



EPA Superfund – Nothing super about it

Joshua Sung
Data Scientist, Anpisa Consulting & Analytics



GUIDING QUESTIONS

- ▣ For communities that have a Superfund site, what are the common population and socioeconomic features of these areas?
- ▣ Does this problem reveal an environmental justice issue?



WHAT IS SUPERFUND?

- Thousands of contaminated waste sites exist
 - Hazardous waste, improper management
 - Manufacturing facilities, processing plants, landfills, abandoned mines
- Love Canal
- CERCLA (i.e. Superfund)
 - Program that allows EPA to clean contaminated sites
 - Forces responsible parties to clean up or reimburse government



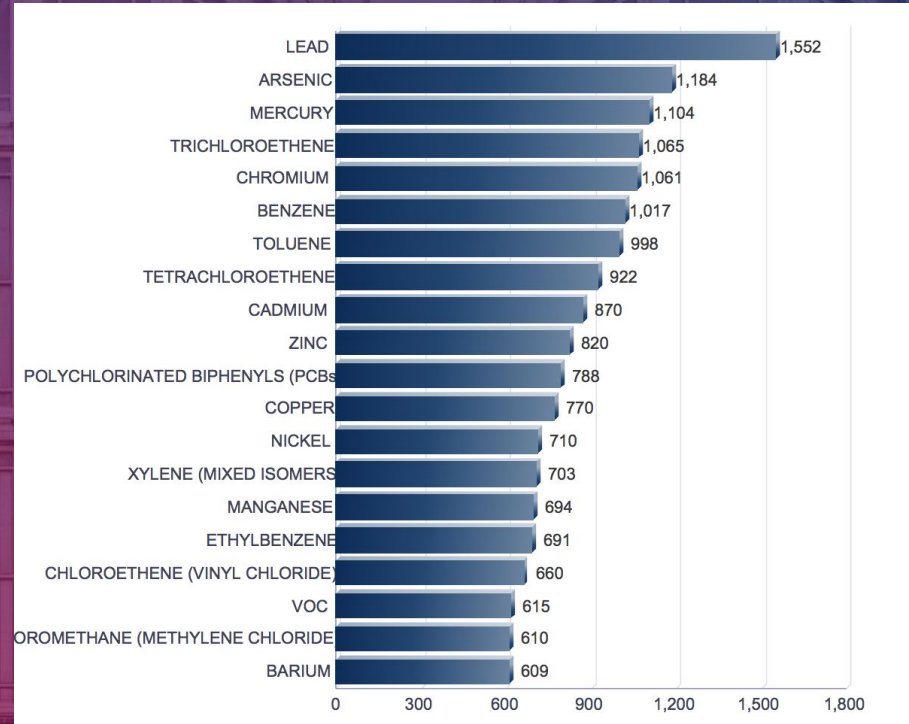
GOALS OF SUPERFUND

- ▣ Protect human health, environment
- ▣ Make responsible parties pay for cleanup work
- ▣ Involve communities in the process
- ▣ Return Superfund sites to productive use



COMMON CONTAMINANTS AT THESE SITES

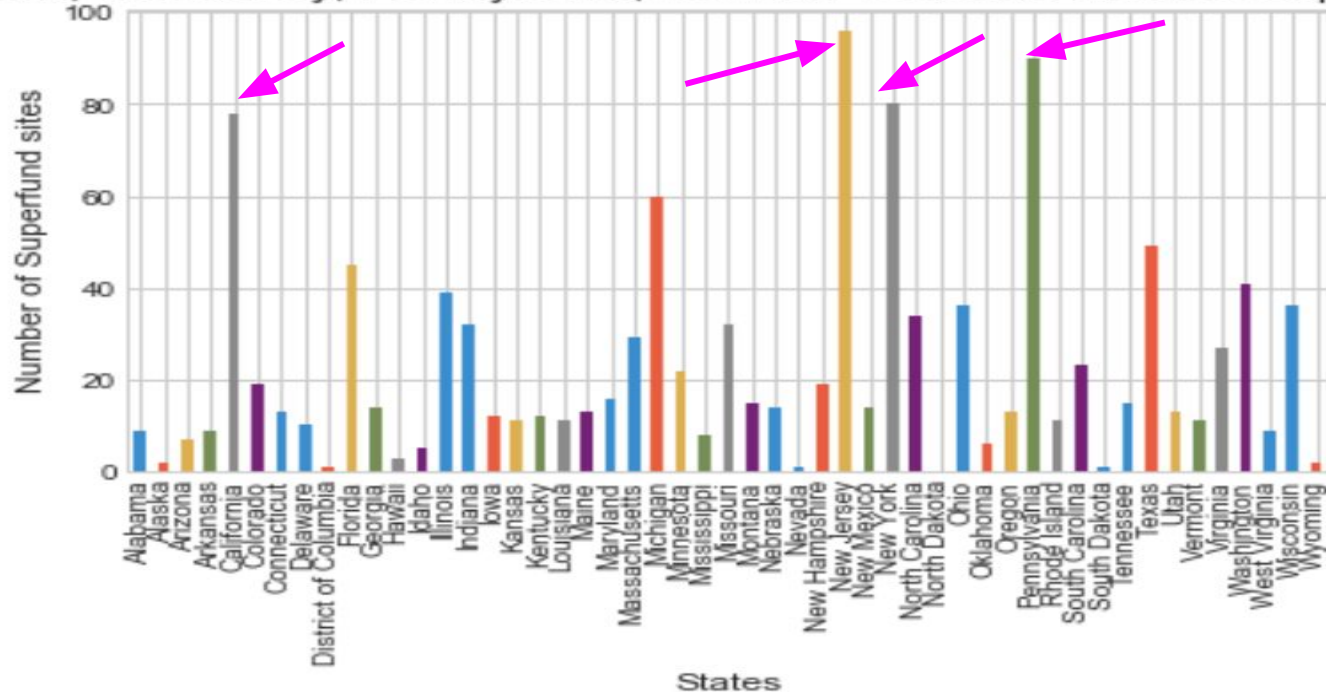
- Lead
- Arsenic
- Mercury
- Cadmium
- Radiation





HOW MANY SUPERFUND SITES ARE OUT THERE?

California, New Jersey, Pennsylvania, and New York have the most Superfund sites





BRIDGESTONE WEST-LAKE LANDFILL SUPERFUND SITE





A PERSONAL STORY: ROBBIN AND MICHAEL DAILEY





WHAT IS OUR DATA?

- ▣ 70,000 observations by census tract
- ▣ 500+ population demographic, socioeconomic data
- ▣ Whether a Superfund site is nearby or not
- ▣ Sites are on the National Priorities List (NPL)
 - ▣ Eligible for long-term remedial action



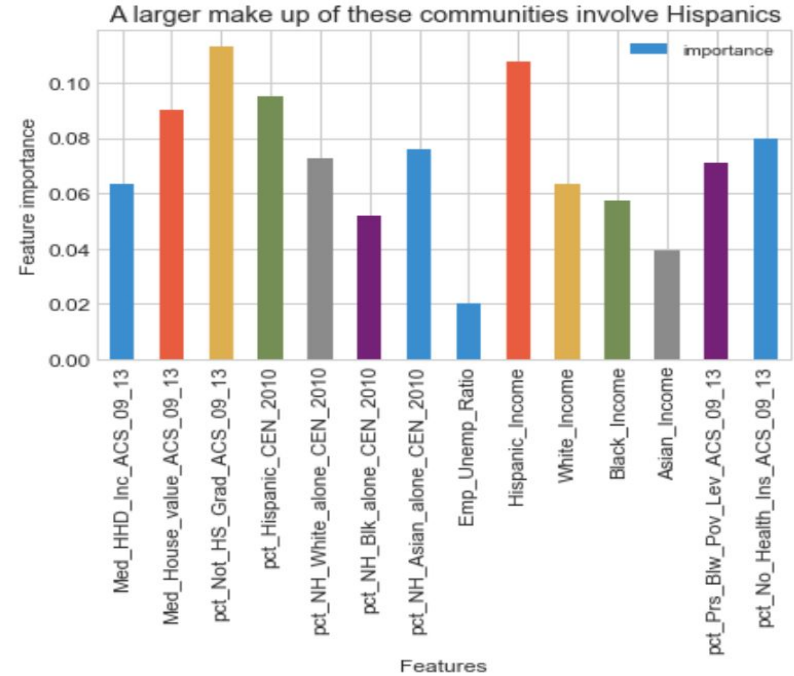
FEATURES ANALYZED

- ▣ Population and economic features
 - Ethnicity
 - Age group
 - US, foreign born
 - Education
 - Whether in poverty or not
 - Gender
 - Employment
 - Median household income
 - Median House value



MODEL BUILDING

- Based on these models:
 - Models don't agree with each other
 - No feature stands out
 - Could indicate possible exogenous features





IMPROVEMENTS TO THE MODEL

- ▣ Superfund site characteristics
 - Longitude, latitude
 - Contamination, radiation
 - Why it was designated there?
 - How long has it been a site?
 - Stage in its remediation process (if any)
- ▣ Identify possible exogenous features to add
- ▣ Other population features:
 - Disease occurrence among residents



CONCLUSIONS

- ▣ Model results are inconclusive → not one features stands out
- ▣ More data needs to be collected → distinguish near Superfund site or not
- ▣ Superfund analysis possibly more complicated
 - ▣ Perhaps beyond the scope of data science?
 - Epidemiology, environmental