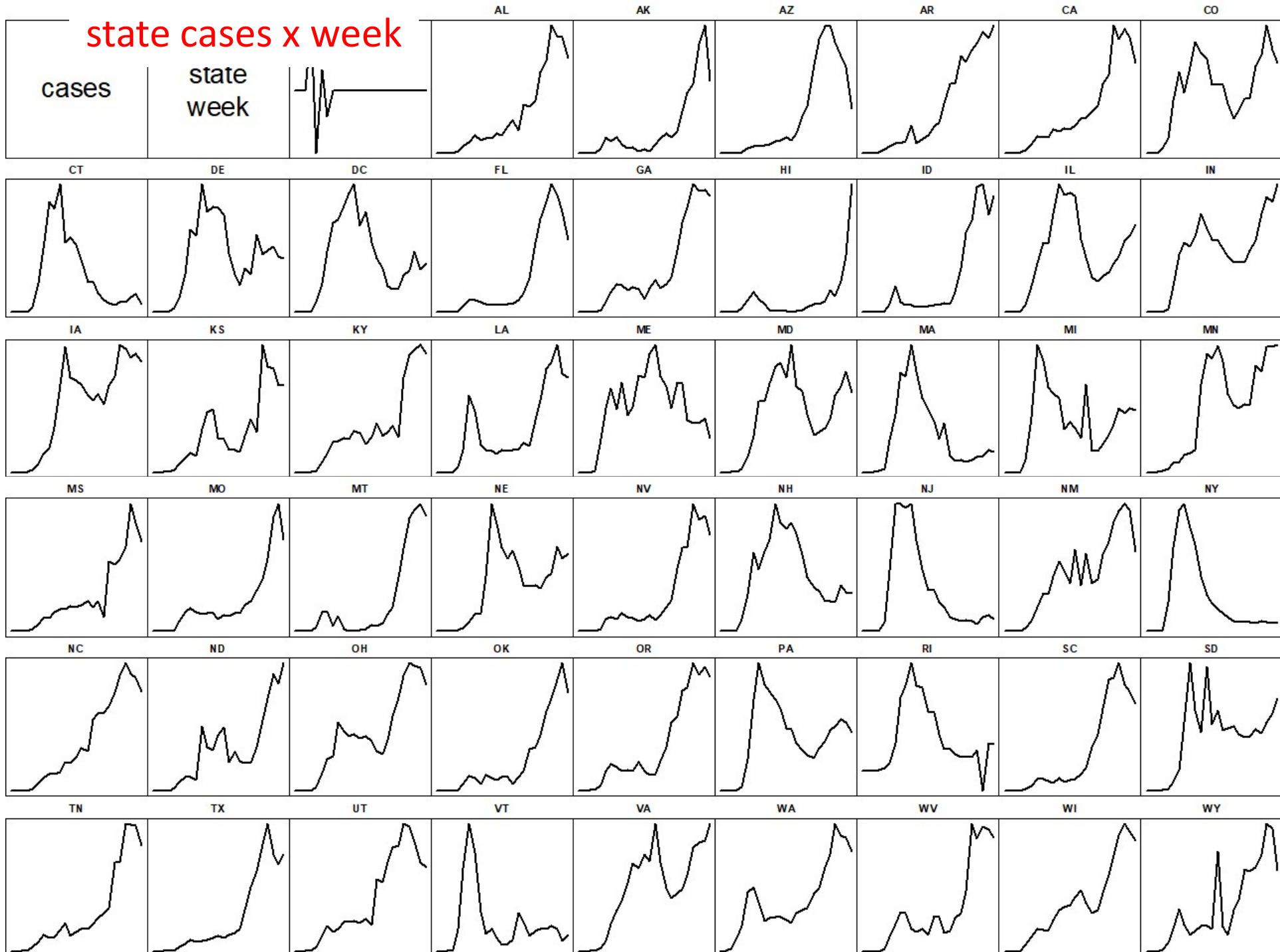
Counties with at least 1 death



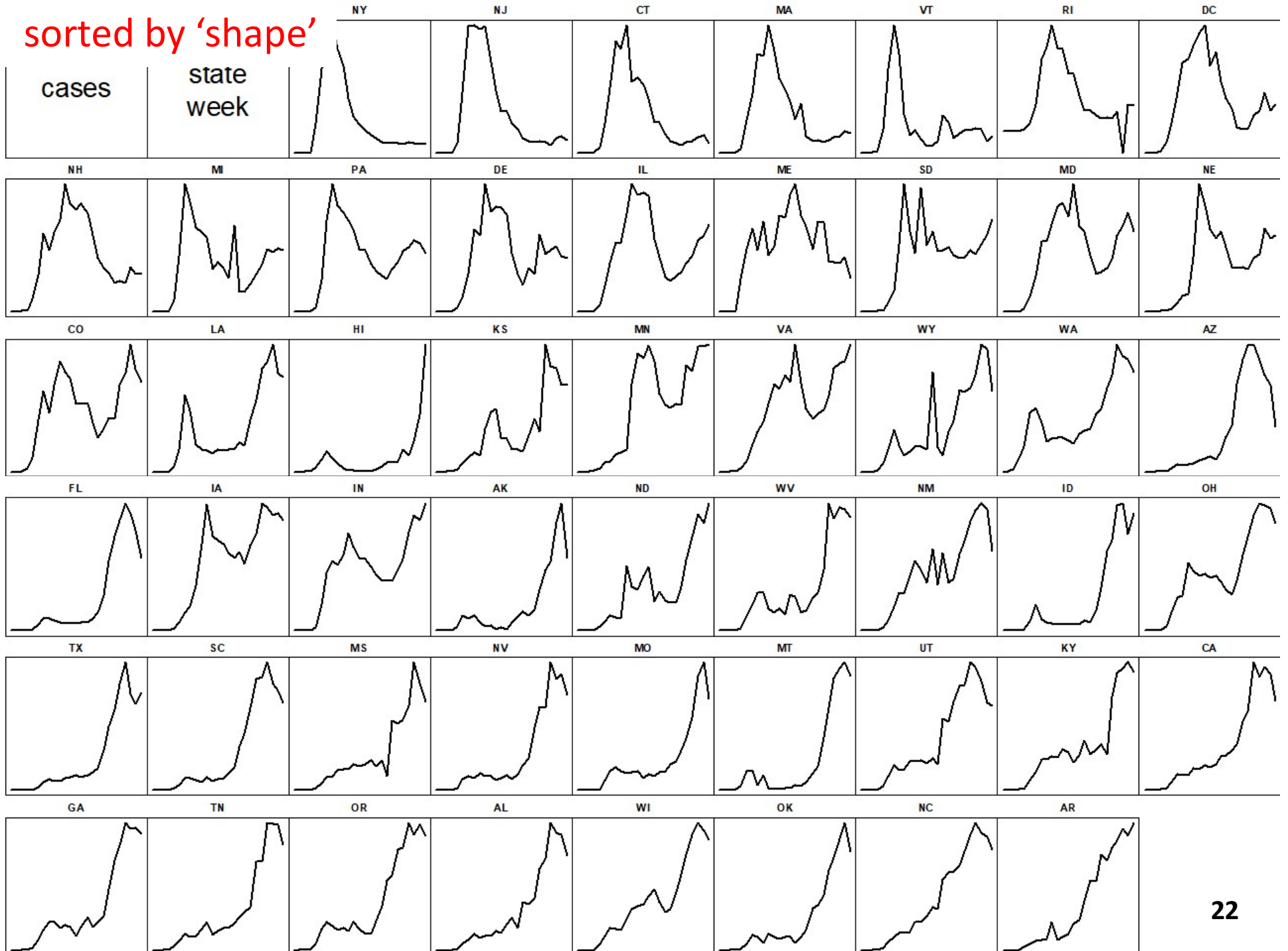
log incidence per 100,000



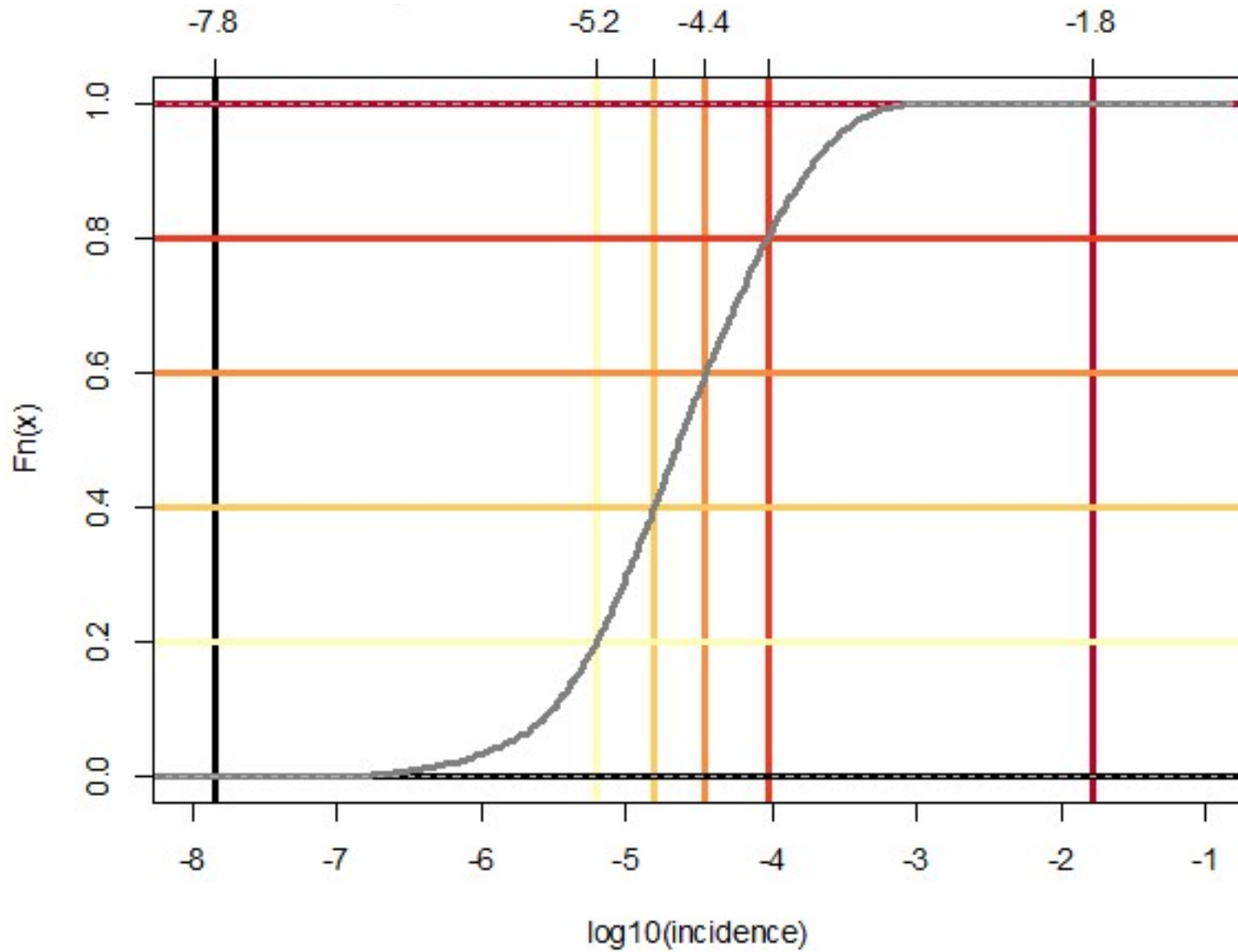
state cases x week



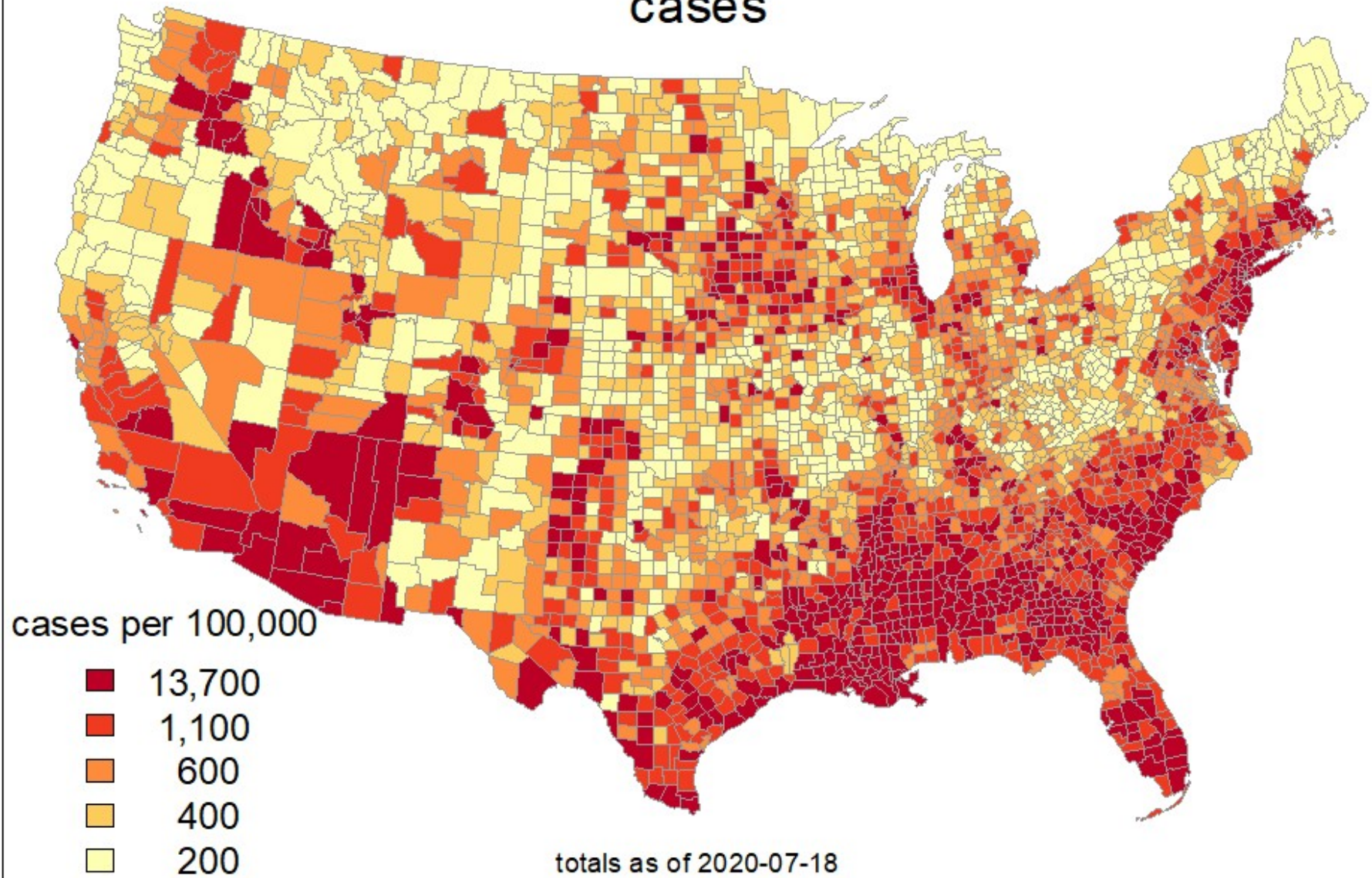
sorted by 'shape'



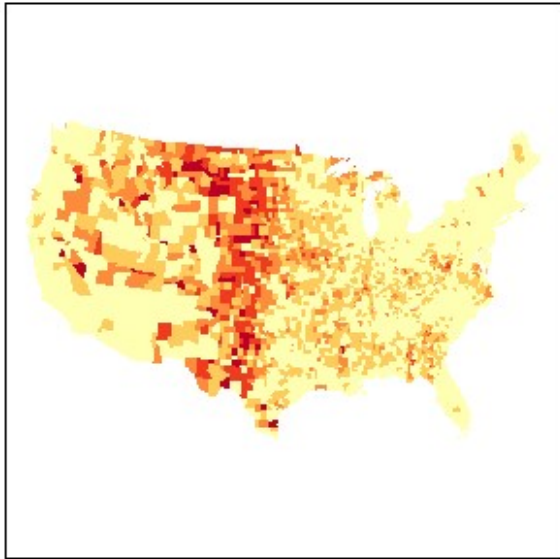
empirical cumulative distribution function



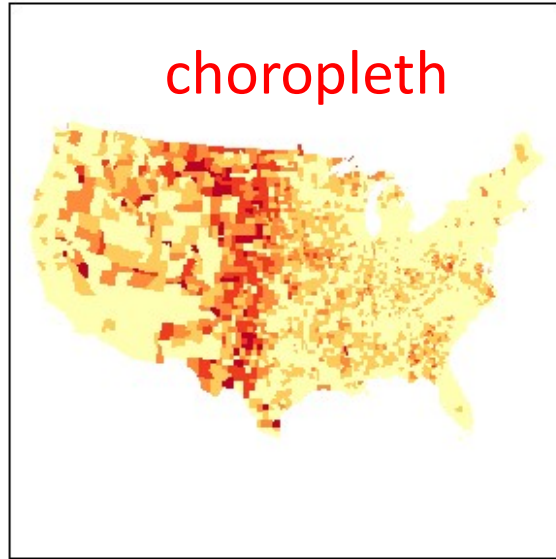
cases



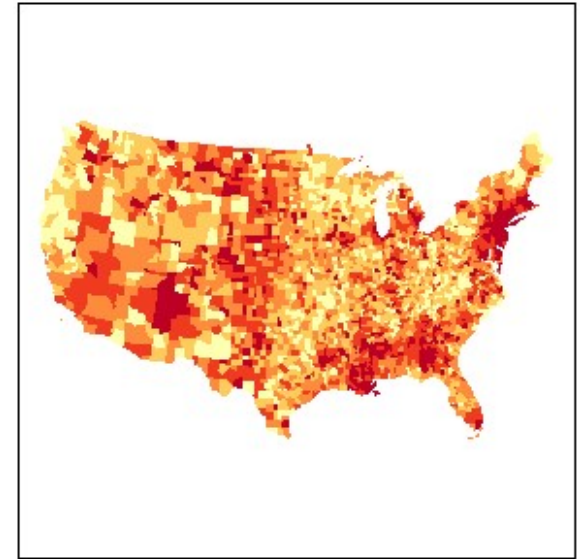
week 5



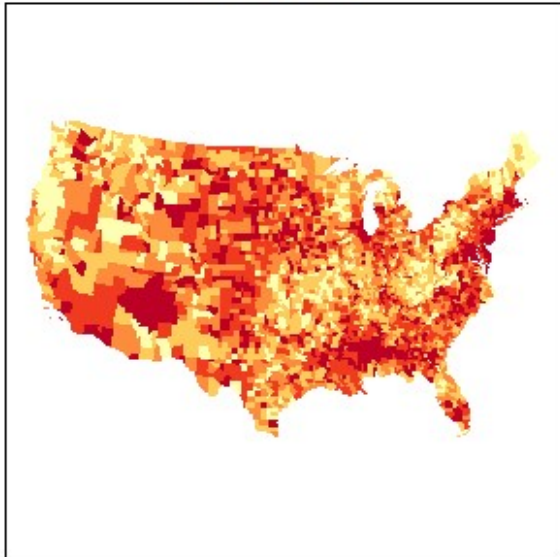
week 10



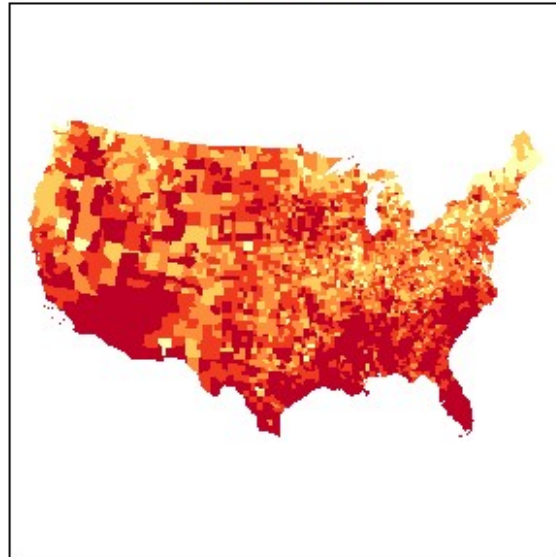
week 15



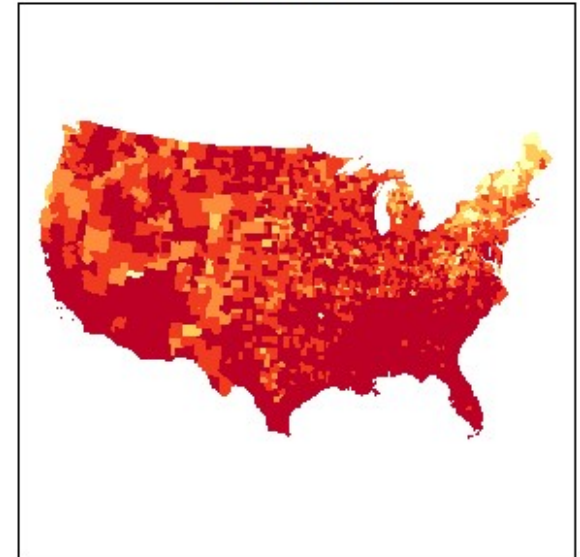
week 20



week 25

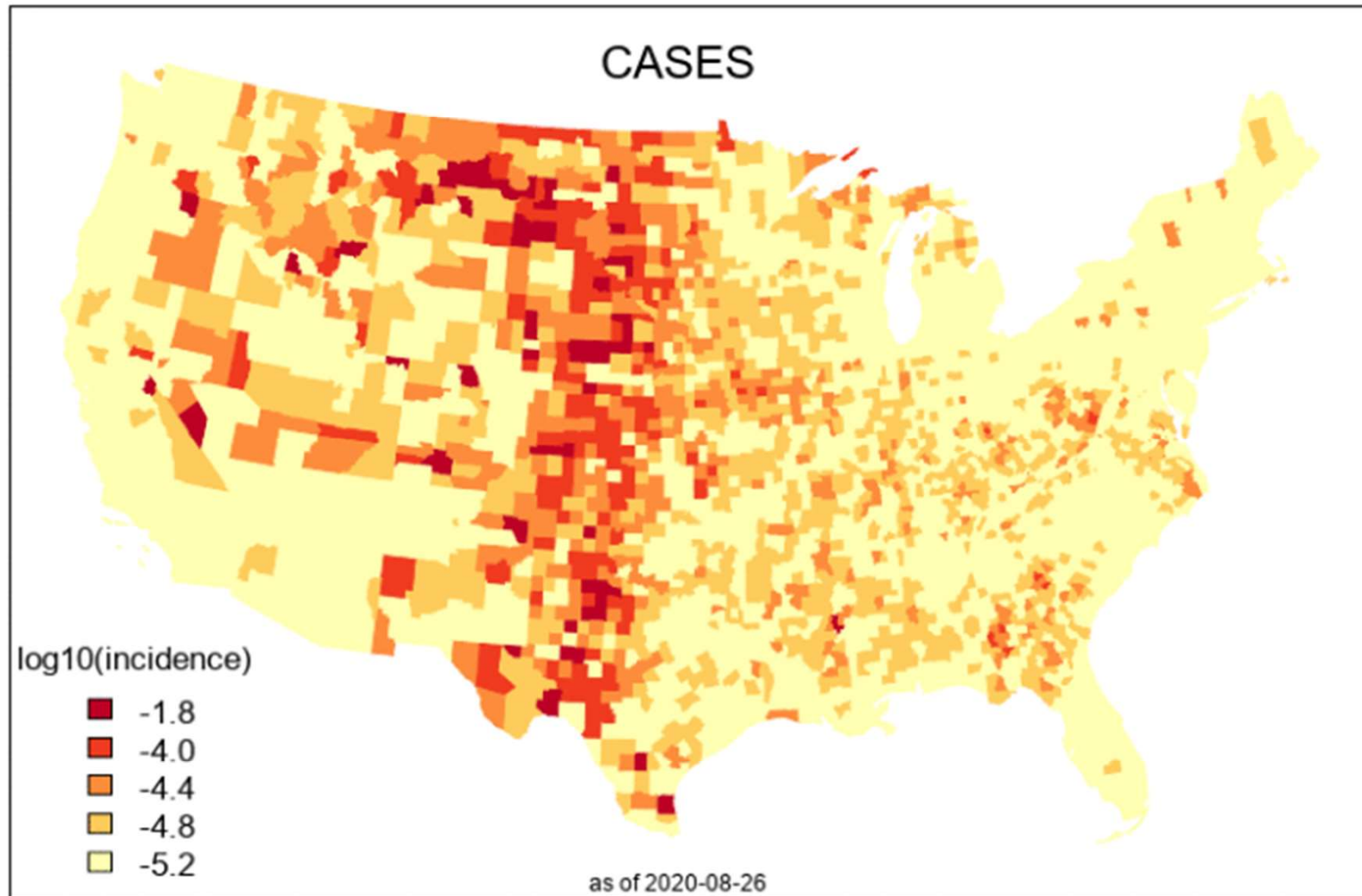


week 30



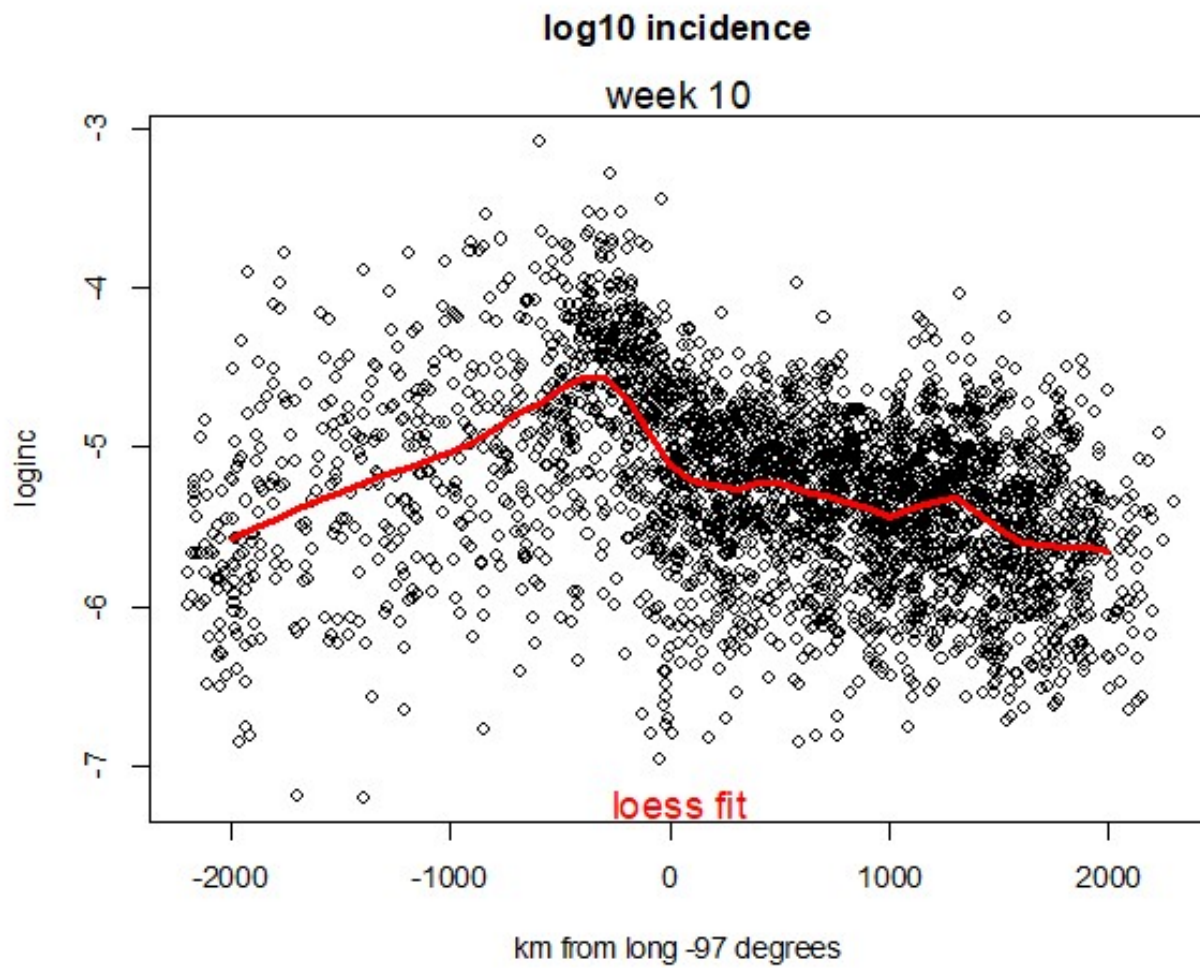
animation of choropleth maps

week 05



usafacts.org/visualizations/coronavirus-covid-19-spread-map/

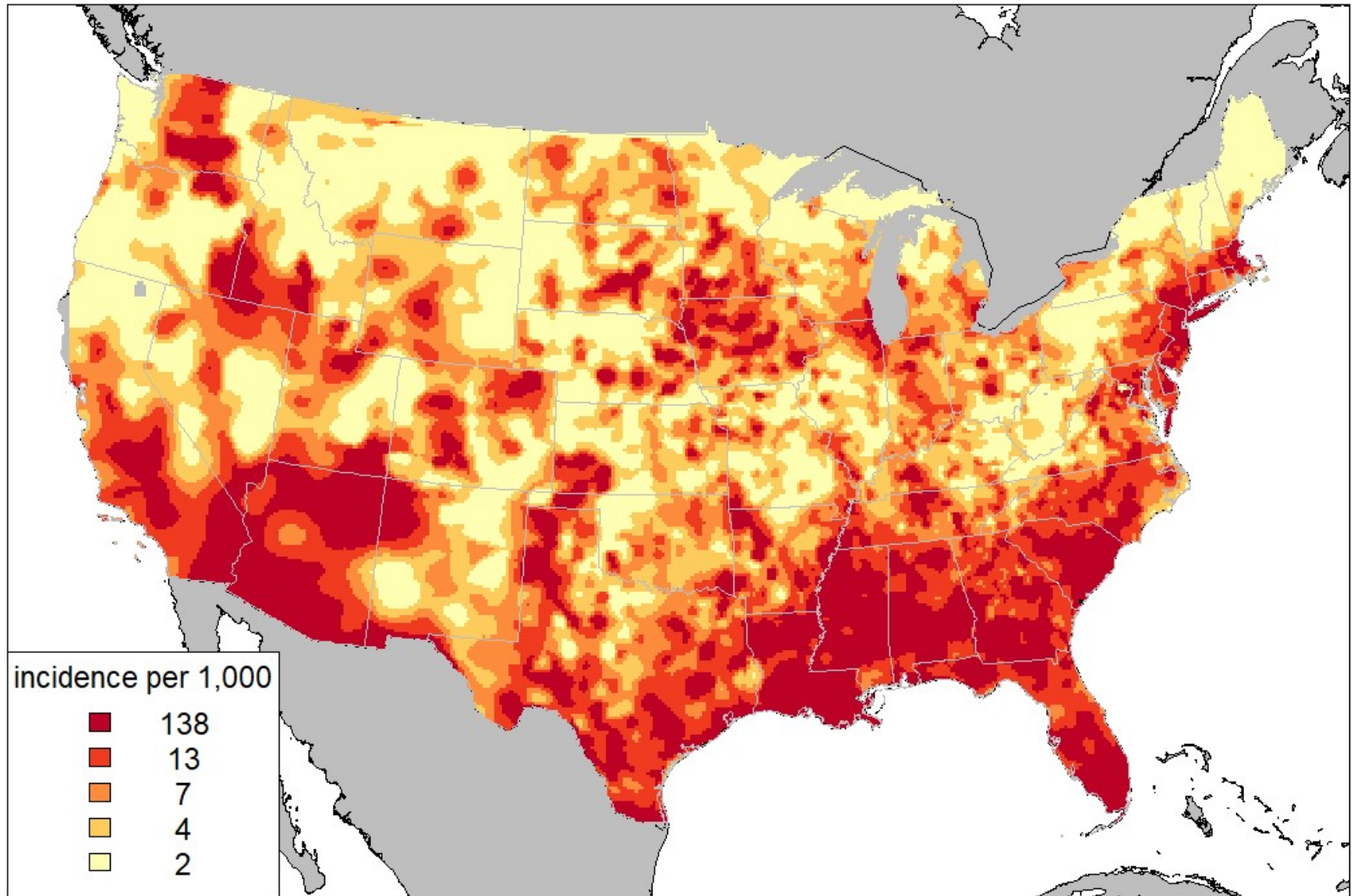
Midwest 'hump'



INTERPOLATION

spline interpolation

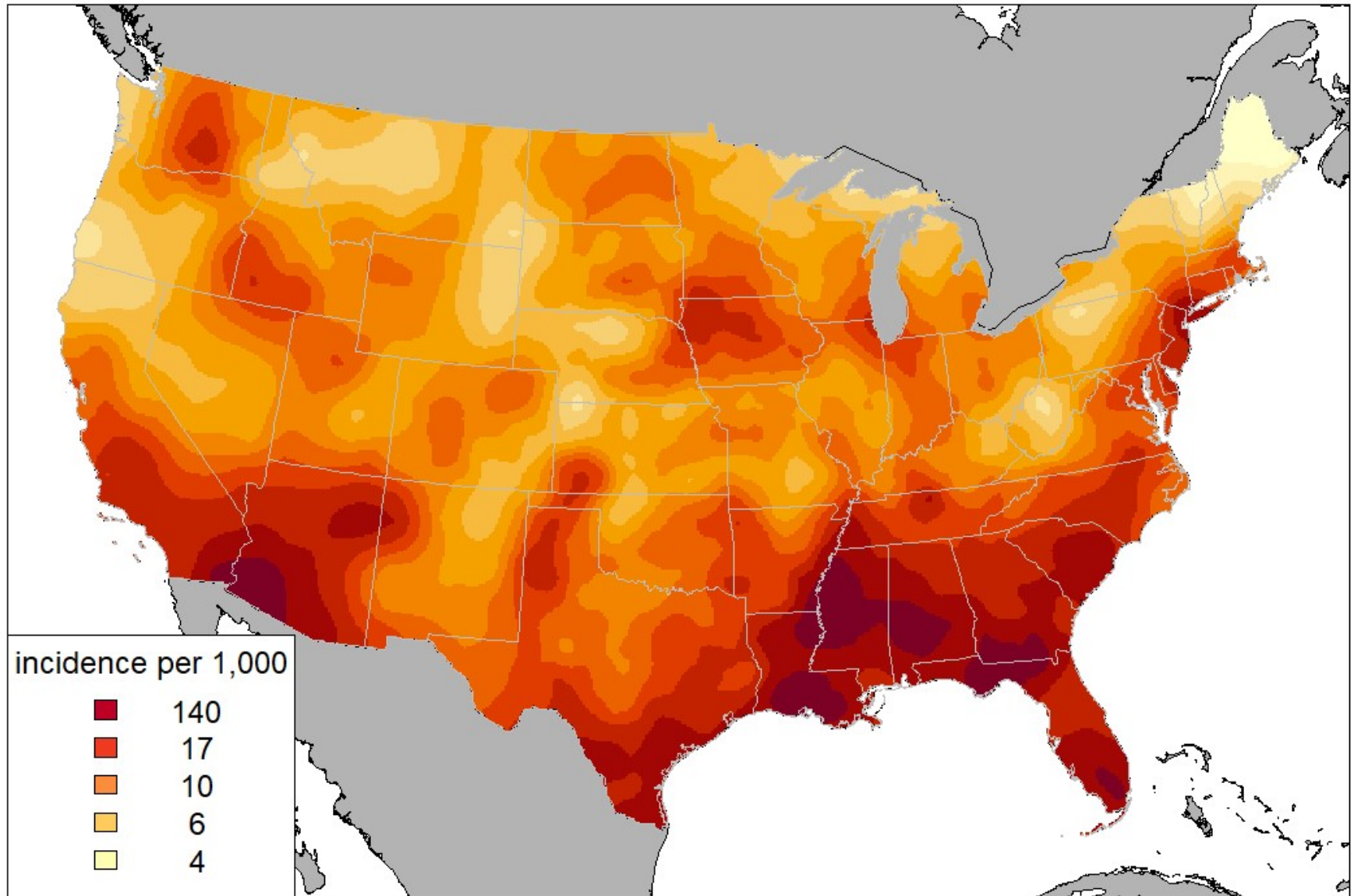
cases as of 2020-07-24



usafacts.org/visualizations/coronavirus-covid-19-spread-map/

krige interpolation

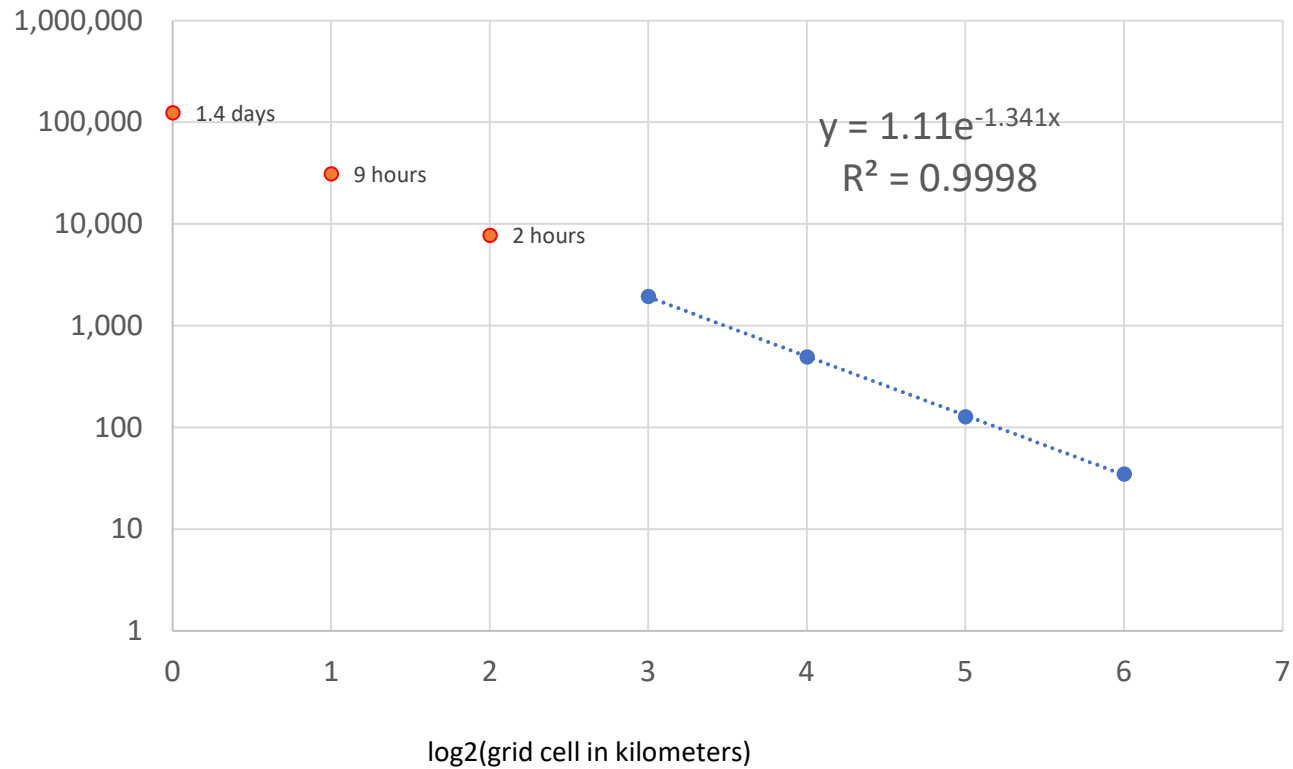
cases as of 2020-08-10



cell size in km: 8

usafacts.org/visualizations/coronavirus-covid-19-spread-map/

timings for kriging interpolation of 1 week's data



CONCLUSIONS

- Time and space must be analyzed together
- Spatial patterns: Midwest hump, spread from core regions, northward from south, etc.
- Animation helps highlight features of the space/time process
- Temporal and spatial SCALES are critical (week and maybe regional groupings?)

NEXT STEPS

- Geostatistics is complicated and CPU/cycle intensive
- Use network (neighbor graph) model?
- Need better response (incidence) index (negative binomial?)
- R library(spacetime)
- Other data sources?
- **ANYBODY WANT TO (GET) HELP (FROM) ME?**

Lee De Cola

LDECOLA@COMCAST.NET