

Homework Assignment 1

Cancer Risk in California’s East Bay

AUTHOR
Jay Kim

PUBLISHED
October 6, 2025

Introduction

Have you ever been to the Bay Area? San Francisco? San Jose? These are much more popular destinations. I would like to discuss a part of the Bay Area that is not as glamorous and somewhat removed from the tech bubble.

This project will use maps to illustrate the following:

1. Retrieve and load the data, subset the data, and save the data to minimize the time spent on the rendering process
2. Plot and render a map that shows the East Bay, which consists of Alameda, Contra Costa, and Solano Counties
3. Understand the plots as we examine them together
4. We will explore the lifetime health risks associated with air quality, demographics (including people of color), and socioeconomic levels.

Initial Setup

```
# Load in all the Necessary Packages for upcoming tasks
pacman::p_load(tidyverse, sf, stars, tmap, here, viridisLite, RColorBrewer)

# Read in EJScreen data
ejscreen <- sf::st_read(here::here("data",
                                   "ejscreen",
                                   "EJSCREEN_2023_BG_StatePct_with_AS_CNMI_GU_VI.gdb"),
                      quiet = TRUE)

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

# Filter to a counties in Bay Area (East Bay)
east_bay_ejscreen <- ejscreen %>%
  dplyr::filter(CNTY_NAME %in% c("Alameda County",
                                "Solano County",
                                "Contra Costa County"))

# Save the subset as a new `gpkg` file
st_write(california,
         here("data",
              "ejscreen",
              "california.gpkg"),
         quiet = TRUE)
st_write(east_bay_ejscreen,
         here("data",
              "ejscreen",
              "east_bay_ejscreen.gpkg"),
         quiet = TRUE)

# Read subsetted data that's been saved as `gpkg` files
california <- st_read(here("data",
                          "ejscreen",
                          "california.gpkg"),
                    quiet = TRUE) %>%
  st_transform(3857)

east_bay_ejscreen <- st_read(here("data",
                                "ejscreen",
                                "east_bay_ejscreen.gpkg"),
                          quiet = TRUE) %>%
  drop_na("P_D2_CANCER") %>%
  st_transform(3310)
```

Maps

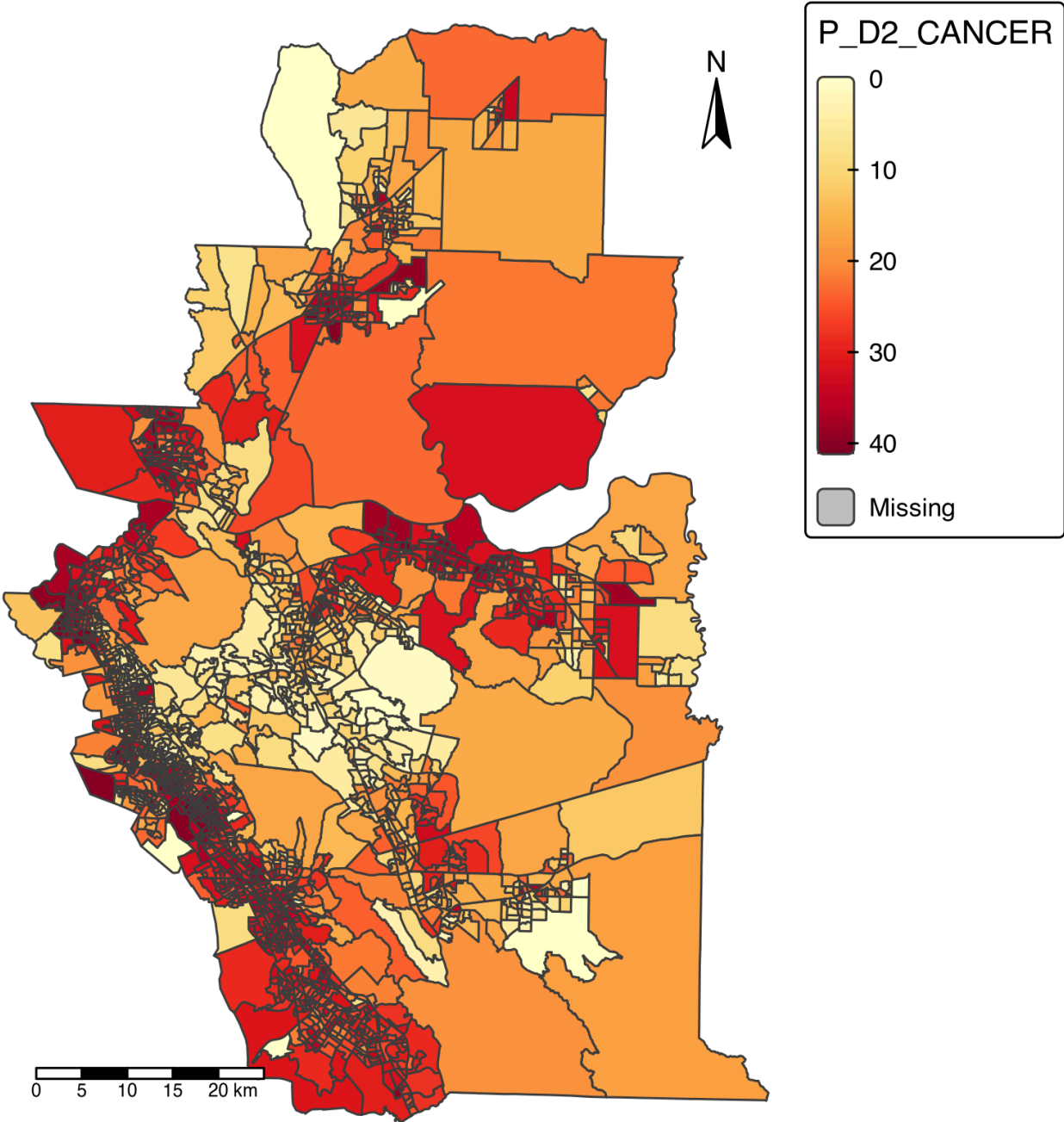
```
bbox_eastbay <- st_bbox(east_bay_ejscreen)

# Create static map
tmap_mode("plot")

# Plotting a spatial object
map_p_cancer <- tm_shape(east_bay_ejscreen, bbox = bbox_eastbay) +
  tm_polygons(fill = "P_D2_CANCER",
              palette = "YlOrRd",
              style = "cont",
              fill.legend = tm_legend(title = "Percentile for Cancer Risk")) +
  tm_layout(legend.outside = TRUE, frame = FALSE) +
  tm_scalebar(position = c("left", "bottom")) +
  tm_compass(type = "arrow", position = c("right", "top")) +
  tm_title("Cancer Risk from Air Toxics (East Bay Counties)")

tmap_save(map_p_cancer,
          filename = here("output", "map_p_cancer.png"),
          width = 6,
          height = 6,
          units = "in",
          dpi = 300)
```

Cancer Risk from Air Toxics (East Bay Counties)



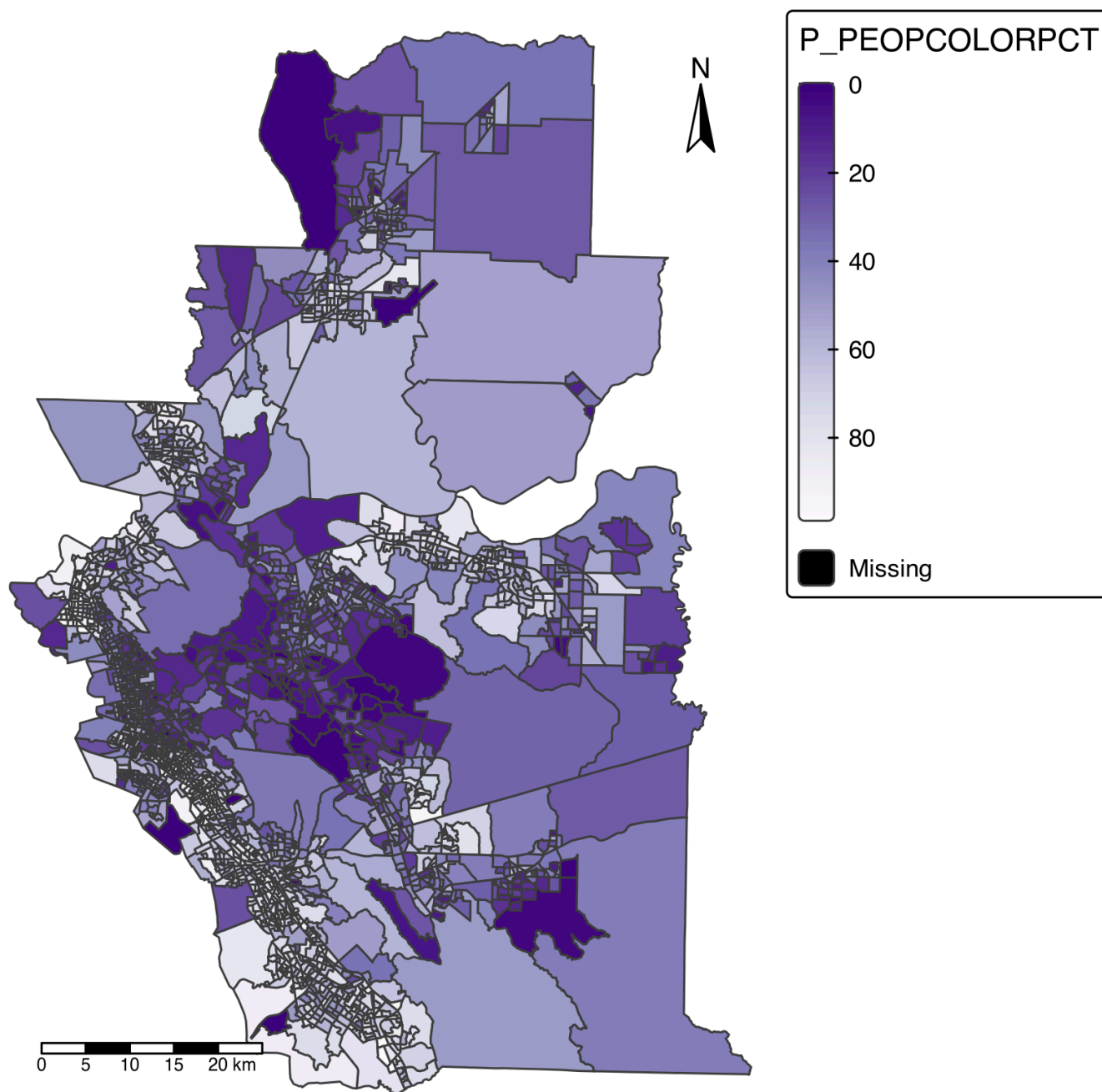
Map of health risk due to air quality

```
# Create static map
tmap_mode("plot")

# Plotting a spatial object
map_p_poc <- tm_shape(east_bay_ejscreen, bbox = bbox_eastbay) +
  tm_polygons(fill = "P_PEOPCOLORPCT",
              palette = "-Purples",
              style = "cont",
              fill.legend = tm_legend(
                title = "Percentile for People of Color")) +
  tm_layout(legend.outside = TRUE, frame = FALSE) +
  tm_scalebar(position = c("left", "bottom")) +
  tm_compass(type = "arrow", position = c("right", "top")) +
  tm_title("Demographic Vulnerability (East Bay Counties)")
```

```
tmap_save(map_p_poc,
  filename = here("output", "map_p_poc.png"),
  width = 6,
  height = 6,
  units = "in",
  dpi = 300)
```

Demographic Vulnerability (East Bay Counties)



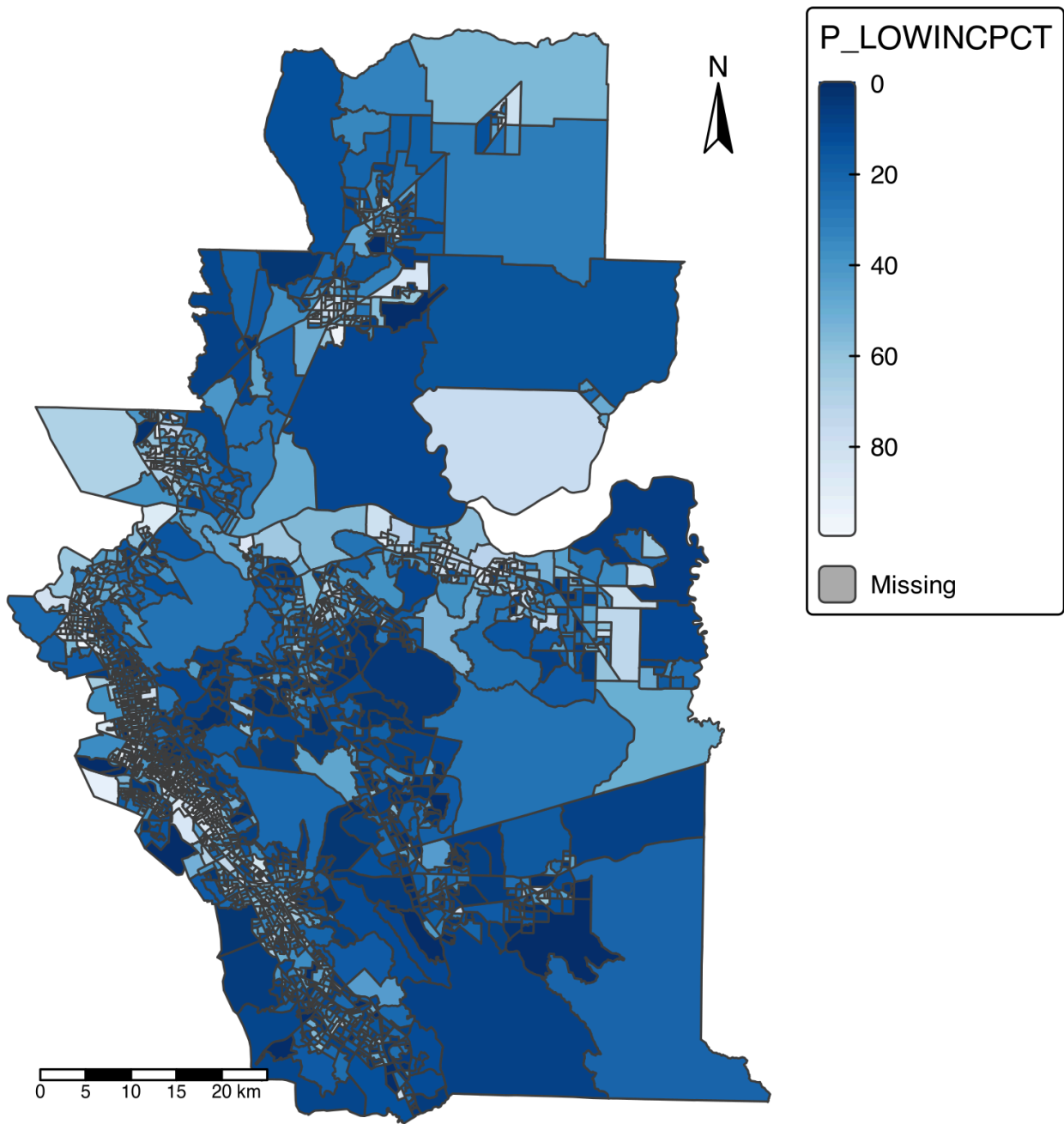
Map of demographic (people of color)

```
# Create static map
tmap_mode("plot")

# Plotting a spatial object for Low-Income Percentage
map_low_income <- tm_shape(east_bay_ejscreen, bbox = bbox_eastbay) +
  tm_polygons(fill = "P_LOWINCPCT",
    palette = "-Blues",
    style = "cont",
    fill.legend = tm_legend(
      title = "Percentile for Low-Income Population")) +
  tm_layout(legend.outside = TRUE, frame = FALSE) +
  tm_scalebar(position = c("left", "bottom")) +
  tm_compass(type = "arrow", position = c("right", "top")) +
  tm_title("Low-Income Residents (East Bay Counties)")

# Save map
tmap_save(map_low_income,
  filename = here("output", "map_low_income.png"),
  width = 6,
  height = 6,
  units = "in",
  dpi = 300)
```

Low-Income Residents (East Bay Counties)

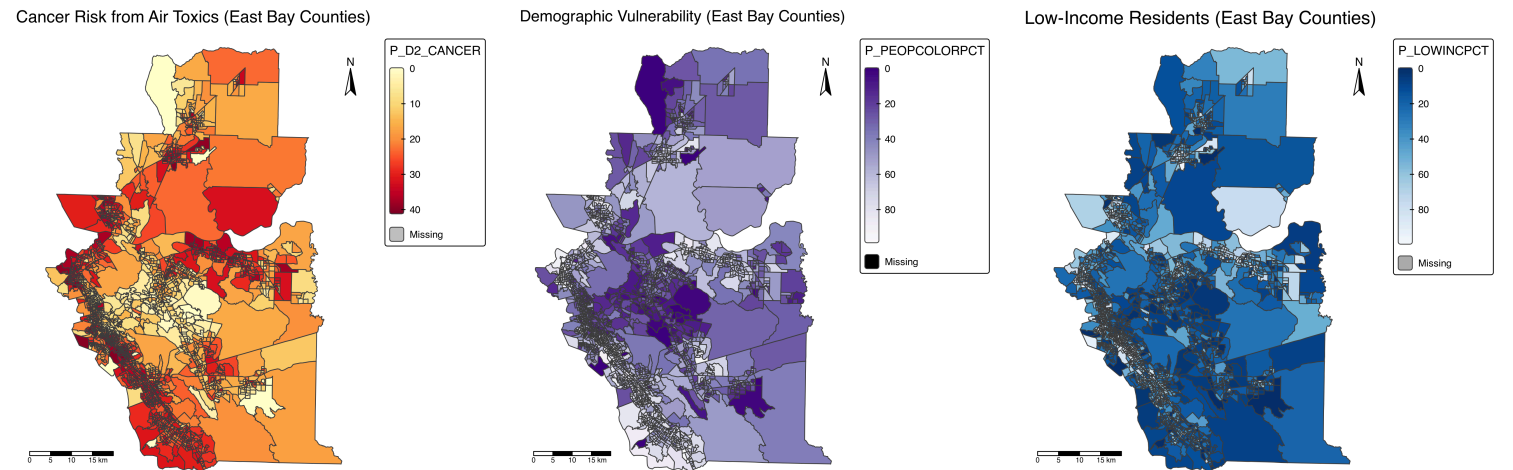


Map of low-income communities

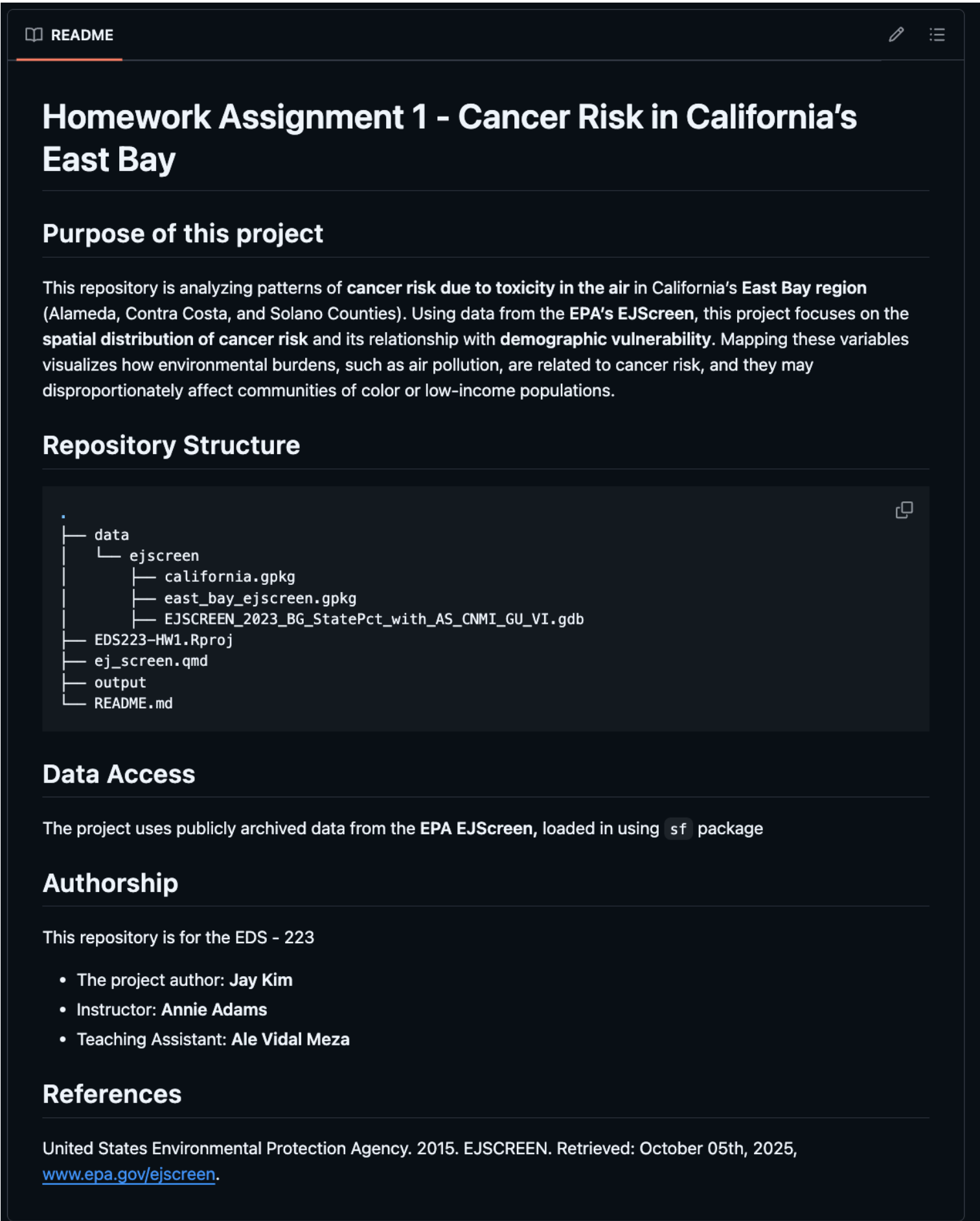
```
# Arrange two maps side-by-side
three_map <- tmap_arrange(map_p_cancer,
  map_p_poc,
  map_low_income,
  ncol = 3,
  widths = c(1, 1, 1))

# Save map as PNG file
tmap_save(three_map,
  filename = here("output", "three_map.png"),
  width = 18,
  height = 6,
  units = "in",
  dpi = 300)
```

The maps below convey a spatial relationship between race, income, and environmental vulnerability across East Bay counties in California. Areas with higher percentages of people of color and the low-income population overlap with regions that have elevated cancer risk. The environmental burden is disproportionately heavier in the places where marginalized communities are, related to their socioeconomic status and racial composition. The map clarifies how environmental injustice occurs in the local vicinity of all other areas, except retirement communities in the northern part of the East Bay.



Comparison of three maps



Screenshot of README