**Data pipeline**

Hierarchy (roughly):

State > Municipality > Localidad > Manzana > lat/lon

Features

Water quality: at the municipality level, and with lat/lon

Mining: Down to the Manzana with lat/lon

Harmful chemicals: lat/lon

Manufacturing: lat/lon

Labels:

Deaths at the municipality level. (we think available for all states? Definitely have for SLP)

Vision for diffusion function:

* Pick a municipality
* Iterate over manzanas in municipality
  + Have population per manzana and centroids
  + Iterate over sources
    - Compute a per-source contribution weight
  + Normalize
* Compute average impact on municipality level

For each manzana, we have a set of weights and population.

For the state, take this (population-weighted) sum of computed effects

Naive idea: use per-source contribution weights