Exploring Patterns of Environmental Justice

Hylaea Miller

2025-10-18

Background

This assignment is part of the course EDS 223: Geospatial Analysis & Remote Sensing, which aims to explore patterns of environmental justice, with focus on:

- Practicing manipulating vector and raster data to build multi-layer maps.
- Practicing making maps in R, specifically using tmap.
- Building effective, responsible, accessible and aesthetically-pleasing maps.

Data

There were three datasets used in this assignment:

- **EJSCREEN**, which is an environmental justice mapping and screening tool that provides the EPA with a nationally consistent dataset and approach for combining environmental and demographic indicators. The data are available through the U.S. Environmental Protection Agency and can be found here: Ejscreen
- **HOLC Redlining**, which is part of the Mapping Inequality project, includes Home Owners' Loan Corporation (HOLC) rated neighborhoods based on their perceived safety for real estate investment, using a ranking system for Los Angeles. The data can be found here: Mapping Inequality.
- **Biodiversity observations**, which contain global biodiversity information.The dataset can be found here: Biodiversity observations

Load Packages

```r
# Load libraries
library(sf)
library(here)
library(dplyr)
library(tmap)
library(janitor)
library(spData)
library(ggplot2)
library(patchwork)
library(gt)
```

Import Data

1. Import the three datasets: EJSCREEN, HOLC Redlining (*map_ineq*), and Biodiversity observations (*birds*), and update the column names to snake case using `clean_names()`.

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

# Read in geodatabase of gbif-birds-LA data
birds <- st_read(here("data", "gbif-birds-LA")) %>%
  clean_names()

# Read in geodatabase of mapping-inequality
map_ineq <- st_read(here("data", "mapping-inequality",
                         "mapping-inequality-los-angeles.json"))
```

2. Check if the datasets have the same coordinate reference systems

```r
print(st_crs(map_ineq) == st_crs(birds)) # True
```

```
[1] TRUE
```

```r
print(st_crs(ejscreen) == st_crs(map_ineq)) # False
```

```
[1] FALSE
```

3. Transform coordinate reference systems to match.

```r
# Check if the datasets have the same coordinate reference system.
# If not, print a warning and transform the second dataset to match the first.

if(st_crs(ejscreen) != st_crs(map_ineq)){
  warning("coordinate refrence systems do not match")
  map_ineq <- st_transform(map_ineq, crs = st_crs(ejscreen))
}
```

```r
# Verify that the CRS transformation was successful
print(st_crs(ejscreen) == st_crs(map_ineq)) # True
```

```
[1] TRUE
```

Subset Data

Filter the EJScreen dataset for Los Angeles County.

```r
# Subset data to Los Angeles County
la <- ejscreen %>%
  filter(cnty_name == "Los Angeles County")
```

Part 1: Legacy of redlining in current environmental (in)justice

1. Create a map of historical redlining neighborhoods

```r
# Create a map of historical redlining neighborhoods in Los Angeles.
tm_shape(map_ineq) +
  tm_polygons("grade",
    title = "HOLC Neighborhood Grade",
    palette = c("A" = "#76a865",
                "B" = "#7cb5bd",
                "C" = "#ffff00",
                "D" = "#d9838d"))  +
  tm_title("Historical Redlining in Los Angeles") +
  tm_scale_bar(position = c("right", "top")) +
  tm_compass(position = c("right", "bottom")) +
  tm_basemap("OpenStreetMap", alpha = 0.6) +
  tm_layout(inner.margins = c(0.01, 0.01, 0.01, 0.1),
            component.autoscale = FALSE)
```

Historical Redlining in Los Angeles

HOLC Neighborhood Grade
- A
- B
- C
- D
- Missing
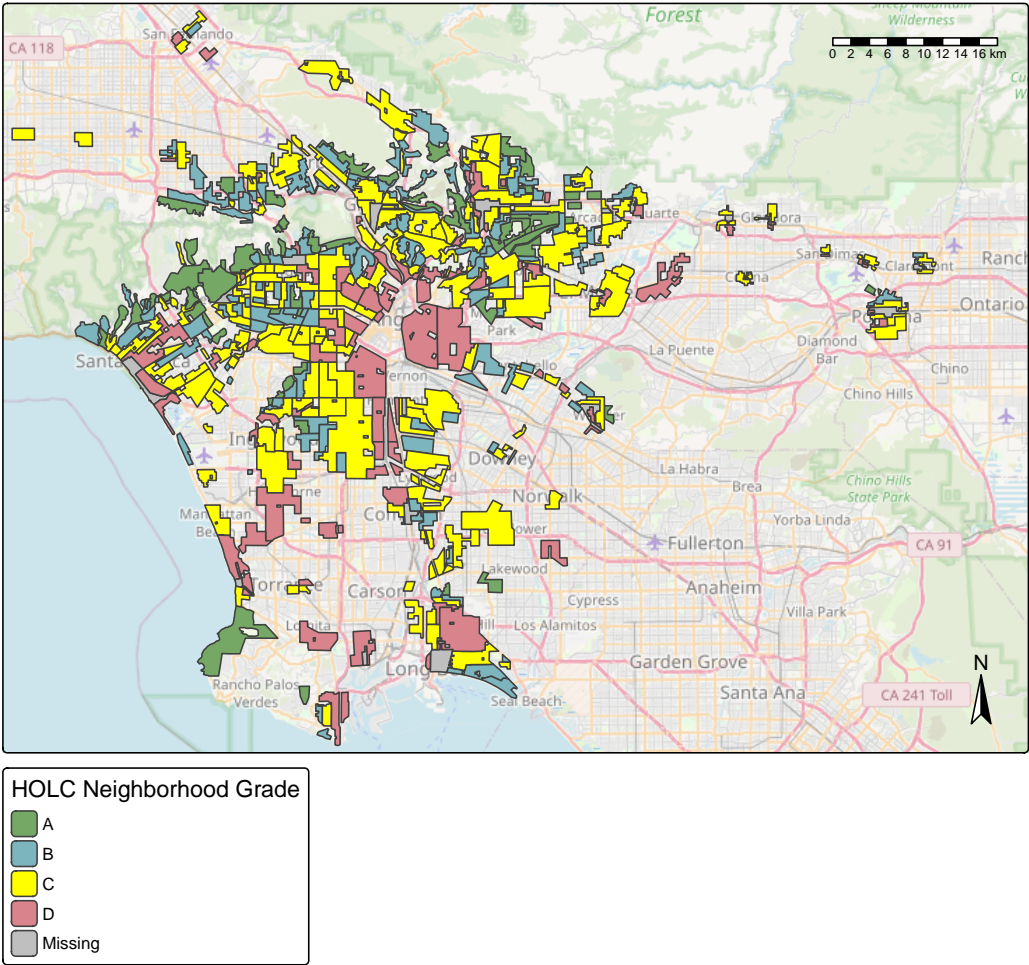
HOLC Grade	Count	Percentage
A	449	7.029
B	1239	19.396
C	3058	47.871
D	1346	21.071
NA	296	4.634

2. Create a table summarizing:

1. Check if the coordinate reference systems match and join the datasets using a spatial intersection.

```
# Check if coordinate reference systems match and add a stop error
if(st_crs(la) != st_crs(map_ineq)){
    stop("coordinate refrence systems must match to join datasets")}

# Join datasets using spatial intersection
la_joined <- st_join(map_ineq, la, join = st_intersects)
```

2. Calculate the percentage of census block groups that fall within each HOLC grade.

```
# Create a summary table with the percentage of census block within each HOLC grade

summary_table <- la_joined %>%
  group_by(grade) %>%
  summarize(n = n()) %>%
  mutate(percentage = round(n / sum(n) * 100, 3)) %>%
  st_drop_geometry()
```

3. Customize the table using the gt package.

```
# Customize table with gt library
summary_table %>%
  ungroup() %>%
  gt() %>%
  cols_label(
    grade = "HOLC Grade",
    n = "Count",
    percentage = "Percentage"
  )
```

3. Create at least two visualizations summarizing current conditions (from the EJScreen data) within HOLC grades using the mean of the following variables (you may combine variables or create separate plots):

- % low income
- percentile for Particulate Matter 2.5
- percentile for low life expectancy

1. Group by HOLC grade and calculate the average for each variable we want to plot.

```
# Find the average values of the variables listed above within each HOLC grade
la_mean <- la_joined %>%
  group_by(grade) %>%
  summarise(
    lowincpct = mean(lowincpct, na.rm = TRUE),
    p_pm25 = mean(p_pm25, na.rm = TRUE),
    lifeexppct = mean(lifeexppct, na.rm = TRUE)) %>%
  filter(grade != 'NA')
```
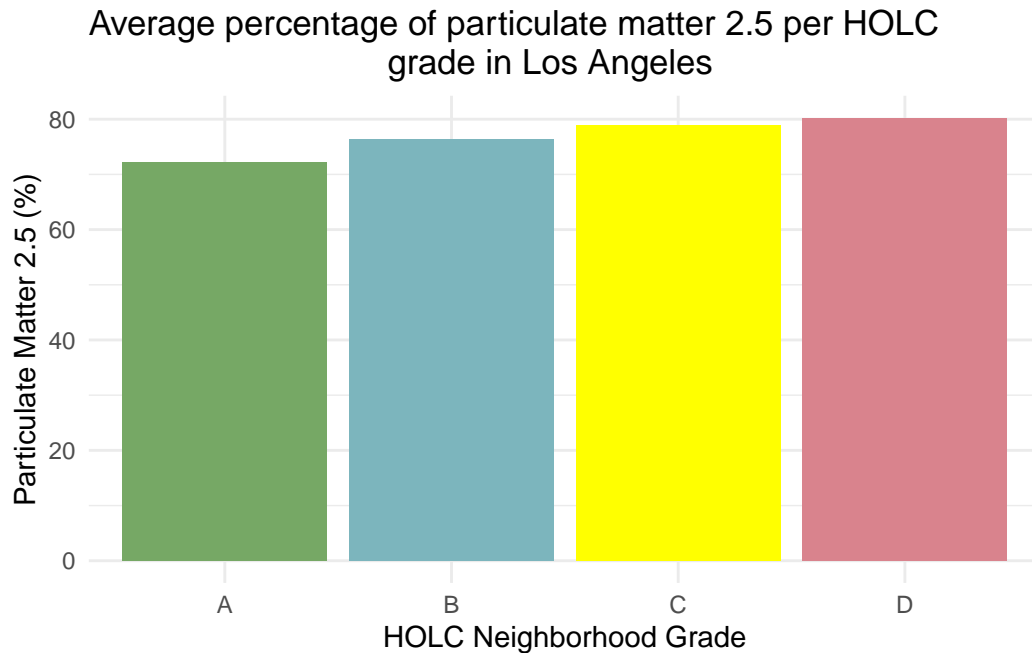
2. Plot the graph of "Average percentage of low-income population per HOLC grade in Los Angeles"

```
# Plot of the percentage of low-income population within HOLC grade
ggplot(data = la_mean, aes(x = grade, y = lowincpct, fill = grade)) +
  geom_col(stat = "identity") +
  scale_fill_manual(values = c("A" = "#76a865",
                               "B" = "#7cb5bd",
                               "C" = "#ffff00",
                               "D" = "#d9838d")) +
  labs(title = "Average percentage of low-income population per HOLC
                          grade in Los Angeles",
       x = "HOLC Neighborhood Grade",
       y = "Low Income (%)") +
  theme_minimal() +
  theme(legend.position = "none")
```

Average percentage of low–income population per HOLC grade in Los Angeles
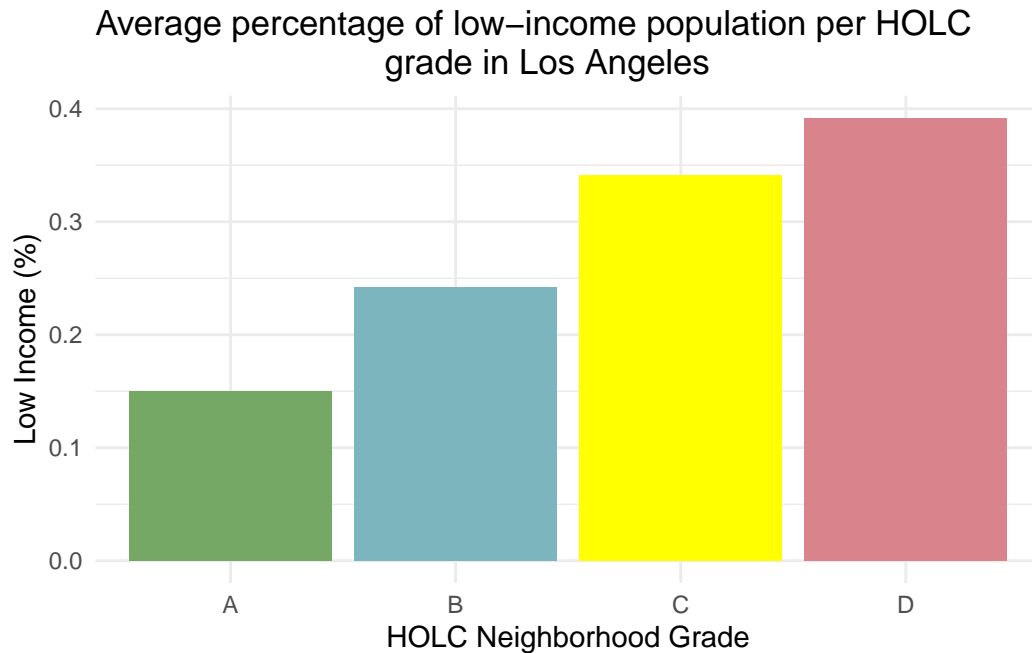
3. Plot the graph of "Average percentage of particulate matter 2.5 per HOLC grade in Los Angeles"

```
# Plot for PM_25 within HOLC grade
ggplot(data = la_mean, aes(x = grade, y = p_pm25 , fill = grade)) +
geom_bar(stat = "identity") +
scale_fill_manual(values = c("A" = "#76a865",
                             "B" = "#7cb5bd",
                             "C" = "#ffff00",
                             "D" = "#d9838d"), ) +
labs(title = "Average percentage of particulate matter 2.5 per HOLC
                       grade in Los Angeles",
    x = "HOLC Neighborhood Grade",
    y = "Particulate Matter 2.5 (%)") +
  theme_minimal() +
  theme(legend.position = "none")
```

Average percentage of particulate matter 2.5 per HOLC grade in Los Angeles

4. Plot the graph of "Average percentage of low life expectancy per HOLCngrade in Los Angeles"

```
# Plot for percentile of low life expectancy within HOLC grade
ggplot(data = la_mean, aes(x = grade, y = lifeexppct , fill = grade)) +
geom_bar(stat = "identity") +
scale_fill_manual(values = c("A" = "#76a865",
                             "B" = "#7cb5bd",
                             "C" = "#ffff00",
                             "D" = "#d9838d"), ) +
labs(title = "Average percentage of low life expectancy per HOLC
                        grade in Los Angeles",
     x = "HOLC Neighborhood Grade",
     y = "low life expectancy (%)") +
  theme_minimal() +
  theme(legend.position = "none")
```

Average percentage of low life expectancy per HOLC grade in Los Angeles
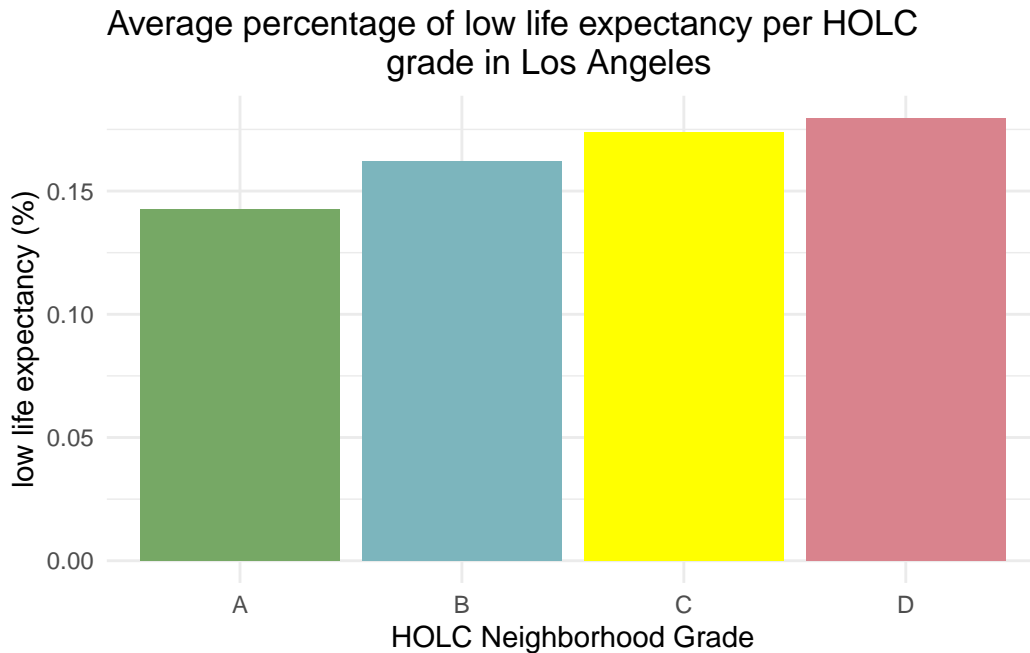
4. Write a brief paragraph reflecting on these results

For every figure, the HOLC neighborhoods with the highest values were the D grade. The first graph shows the average percentage of the low-income population per grade, where Grade D neighborhoods have the largest average number of residents with low income. This result represents a possible reflection of the economic consequences of the redlining policies that these communities went through, such as restricted public investment and access to credit. Grades A and B show the lowest percentages of low-income residents, revealing possible patterns of economic privilege and pointing out the probability that historical housing discrimination still continues to the present day, contributing to wealth inequality.

The second graph shows the average percentage of particulate matter 2.5, where the same pattern from the first figure can still be observed, though with a lower proportion. The grade D neighborhoods show the highest percentage of particulate matter 2.5. This result suggests possible environmental injustice, highlighting potential disproportionate exposure of the population within this grade to health risks.

The third graph shows the average percentage of low life expectancy per grade. Grade D neighborhoods show the highest percentage of low life expectancy. The lower life expectancy in these areas can be associated with historically restricted public investment in healthcare and limited access to medical assistance, suggesting possible environmental injustice concerns and their effect on health inequality

Part 2: Legacy of redlining in biodiversity observations

1. A figure summarizing the percent of observations within redlined neighborhoods within each HOLC grade

```r
# Transform coordinates from dataset to match
birds <- st_transform(birds, crs = st_crs(map_ineq))

# Check if coordinate reference systems match and add a stop error
if(st_crs(birds) != st_crs(map_ineq)){
    stop("coordinate reference systems must match to join datasets")
}

# Join datasets
birds_joined <- st_join(birds, map_ineq, join = st_intersects)
```

```r
# Summarize bird count and percentage distribution within HOLC grades
summarize_birds <- birds_joined %>%
  group_by(grade) %>%
  summarise(count = n()) %>%
  tidyr::drop_na() %>%
  mutate(percentage = (count / sum(count)) * 100)
```

```r
# Plot for percentile of bird observations within HOLC grade
ggplot(data = summarize_birds, aes(x = grade, y = percentage , fill = grade)) +
  geom_bar(stat = "identity") +
  scale_fill_manual(values = c("A" = "#76a865",
                               "B" = "#7cb5bd",
                               "C" = "#ffff00",
                               "D" = "#d9838d"), ) +
  labs(title = "Percent of bird observations
                 per HOLC in Los Angeles",
       x = "HOLC Neighborhood Grade",
       y = "Bird Observations (%)") +
    theme_minimal() +
    theme(legend.position = "none")
```

Percent of bird observations per HOLC in Los Angeles
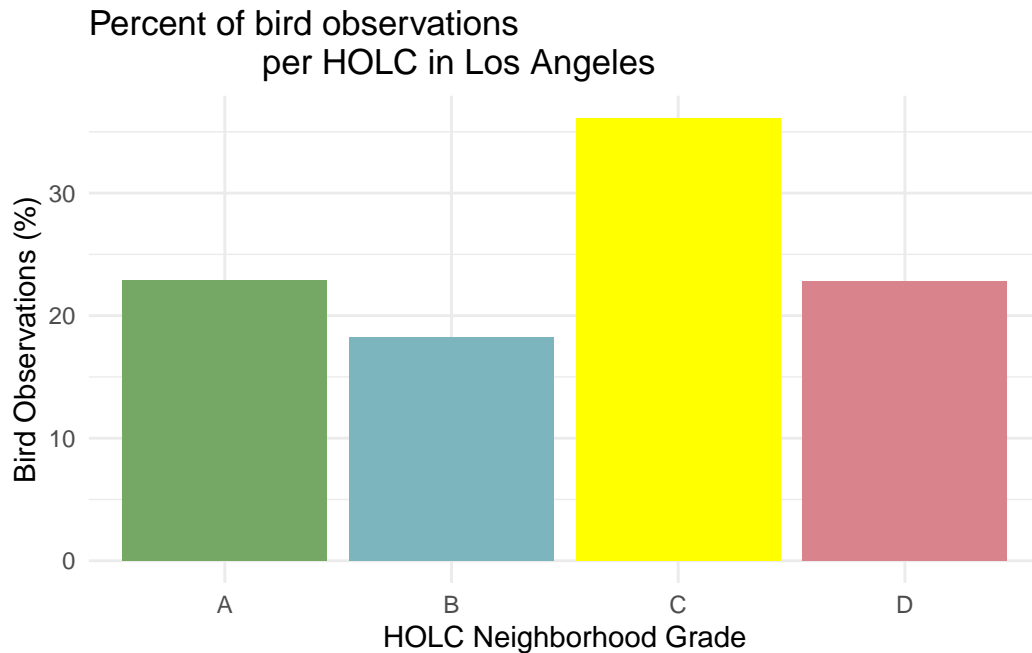
2. Our results don't match the findings from Ellis-Soto et al. 2023! Read the abstract of the study. Why might we have obtained different results in our analysis? What did the paper consider that we did not?

The paper states that the lowest bird observations are from the redlined neighborhoods, which differs from the results found here, where the B-grade neighborhoods have the lowest percentage of bird observations. One possible reason we might have obtained these different results is the influence of other variables not accounted for; for example, the population of people within each grade could affect the number of observations recorded. In the article, it is mentioned that the analyses accounted for present-day vegetation, open space, population density, and climate. They also used a much larger sample, including 195 metropolitan areas in the United States, compared with our results, which were calculated only as the percentage of bird observation counts per grade level in Los Angeles.

Acknowledgement

Ellis-Soto, D., Chapman, M., & Locke, D. H. (2022, June 9). Uneven biodiversity sampling across redlined urban areas in the United States. Ecoevorxiv.org; EcoEvoRxiv. https://ecoevorxiv.org/repository/view/3736/