Analysis of demographic and environmental conditions near selected facilities

2023-04-16

Abstract - Executive Order 14008 calls on EPA and other Agencies to make achieving environmental justice part of their missions, and EO 12989 directed EPA to make environmental justice part of their mission by identifying and addressing, as appropriate, disproportionately high and adverse human health or environmental effects of their programs, policies, and activities on minority populations and low-income populations in the United States. The United States Environmental Protection Agency (US EPA) analyzed baseline demographic and environmental conditions in communities living NA these sites. The analysis used EPA’s EJAM tool and EJScreen version 2.1 with demographic data based on the Census Bureau’s 2016-2020 American Community Survey (ACS). The analysis found that PLACEHOLDER EXAMPLE PLACEHOLDER EXAMPLE PLACEHOLDER EXAMPLE PLACEHOLDER EXAMPLE PLACEHOLDER EXAMPLE PLACEHOLDER EXAMPLE PLACEHOLDER EXAMPLE PLACEHOLDER EXAMPLE

Table of Contents

# 1 Executive Summary

Executive Order 14008 calls on EPA and other Agencies to make achieving environmental justice part of their missions, and EO 12989 directed EPA to make environmental justice part of their mission by identifying and addressing, as appropriate, disproportionately high and adverse human health or environmental effects of their programs, policies, and activities on minority populations and low-income populations in the United States.

EPA has conducted an analysis to characterize baseline environmental conditions faced by communities living near NA. The United States Environmental Protection Agency (US EPA) analyzed baseline demographic and environmental conditions in communities living NA. The analysis used EPA’s EJAM tool and EJScreen version 2.1 with demographic data based on the Census Bureau’s 2016-2020 American Community Survey (ACS).

## 1.1 Broad overview of findings

PLACEHOLDER EXAMPLE The environmental indicators – especially proximity scores – are more notable than the demographic indicators at these sites. Demographic indicators are moderately above average overall.

PLACEHOLDER EXAMPLE The environmental indicators – especially proximity scores – are more notable than the demographic indicators at these sites. Demographic indicators are moderately above average overall.

PLACEHOLDER EXAMPLE The environmental indicators – especially proximity scores – are more notable than the demographic indicators at these sites. Demographic indicators are moderately above average overall.

## 1.2 Summary of Findings

* PLACEHOLDER EXAMPLE Overall, % people of color, % limited English proficiency, and the Demographic Indicator are more than 1.5x the State rate, for the population within 1 mile.
* PLACEHOLDER EXAMPLE About a third of these sites are above the 80th percentile in State for the Demographic Indicator. The same is true for % low income and % with less than high school.
* PLACEHOLDER EXAMPLE About half of these 57 sites are in just 4 states: FL, NY, PA, or MA. Most of the people here live near just 8 sites (15% of sites). Most of the sites with the higher demographic indicators are owned by Covanta or Wheelabrator.
* PLACEHOLDER EXAMPLE Many of the sites with the highest demographic indicators also have proximity scores that are 5 to 10 times the State average.
  + PLACEHOLDER EXAMPLE The average person’s RMP score is more than 3x their State’s average.
  + PLACEHOLDER EXAMPLE The average person’s NPL and TSDF scores are about 2.5x State averages.
  + PLACEHOLDER EXAMPLE Most of these sites are >=80th in State for NATA. Same for RMP.
* PLACEHOLDER EXAMPLE
* PLACEHOLDER EXAMPLE
* PLACEHOLDER EXAMPLE

# 2 Introduction

Executive Order 12898 (59 FR 7629; February 16, 1994) established federal executive policy on environmental justice. Its main provision directed federal agencies, to the greatest extent practicable and permitted by law, to make environmental justice part of their mission by identifying and addressing, as appropriate, disproportionately high and adverse human health or environmental effects of their programs, policies, and activities on minority populations and low-income populations in the United States.

EPA defines environmental justice as the fair treatment and meaningful involvement of all people regardless of race, color, national origin, or income with respect to the development, implementation, and enforcement of environmental laws, regulations, and policies.

Executive Order 14008 (86 FR 7619; January 27, 2021) also calls on Agencies to make achieving environmental justice part of their missions “by developing programs, policies, and activities to address the disproportionately high and adverse human health, environmental, climate-related and other cumulative impacts on disadvantaged communities, as well as the accompanying economic challenges of such impacts.” It also declares a policy “to secure environmental justice and spur economic opportunity for disadvantaged communities that have been historically marginalized and overburdened by pollution and under-investment in housing, transportation, water and wastewater infrastructure and health care.”

EPA also released its “Technical Guidance for Assessing Environmental Justice in Regulatory Analysis” (U.S. EPA, 2016) to provide recommendations that encourage analysts to conduct the highest quality analysis feasible, recognizing that data limitations, time and resource constraints, and analytic challenges will vary by media and circumstance.

# 3 Methods

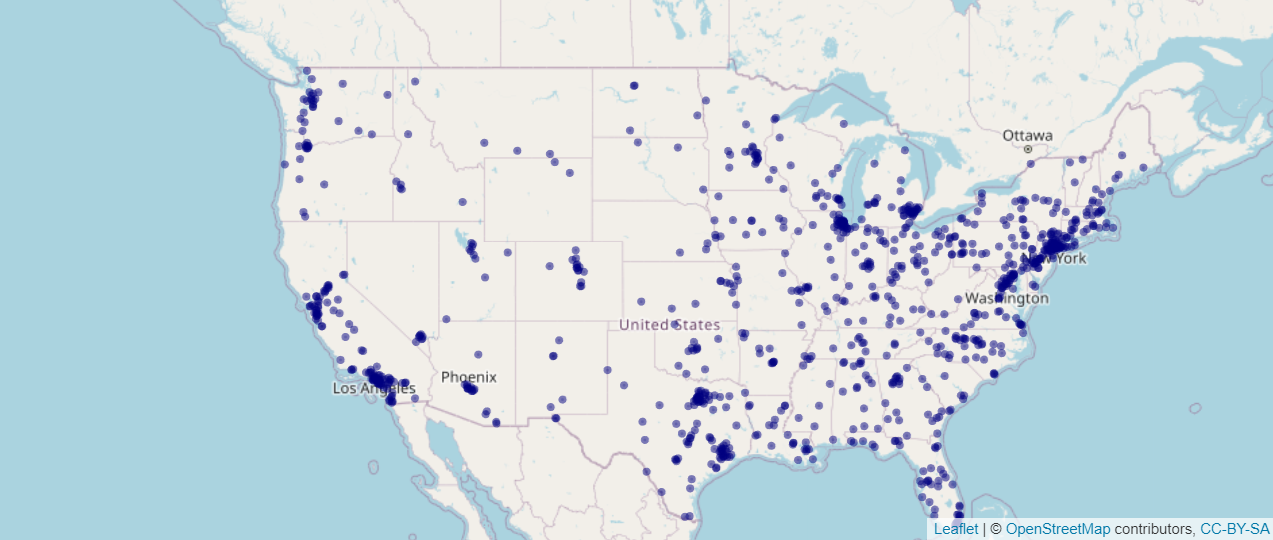
EPA’s EJAM (Environmental Justice Analysis Multisite) tool was used to develop this analysis. EPA’s [EJAM tool](https://rstudio-connect.dmap-stage.aws.epa.gov/content/dc3cda00-20a2-47ed-a753-0dcb89eb8f2a/) is a user-friendly web app that can summarize demographics and environmental conditions for any list of places in the nation. It provides interactive results and a formatted, ready-to-share report with written explanations of the results, tables, and graphics. The report can provide EJ-related information about people who live in communities near any of the industrial facilities on a list, for example.

The basic methodology and data used for this analysis are the same as those used in EPA’s EJScreen tool, with a few exceptions described below, and are described in detail in EJScreen’s documentation, at [EJScreen](https://epa.gov/ejscreen) and with more technical details available at [EJScreen technical documentation page](https://www.epa.gov/ejscreen/technical-information-about-ejscreen). The only notable differences are the following:

* EJAM may include additional demographic or environmental indicators.
* For a proximity analysis (to characterize everyone living within a certain certain distance from a point such as a facility), EJAM identifies which residents live nearby using a slight variation on how the distance to each Census block is measured. While EJScreen uses ESRI’s ArcGIS calculations, EJAM calculates the distance using formulas implemented in the R language for statistical computing (R Core Team 2022). These measurements provide almost identical results for estimated distance from the average person in a block to a given site point. PLACEHOLDER MORE INFO
* EJAM aggregates indicator values within and across locations, converts them to percentiles, and does other summary calculations using the same formulas to the greatest extent possible, but in R rather than within EJScreen itself. There may be slight differences between raw scores and percentiles in EJScreen and EJAM in some cases.

## 3.1 Selection of sites analyzed

## 3.2 Data Viz 1 – Map



## 3.3 Estimating locations and population counts of residents

### 3.3.1 Spatial resolution of data

The analysis used EPA’s EJAM tool and EJScreen version 2.1 with demographic data based on the Census Bureau’s 2016-2020 American Community Survey (ACS) and the corresponding version of Decennial Census information on geographic boundaries and FIPS codes for blocks, block groups, tracts, counties, and States.

See EJScreen methodology for details.

### 3.3.2 Analytic method for buffering, and tools used to implement that method

The buffering method was the same as EJScreen’s method in principle, i.e., Percent of each block group’s population that is estimated as inside a buffer is based on which Census block internal points are included in the buffer and using a block weight that is the Census 2020 block population as a fraction of the parent block group’s Census 2020 population (which is not quite the same as the ACS population count). That block weight is a fraction of the parent block group. A slightly different method was used to identify which block points are inside the buffer than EJScreen uses, but results should be almost identical [confirm this and clarify]

See EJScreen methodology for details. EJAM finds nearby block points using a very fast approach but otherwise uses the EJScreen method of estimating nearby residential populations for proximity analysis.

## 3.4 Demographic and environmental indicators

PLACEHOLDER - The demographics included here are those in EJScreen and also race/ethnicity subgroups that comprise the total count of people of color. POC are defined as all other than those self-identifying in ACS survey data as white, single race, not Hispanic or Latino - i.e., non-hispanic white alone (NHWA). The subgroups include Hispanic or Latino (“hispanic”), several groups that are not hispanic but of only single race (e.g., Asian, or more specifically non-hispanic asian alone), non-hispanic other single race, and non-hispanic multiracial.

See EJScreen methodology for details.

# 4 Findings

## 4.1 Text on Findings

### 4.1.1 Results – Basic information about the locations analyzed and number of people nearby

• Explain whether the sites are to be analyzed as a whole or in some type of subgroups, such as 2 different source categories, or large vs small facilities, or some other categories we will use to compare all these stats?

• Count of locations

• Locations missing data

• Where are the sites (regions, states, cities?, urban/rural?)

• Clustering (are they near each other? How close? Which ones are in clusters, maybe?)

• Total population near any of the sites (count of unique residents) and some summary of site-specific Population sizes and population density nearby (what % of all people are at x% of the sites? Range and Interquartile range of pop counts nearby? etc.). Make clear to what extent some people in the overall summary stats are near more than a single one of these sites.

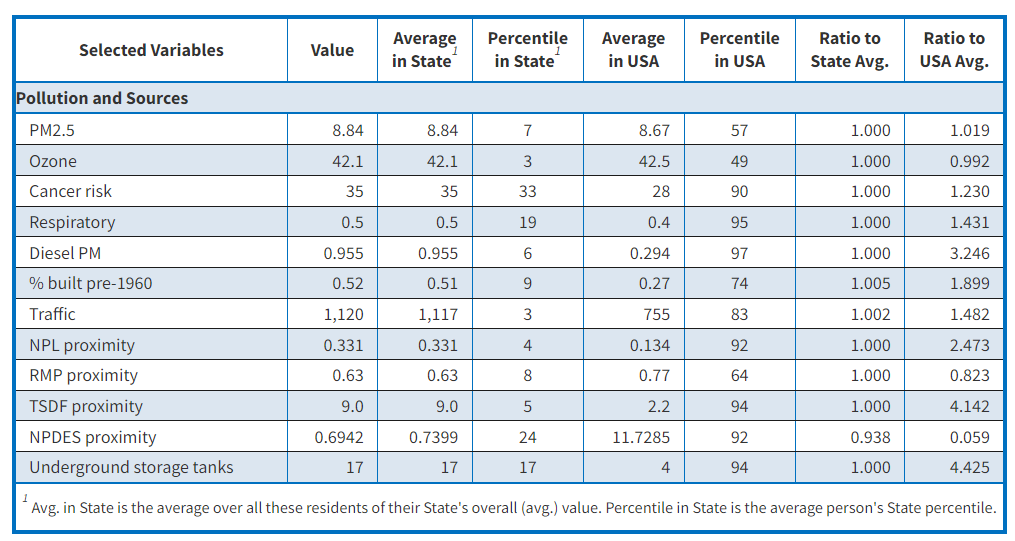
### 4.1.2 Demographics overall

Overall summary statement of some kind - e.g., a “\_\_\_very large” number of all the (12?) envt (or D) indicators were “\_\_\_very high” at a “\_\_\_very large” share of the sites or preferably for a “\_\_\_very large share of the people”? Or, how many of 12 indicators were “high” for the average person and/or site overall? Somehow put that in context ?? vs other rules, usA, other sites, etc.???

discussion of table here

discussion of table here

## 4.2 Data Table 1. Demographic Indicators



Almost all the EJ-relevant groups (low-income, people of color, etc.) are at least somewhat over-represented near these sites overall (at the collection of sites as a whole).

Most notably, % with limited English, % low income, and % with less than high school education near these sites are about 1.5 to 1.7 times the US overall rates.

The people living near these sites are 40% more likely to be in Limited-English Households than the average US resident.

The % with limited English is driven by high scores at only a few very highly population sites – it is high enough that the rate overall is in the top quintile nationally (83 percentile), but other demographics do not reach the top quintile for the entire population across all sites as a whole.

Near most of these 72 sites, % low income is at least 1.3x the rate in the US overall, and near 1 in 4 it is at least 1.5x the US rate.

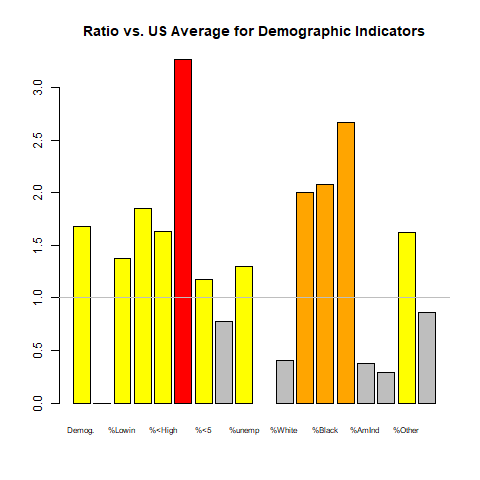
discussion of plot goes here

discussion of plot goes here

discussion of plot goes here

discussion of plot goes here

## 4.3 Data Viz 2– Barplot



### 4.3.1 Key D group(s) based on policy or default / Key D group(s) based on observed magnitude of disparities here

• Sdf

• Asdf

• Asdf

Largest disparity in presence of any group – which group was most over-represented here?

* Relative disparity vs US: Which group(s) had the largest ratio of local % (for the overall set of sites) to US % overall?
* Absolute disparity vs US: which group(s) had

“Large” disparities: which groups were “very” over-represented here?

* Which indicator has max mean percentile? And how high is that?
* Which indicator has largest ratio to US? To state? And how large is that?

### 4.3.2 Demographics at key sites

Overall summary statement of some kind - e.g., a “\_\_\_very large” number of all the (12?) envt (or D) indicators were “\_\_\_very high” at a “\_\_\_very large” share of the sites or preferably for a “\_\_\_very large share of the people”? Or, how many of 12 indicators were “high” for the average person and/or site overall? Somehow put that in context ?? vs other rules, usA, other sites, etc.???

Demographic stats on the distribution across sites/people

• What % or count of people/sites have score > key cutoffs?

• What is mean or median or 95th pctile score of people/sites ?

• NOTABLE “WORST” SITES NAMES?

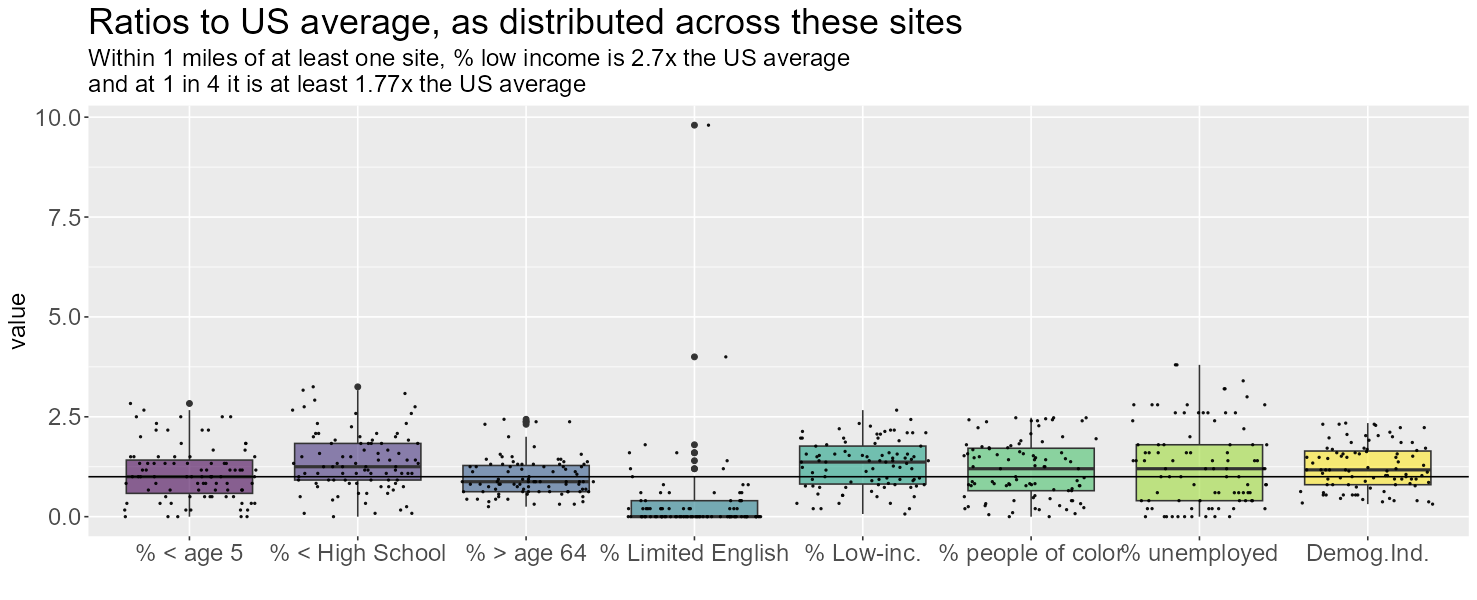
There are a handful of sites each with at least one very high demographic stat within 1 mile, however this may be within the normal range of what one would expect across the range of US residential areas -- there does not appear to be a pattern of an unusually large share of these 72 sites having any given demographic stat in the top 5%, for example.

Seven key sites have at least some demographic percentages more than 2x the US average: Crawford in Chicago IL (densely populated location), Bay Shore in Ohio, Watts Bar Fossil Plant in Spring City TN (but has almost no nearby residents), Arkwright in Macon GA, Venice in IL, Lake Shore in Cleveland, and Fair Station in Muscatine IA.

At two sites, percent people of color and percent low income are both more than twice the US averages (Venice and Lake Shore).

A few sites have over triple the US average unemployment rate (two of which are in the top 5% nationwide for their rate of unemployment, Bay Shore and Watts Bar).

## 4.4 Data Viz 3 – Boxplots



### 4.4.1 Environment overall

The people living nearby these sites as a whole are facing relatively high exposure to indicators of RMP proximity and possible lead paint due to older buildings. Overall the average person nearby has RMP proximity more than 3 times the US average. Lead paint and traffic are also notable, at more than 2x the US average. The average person near any of these sites lives in a blockgroup that is at the 80th percentile (worst quintile) of the US for RMP and lead paint – that is unusual because it is a pattern for these populations as a whole not just at one site. The same is almost true for traffic and UST – the average person nearby has those indicators in the worst quartile of the US. (Wastewater and Superfund also tend to be very high at an unusually large share of these sites but not necessarily at the ones in highly populated areas).

• Overall summary statement of some kind. e.g., a “\_\_\_very large” number of all the (12?) envt (or D) indicators were “\_\_\_very high” at a “\_\_\_very large” share of the sites or preferably for a “\_\_\_very large share of the people”? Or, how many of 12 indicators were “high” for the average person and/or site overall? Somehow put that in context ?? vs other rules, usA, other sites, etc.???

• Which indicator has max mean percentile? And how high is that?

• Which indicator has largest ratio to US? To state? And how large is that?

discussion of table here

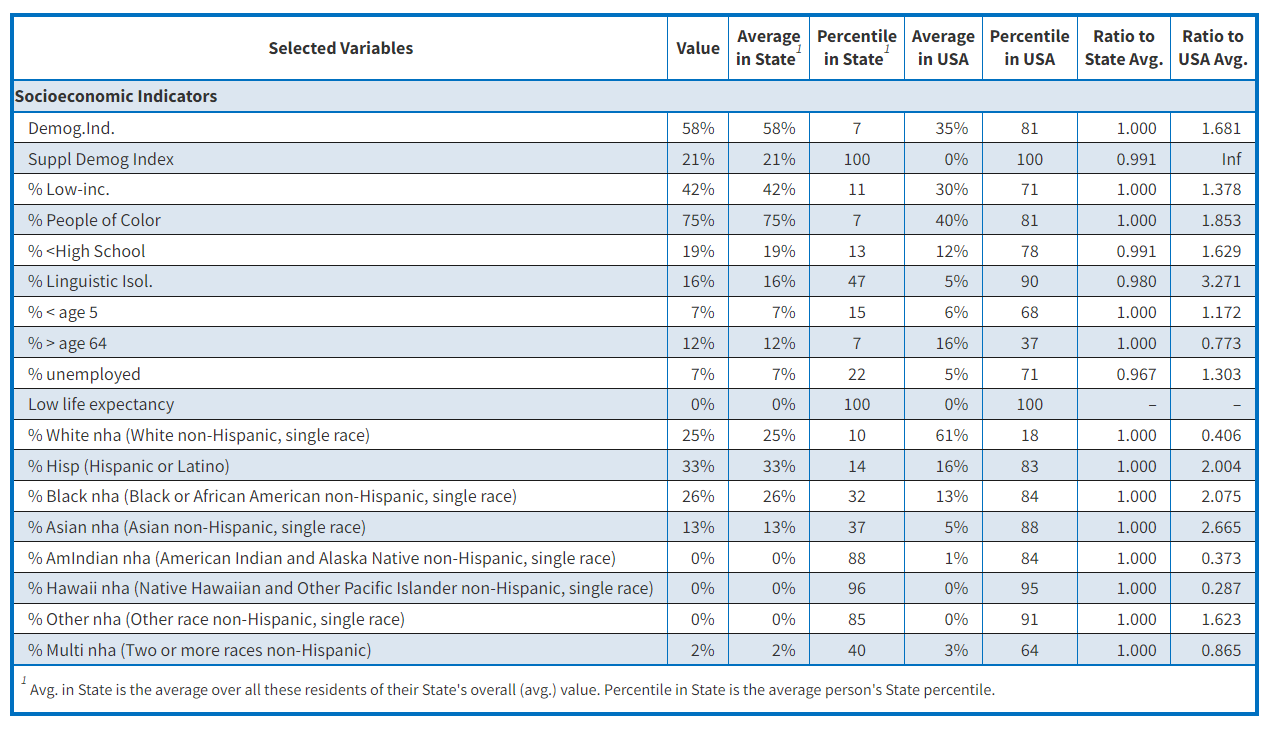
discussion of table here

discussion of table here

discussion of table here

discussion of table here

## 4.5 Data Table 2. Environmental Indicators



### 4.5.1 Environmental indicators distribution across the residents and sites

• What % or count of people/sites have score > key cutoffs?

• What is mean or median or 95th pctile score of people/sites ?

• NOTABLE “WORST” SITES NAMES?

### 4.5.2 Environment at key sites

There are an unusually large number of sites with very high environmental stats within 1 mile. Surprisingly, 27 of these 72 sites have at least one above the 95th percentile. One might expect 5% of these sites (i.e., 3.6 sites) to have a given score in the top 5% nationwide, but there are 8 sites (2x what one might expect) with RMP proximity scores in the top 5%, and the same is true for Superfund NPL proximity scores (there are 8 sites >=95th %ile instead of the expected 3.6). For the wastewater discharge indicator, there are 11 such sites, over 3x as many as one might expect. Most sites have at least one environmental indicator >80th percentile.

### 4.5.3 Cumulative impacts at key sites

Multiple environmental stressors are also an issue in some cases - At two of the sites, there are five environmental indicators that are all more than twice the US average (Valley and Crawford).

discussion of map

discussion of map

discussion of map

discussion of map

discussion of map

### 4.5.4 Combination of demographic and environmental conditions in these locations

Summary of Envt and Demog across all indicators? In other words, a single score like the combination of 12 EJ indexes as a threshold approach summary? e.g. the average person had 5 or more of the 12 EJ indexes at least at the 95th percentile nationwide?

• Individual EJ indicator of 12 that generally had highest overall percentiles? On average highest vs the one where at least a few sites had a very high number? Other?

• User specified EJ index(es)?

# 5 Appendices

### 5.0.1 Appendix 1 - Detailed tables of statistics for each site

e.g. how many of 12 are >x? does it have any high scores for any E? any D? etc.

### 5.0.2 Appendix 2 - Detailed tables of statistics for each indicator

#### 5.0.2.1 \*\*What is Score of people/sites at key cutoffs: What is the X (useful) percentile

value, for this indicator (as %ile of people nearby or of sites)?\*\*

* min
* min other than zero?
* mean
* 25th %ile of these sites or people (25% have lower than this, and 75% have higher than this – those are same if no ties with this value, but can differ if multiple places have this same exact indicator score)
* median (half of these sites or people have a score that is >= this, and half have <= this)
* 75th %ile of these sites or people (if a tied value, would want both >75 and <25%)
* max

#### 5.0.2.2 **How many people/sites have score > key cutoffs:  What % & what # of sites & people have this indicator score (raw or percentile) >= x (useful cutoff)?**

* **For percentiles**
  + % of sites w data where >=80
  + % of sites w data where >=90
  + % of sites w data where >=95
  + # of sites where >=95
* **For ratios to State or USA average overall % of sites w data where >1 not=1**
  + % of sites w data where ratio is >=2
  + % of sites w data where >=3
  + % of sites w data where >=5
  + % of sites w data where >=10
  + # of sites w data where >=10
* **for 12 EJ indexes, how many of the indexes are above a given threshold?**
  + # of sites w data where >=1
  + # of sites w data where >1 not =1
  + # of sites w data where >=4

### 5.0.3 Appendix 3 - notes on how to describe places generically/ in parameters

HOW TO REFER TO THE PLACES STUDIED (near these facilities vs more general language)

Note: this could be a “proximity analysis” in cases where it relies on circular buffers around facility points to define buffers that include residents within a fixed distance from one or more facilities/sites. But it more generally could be an EJ analysis that describes environmental and demographic conditions among residents in any specified places, such as all the places where air quality modeling suggests risk is currently above 1 in 1 million, for example. So the language should be flexible and refer to something like this:

FOR NON-PROXIMITY ANALYSIS, GENERALLY ANY KIND OF BUFFERS/PLACES ANALYZED:

The demographics of residents in …

The demographic / environmental indicators in …

Percent low income among residents in …

The environmental conditions in …

The environmental indicators for the average resident in …

The PM2.5 levels in …

Residents in …

… these locations

… these places

… these areas

… the analyzed locations

FOR PROXIMITY ANALYSIS SPECIFICALLY, EASIER TO SAY one of these:

Residents / conditions …

… within x miles of these sites…

… within x miles of any of these sites…

… nearby

… near these sites

… near any of these sites

## 5.1 Acknowledgements

## 5.2 List of Abbreviations

# References

R Core Team. 2022. *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing. <https://www.R-project.org/>.