

Exploring Environmental Injustice: Air Quality

Leela Dixit

2025-10-18

Exploring Environmental (in)justice

The United States Environmental Protection Agency defines Environmental Justice as “the just treatment and meaningful involvement of all people, regardless of income, race, color, national origin, Tribal affiliation, or disability”. In this project, I will be exploring particulate pollution in California and exploring which groups might be disproportionately affected by poor air quality.

Step 1: Load in necessary packages.

```
rm(list = ls())
library(tidyverse)
library(sf)
library(here)
library(tmap)
```

Step 2: Read in environmental justice data and clean for mapping.

```
# read in geodatabase of EJScreen data at the Census Block Group level
ejscreen <- sf::st_read(here::here("data", "data",
                                "ejscreen",
                                "EJSCREEN_2023_BG_StatePct_with_AS_CNMI_GU_VI.gdb"))

# filter to California only
california <- ejscreen %>%
  dplyr::filter(ST_ABBREV == "CA") %>%
```

```

drop_na(c("P_D2_CANCER", "P_PTRAF", "LIFEEXPPCT"))

# calculate the mean for each EJ variable, by county
ca_counties <- aggregate(california, by = list(california$CNTY_NAME), FUN = mean)

# separate the county names into name and county, and remove the "county" string
ca_counties <- separate(ca_counties, Group.1,
                        into = c("county", "county_name"), sep = " ")
ca_counties <- select(ca_counties, -c("county_name"))

```

Step 3: Map 1

The map below represents the estimated lifetime inhalation cancer risk as a result of carcinogens present in the air. This value is persons per million over a 70-year lifetime. EJScreen presents these data as a percentile for better comparison. Each inset map is one of the top three highest counties for lifetime inhalation cancer risk.

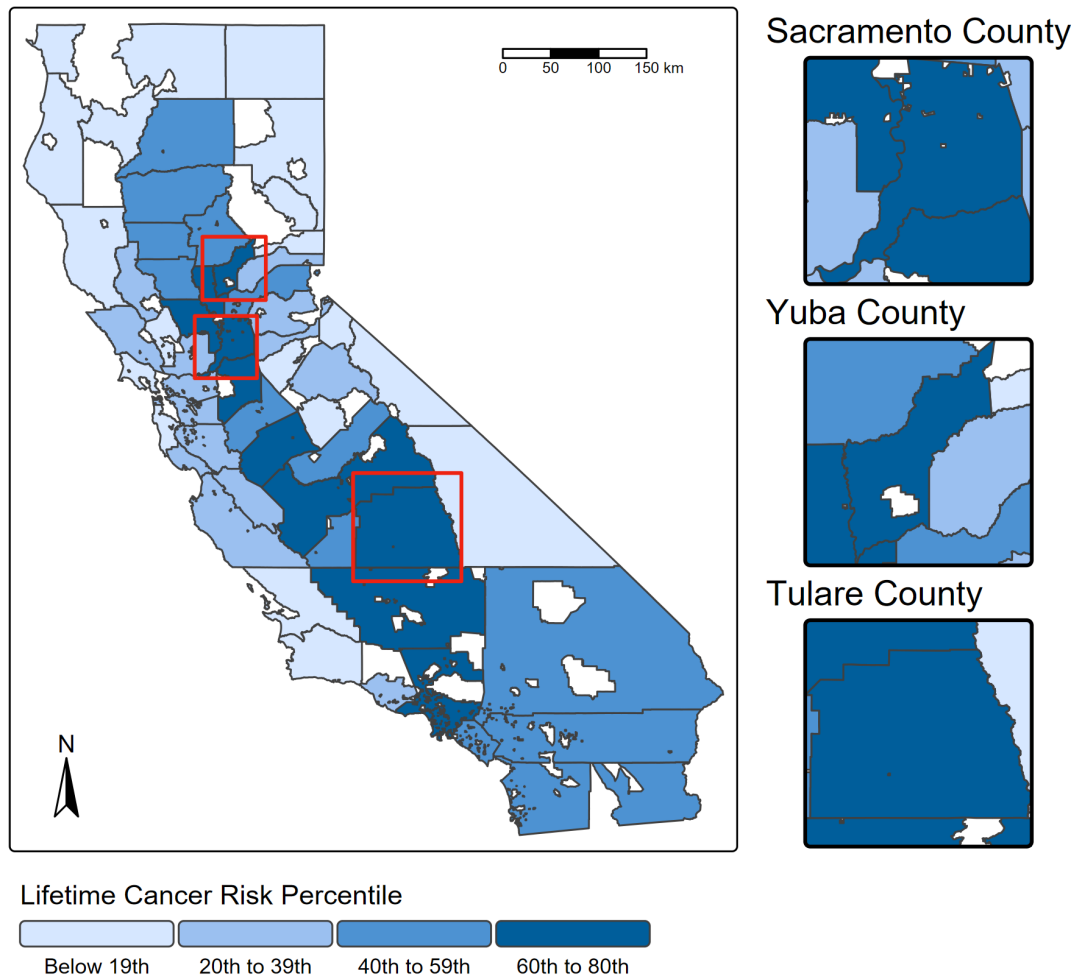
```

# percentile of airborne toxics that can lead to cancer, by county in California
counties_map <- tm_shape(ca_counties) +
  tm_polygons(fill = "P_D2_CANCER",
              fill.legend = tm_legend(title = "Lifetime Cancer Risk Percentile",
                                      orientation = "landscape",
                                      position = tm_pos_out("center", "bottom"),
                                      frame = FALSE),
              fill.scale = tm_scale(labels = c("Below 19th", "20th to 39th",
                                              "40th to 59th", "60th to 80th"))) +
  tm_title("Lifetime cancer risk via inhalation of toxins, by county",
           frame = FALSE) +
  tm_scalebar(position = "right") +
  tm_compass(position = c("left", "bottom")) +
# adding inset maps for top 3 counties with highest airborne toxics levels
tm_title("Sacramento County", group_id = 1) +
  tm_inset(tmaptools::bb("Sacramento County"), group_id = 1) +
tm_title("Yuba County", group_id = 2) +
  tm_inset(tmaptools::bb("Yuba County"), group_id = 2) +
tm_title("Tulare County", group_id = 3) +
  tm_inset(tmaptools::bb("Tulare County"), group_id = 3) +
  tm_components(1, position = tm_pos_out()) +
  tm_components(2, position = tm_pos_out()) +
  tm_components(3, position = tm_pos_out())

```

```
# save the plot as a png
tmap_save(counties_map, here::here("figs", "ca_counties.png"))
```

Lifetime cancer risk via inhalation of toxins, by county



This map shows that counties of the central valley (San Joaquin Valley) are generally in

a higher percentile for cancer risk via inhalation of airborne toxins. There could be a few potential causes of increased carcinogens in outside air:

- Nearby I-5 and Highway 99 introduce pollution sources, such as diesel locomotives.
- Vehicle emissions from these roadways, tractors and irrigation pumps from local agriculture, and wood burning stoves/fireplaces all contribute to fine particulate matter (PM2.5).
- Topography from surrounding mountains create a “bowl”, trapping these pollutants.

Source: [EPA Activities for Cleaner Air](#)

Step 4: Maps 2 and 3

We will zoom in to a few counties to get a more detailed view of any patterns that might emerge between airborne toxics and underrepresented communities.

```
# zoom in on a few higher risk counties from the cancer index to look at things
# like PM2.5, diesel, and further demographic information

# sort them and choose top 3
ca_counties %>%
  select(county, P_D2_CANCER) %>%
  arrange(desc(P_D2_CANCER))
# In order: Sacramento, Tulare, Yuba Counties

# subset full data set for top 3
top3 <- c("Sacramento County", "Tulare County", "Yuba County")
top3_county <- california %>%
  filter(CNTY_NAME %in% top3) %>%
  rename("Cancer Risk" = "P_D2_CANCER",
         "Diesel Particulate Matter" = "P_D2_DSLPM",
         "PM 2.5" = "P_D2_PM25",
         "People of Color" = "P_PEOPCOLORPCT",
         "Low Life Expectancy" = "P_LIFEEXPPCT",
         "Low Income" = "P_LOWINCPCT")

# subset full data for just Sacramento County,
# the highest inhalation cancer risk county
top3_sac <- california %>%
  filter(CNTY_NAME == "Sacramento County") %>%
  rename("Cancer Risk" = "P_D2_CANCER",
         "Diesel Particulate Matter" = "P_D2_DSLPM",
```

```

"PM 2.5" = "P_D2_PM25",
"People of Color" = "P_PEOPCOLORPCT",
"Low Life Expectancy" = "P_LIFEEXPPCT",
"Low Income" = "P_LOWINCPCT")

```

We will first take a look closer into the three counties with the highest inhalation cancer risk. In order: 1. Sacramento County 2. Tulare County 3. Yuba County

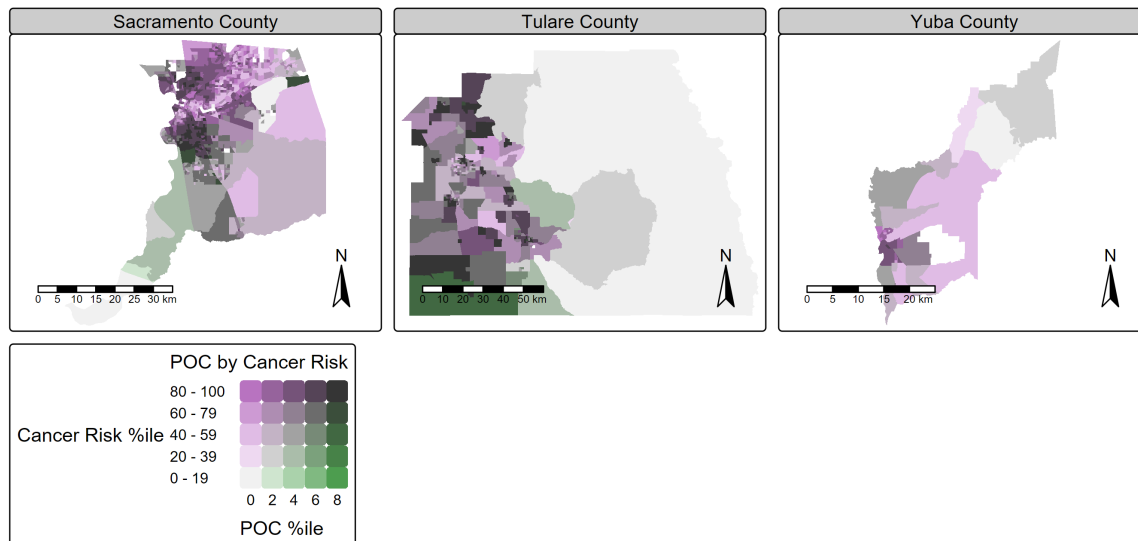
```

# bivariate plot of cancer risk by POC for top 3 counties
bivariate <- tm_shape(top3_county) +
  tm_fill(tm_vars(c("Cancer Risk", "People of Color"),
    multivariate = TRUE),
    fill.legend = tm_legend_bivariate(title = "POC by Cancer Risk",
      xlab = "POC %ile",
      ylab = "Cancer Risk %ile"),
    fill.scale = tm_scale(ylabels = c("Below 19th", "20th to 39th",
      "40th to 59th", "60th to 80th")))) +
  tm_facets("CNTY_NAME") +
  tm_title("Percentile of People of Color by Percentile of Inhalation Cancer Risk
(top 3 highest risk counties)") +
  tm_scalebar(position = c("left", "bottom")) +
  tm_compass(position = c("right", "bottom"))

# save the plot as a png
tmap_save(bivariate, here::here("figs", "bivariate.png"))

```

Percentile of People of Color by Percentile of Inhalation Cancer Risk (top 3 highest risk counties)



From these maps, we can see that there is a co-occurrence of inhalation cancer risk and high percentile of people of color that live in those areas. We can also see that there are not many areas in these counties where there is a high number of people of color and NOT a high inhalation cancer risk, meaning areas with low inhalation cancer risk are most often populated by white residents. This poses an environmental justice issue of people of color being disproportionately exposed to levels of airborne toxins that lead to higher cancer risk through inhalation. Further exploration into other population metrics, such as income level or factors that lead to healthcare accessibility, would be important to look into to get a bigger picture of patterns that could be occurring.

Lastly, we will look closer into just Sacramento County, the county with the highest inhalation cancer risk. We will look at further risk factors such as PM2.5 and diesel particulates, both of which contribute to inhalation risks, along with more environmental justice metrics (low life expectancy and low income).

```
# pivot variables in data frame in order to facet wrap each variables map easily
top3_sac_long <- top3_sac %>%
  select("CNTY_NAME", "Cancer Risk", "Diesel Particulate Matter", "PM 2.5",
         "People of Color", "Low Life Expectancy", "Low Income") %>%
  pivot_longer(col = c("Cancer Risk", "Diesel Particulate Matter", "PM 2.5",
                      "People of Color", "Low Life Expectancy", "Low Income"),
              names_to = "Metric", values_to = "value")

# order the metrics using factor() by how they should be faceted
```

```

top3_sac_long$Metric <- as.factor(top3_sac_long$Metric)
top3_sac_long$Metric <- factor(top3_sac_long$Metric,
                              levels = c("Cancer Risk",
                                           "Diesel Particulate Matter",
                                           "PM 2.5", "People of Color",
                                           "Low Life Expectancy", "Low Income"))

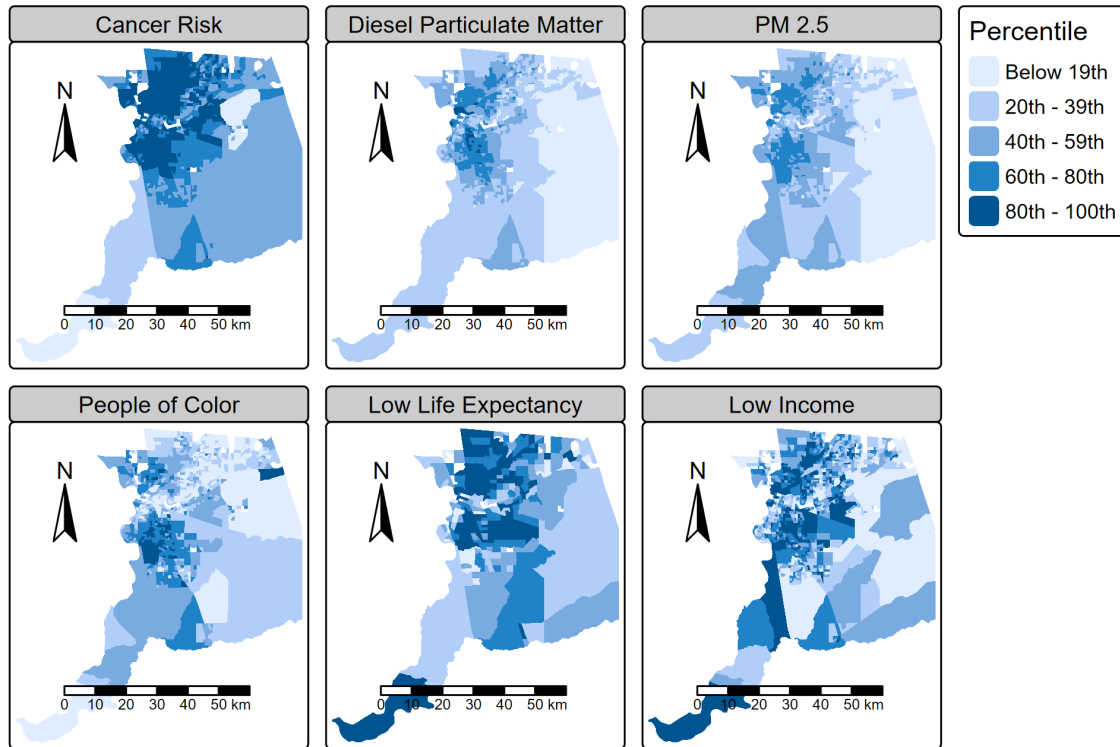
# facet maps for 3 air quality metrics and 3 environmental justice metrics
facet <- tm_shape(top3_sac_long) +
  tm_fill(fill = "value",
          fill.legend = tm_legend(title = "Percentile"),
          fill.scale = tm_scale(labels = c("Below 19th", "20th - 39th",
                                           "40th - 59th", "60th - 80th",
                                           "80th - 100th")) +

  tm_facets(by = "Metric") +
  tm_title("Air Quality and Environmental Justice Metrics for Sacramento County") +
  tm_scalebar(position = tm_pos_in("center", "bottom")) +
  tm_compass(position = tm_pos_in("left", "top"))

# save the plot as a png
tmap_save(facet, here::here("figs", "facet.png"), width=2000)

```

Air Quality and Environmental Justice Metrics for Sacramento County



We can see the additional air quality metrics (diesel airborne particulate matter and PM2.5) both congregate in high percentiles in similar locations as lifetime inhalation cancer risk. This makes sense, as both additional air quality metrics likely contribute to overall inhalation cancer risks. We can see our environmental justice metrics we are looking at, percentiles of people of color, low life expectancy, and low income, are also highest in similar locations as our air quality metrics. This further shows that disadvantaged groups could be experiencing higher health risks due to their locations, potentially more difficult time receiving medical care because of income status, and possibly contributing to a lower life expectancy.

Step 5: Future directions

Exploring causal relationships between these metrics and looking into additional metrics (both environmental justice focused and other air quality or environmental metrics) could help tease apart the potential pattern we see here, and lead to more effective policy and research in line with creating just environmental solutions for those underrepresented in California.