**QUESTION?**

**EXAMPLE BELOW**

In this project, we will examine how uncertainty in estimates annual concentrations in air pollution (PM2.5), derived from the Bayesian Non-parametric Ensemble (BNE) method developed by Marianthi-Anna Kioumourtzoglou, Brent Coull et al., propagate through to estimates of the associations between PM2.5 concentrations and annual death rates. Our outcome will be Medicare annual all-cause death counts at the ZIP Code level, with area-level annual covariates derived from the American Community Survey (ACS). We will use annual denominator files as our offset in the Poisson model to create death rates.

We will obtain 50 nationwide complete random samples of annual ZIP Code-level PM2.5 concentrations from BNE model posterior distribution estimates across 1999-2019 (or as late as possible up to 2019). This will take place at Columbia University's HPC.

We will then build a Bayesian spatiotemporal model to obtain estimates of the association between annual PM2.5 concentrations and annual death rates for each of the 50 complete national exposure samples. Once each model has run, we will take 1000 draws from each of the posterior estimates of the association between PM2.5 annual concentrations and annual death rates. We will combine the draws from each of the 50 models to create a pool of 50,000 draws, from which we will calculate the final estimate of the targeted association between PM2.5 and death rates with a 95% Credible Interval.

To compare, we will also use the same Bayesian spatiotemporal model but with each of the individual PM2.5 concentration estimates from the original models which went into processing the BNE estimates (e.g., CACES estimates and others).

We will run all models involving the Medicare annual all-cause death counts on the Harvard RCE/FASSE cluster, and no health data, raw or processed, will ever leave the Harvard cluster. The team has extensive experience using the RCE/FASSE cluster and fully understands the sensitivity of the Medicare data.