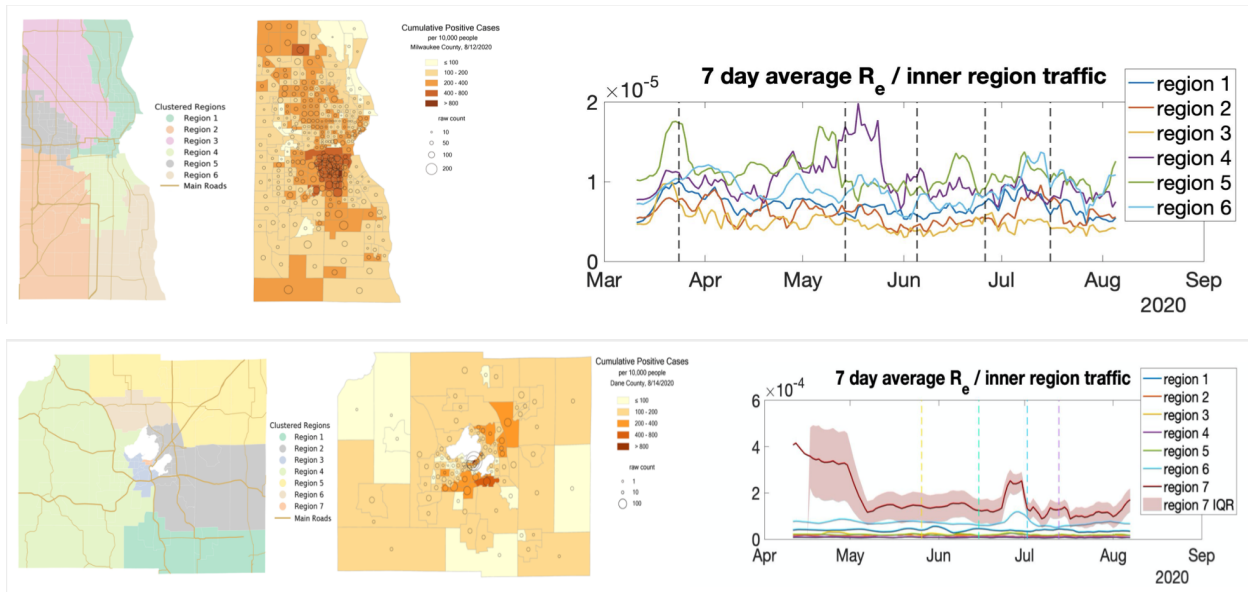


## Data Assimilation to Reconstruct Spreading of COVID-19

How did COVID-19 spread across counties in the US in the summer of 2020, when the country was mostly in a lockdown? A team at the University of Wisconsin-Madison collected public health data along with foot traffic data to reconstruct the pattern of spread in Milwaukee County and Dane County, the two largest counties in their state.

Their study revealed that the evolution of the epidemic on the county level depended crucially on the heterogeneities within a population. Geographic, socioeconomic, and cultural variations all played significant roles. The study investigated associations between the spread of the virus and business foot traffic, race and ethnicity, and age structure. In Dane County, which includes the college town of Madison, the most important heterogeneity is the age structure, while in the large urban area of Milwaukee County, racial and ethnic heterogeneity becomes a more significant factor in the spread of the virus.



Two machine learning techniques were used for the study: Spatial clustering based on origin-to-destination traffic flow and data assimilation constructed on a human mobility flow-augmented stochastic epidemic compartmental model. The former divided the county into (mostly connected) subregions and identified the segregation, while the latter recovered the historical virus transmission rate within each subregion.

**Credits.** IFDS investigators Nan Chen, Jordan Ellenberg, and Qin Li and several collaborators, funded in part by NSF grant 2023239.

**Publications.**

Xiao Hou et al., "Intracounty modeling of COVID-19 infection with human mobility: Assessing spatial heterogeneity with business traffic, age, and race," *Proceedings of the National Academy of Sciences* 118 (2021), <https://www.pnas.org/doi/abs/10.1073/pnas.2020524118>