**Contact**

Matt Nissen, [mnissen1@stanford.edu](mailto:mnissen1@stanford.edu)

**Data Sources**

NIH TOXMAP

**Year**

2016

**Geographies**

All states & counties, lat/long

**Datasets**

Original datasets:

TRI sites- 2016 ([.csv file](https://drive.google.com/open?id=10oq5J8LjFK10hGDYTGVx0rsizYfPZvQU))

* Taken from NIH TOXMAP database, contains all current (as of 2016) TRI sites at the national level

TRI sites- all ([.csv file](https://drive.google.com/open?id=1l4VJH75eUKcp5bAQdfz9FhPHlCfUuovl))

* Taken from NIH TOXMAP database, contains all TRI sites - both current and historical - at the national level

Superfund sites ([.csv file](https://drive.google.com/open?id=1ll9Olx62Fw-vikCqVIl_bNLw8qcCETr4))

* Taken from NIH TOXMAP database, contains all Superfund sites - both current and historical - at the national level

New dataset: [All Industrial Sites (Current and Historical)](https://drive.google.com/open?id=1wCRsyVM2-WQJQd_vYng90Uu4ghaxMo6S)

* Contains lat/long of all sites
* Contains binary variables to determine historical/current and TRI/Superfund
* Contains all additional data about specific TRI/Superfund sites

**Processing**

Prepare TRI files for merge:

* Filter TRI sites (all) for SPARCC regions using [TRI\_sites\_historical.do](https://drive.google.com/open?id=1XHukAZnNiQWlJIeRA91m0V0idvOJ0ASw)
  + Save to [TRI\_historical\_sites.dta](https://drive.google.com/open?id=1LXN4EwLket311sgguH7ny71FsLcxXtGw)
* Using [TRI\_sites\_all\_merge.do](https://drive.google.com/open?id=1K-rfQryfpIMAH1epwWdV2lwt93emeOgN) filter TRI sites (2016) for SPARCC regions
  + Generate a historical flag, turn zipcode into a string to allow for the append
  + Append TRI\_hisotrical\_sites.dta
  + Remove duplicates, keep only current sites with the historical flag
  + Add a flag for site type
  + Save to [TRI\_all\_sites.dta](https://drive.google.com/open?id=11tPvB8iDl15hmUVnPzRZl4tg1q6-KkuH)

Prepare Superfund file for merge:

* Filter Superfund sites (.csv) for SPARCC regions using [superfund\_sites\_all.do](https://drive.google.com/open?id=1DatRQ7twxQlHQUELGZMObu9IM5Dwx6jZ)
* Clean variable names, add site type flag, add historical flag based on NPL status, turn zipcode into a string for append
* Save to [superfund\_sites\_cleaned.dta](https://drive.google.com/open?id=1881HrsCKQa5q3_M0V2WxFuXR2Xqkrh3C)

Merge TRI and Superfund

* Use [merge\_superfund\_TRI.do](https://drive.google.com/open?id=1I22cWfnXhq6shQYIe389h9AjkCoOhiEH) to append superfund\_sites\_cleaned.dta and TRI\_all\_sites.dta
* Save to [industrial\_sites.dta](https://drive.google.com/open?id=1wCRsyVM2-WQJQd_vYng90Uu4ghaxMo6S)

**Variable Description**

[Used variables in final dataset]

epa\_id EPA ID (*Superfund only*)

site\_id Site ID (*Superfund only*)

name Industrial site name

address Industrial site street address

city Industrial site city

county Industrial site county

county\_fips County FIPS Code (2010)

state Industrial site state

zipcode Industrial site zipcode

latitude Industrial site latitude

longitude Industrial site longitude

npl\_status National Priorities List Status (*Superfund only*)

nplstatusdate NPL status date (*Superfund only*)

hrs\_score Hazard Ranking System Score (*Superfund only*)

site\_type Industrial site type flag: 0 - Superfund; 1 - TRI

historical\_site Industrial site historical flag: 0 - 2016 Site; 1 - Historical Site

tri\_id TRI ID (*TRI only)*

frsid FRS ID (*TRI only)*

totalreleasesallyears Total releases (all years) (*TRI only*)

totalreleases2016 Total releases (2016) (*TRI only*)

**Additional Notes**