Exploration of Environmental Justice

Aakriti Poudel

2025-10-08

Load libraries

```
# Import libraries
library(tidyverse)
library(sf)
library(tmap)
library(spData) # preloaded spatial data
library(kableExtra)
library(ggplot2)
```

Read data

```
# Read environmental justice screening data
ejscreen <- sf::st_read(here::here("data",
                                   "ejscreen",
                                   "EJSCREEN_2023_BG_StatePct_with_AS_CNMI_GU_VI.gdb"))

# Read Home Owners' Loan Corporation (HOLC) redlining data
map_inequality <- sf::st_read(here::here("data",
                                   "mapping-inequality",
                                   "mapping-inequality-los-angeles.json"))

# Read biodiversity observations data
la_birds <- sf::st_read(here::here("data",
                                   "gbif-birds-LA",
                                   "gbif-birds-LA.shp"))
```

```r
# Filter to a county you are interested in: Los Angeles County
la <- ejscreen %>%
  filter(ST_ABBREV == "CA") %>%
  filter(CNTY_NAME == c("Los Angeles County"))
```

```r
# Check whether CRS matches
st_crs(map_inequality) == st_crs(la)
```

```
[1] FALSE
```

```r
# Transform the CRS of a dataset
map_ineq_transform <- st_transform(map_inequality, crs = st_crs(la))

# Check CRS of the dataset
if(st_crs(map_ineq_transform) == st_crs(la)){
  print("It's a match!")
} else {
  print("It's not a match!!")
}
```

Create map

Legacy of redlining in current environmental (in)justice

1. Create a map of historical redlining neighborhoods

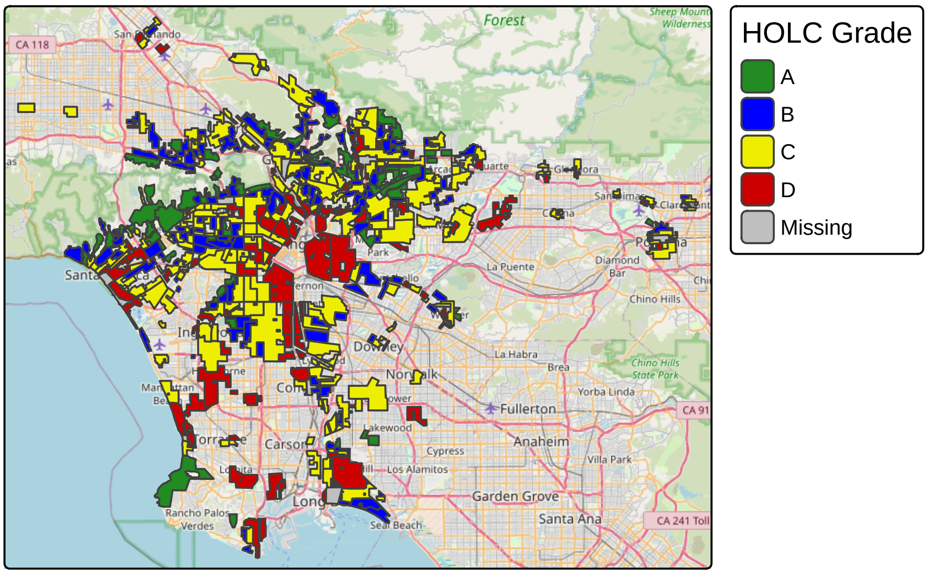
```r
# Plot a map of historical redlining neighborhood

redlining_area <- tm_shape(map_inequality) +
  # neighborhoods colored by HOLC grade
  tm_polygons(fill = 'grade',
              fill.scale = tm_scale(values = c('A' = 'forestgreen',
                                                 'B' = 'blue',
                                                 'C' = 'yellow2',
                                                 'D' = 'red3')),
              fill.legend = tm_legend(title = 'HOLC Grade')) +
  tm_title(text = 'Historical redlining neighborhoods colored by HOLC grade',
           fontface = 'bold',
           size = 1,
           position = tm_pos_out('center', 'top')) +
```

```
  tm_basemap(c('OpenStreetMap', 'Esri.WorldImagery', 'CartoDB.Positron')) +
  tm_legend()

redlining_area
```

Historical redlining neighborhoods colored by HOLC grade




```
# Save HOLC grade block map
holc_block <- tmap_save(redlining_area, here::here('figs', 'holc-block.jpg'),
                        width = 12, height = 12)
```

2. Create a table summarizing:

a. the percentage of census block groups that fall within each HOLC grade
b. Also include the percent of census black groups that don't fall within a HOLC grade.

Hint: The HOLC data contains the grades and the EJScreen data contains the census blocks, so you will need to combine the data spatially before doing summary statistics. Once you combine and no longer need the geometries, you can use `st_drop_geometry()`.

```
# Spatial join between EJScreen data and HOLC grade data
la_ej <- st_join(map_ineq_transform, la, join = st_intersects) %>%
# Drop geometry
  st_drop_geometry()
```

```r
# Total number of rows in dataset
total_blocks <- nrow(la_ej)

# Calculate percentage of each HOLC grade
holc_summary <- la_ej %>%
  # Replace NA with descriptive label
  mutate(holc_grade = replace_na(grade, 'No HOLC grade')) %>%
  group_by(holc_grade) %>%
  summarise(count = n()) %>%
  mutate(percentage = (count / sum(count)) * 100)

# Create a summary table
holc_summary_table <- holc_summary %>%
  kable(col.names = c('HOLC grade', 'No. of blocks', '% of blocks'),
        align = 'ccc') %>%
  kable_styling(bootstrap_options = c('striped', 'bordered', 'hover'),
                full_width = FALSE,
                position = 'center') %>%
  column_spec(1:3, width = '3in')

holc_summary_table
```

HOLC grade	No. of blocks
A	449
B	1239
C	3058
D	1346
No HOLC grade	296

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):

a. % low income
b. percentile for Particulate Matter 2.5
c. percentile for low life expectancy

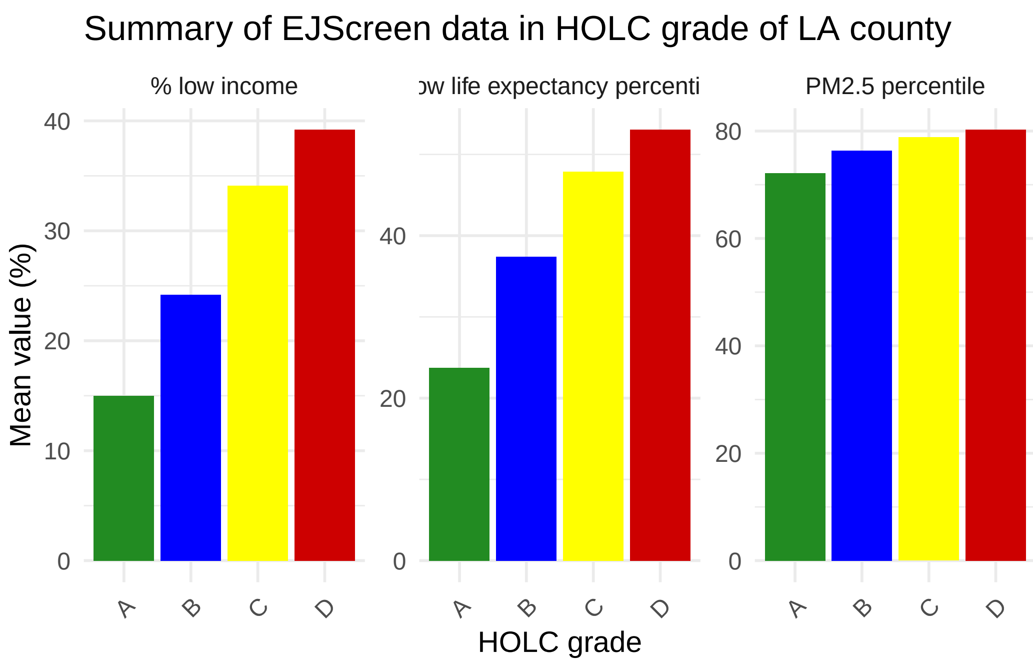
Use ggplot for your visualizations! You will first need to calculate mean of each variable grouped by HOLC grade.

```r
# Calculate mean of % low income, percentile for Particulate Matter 2.5 and percentile for lo
la_ej_mean <- la_ej %>%
  filter(grade != 'NA') %>%
  group_by(grade) %>%
  summarise(lowincome_mean = 100 * mean(LOWINCPCT, na.rm = TRUE),
            pm_mean = mean(P_PM25, na.rm = TRUE),
            lifeexp_mean = mean(P_LIFEEXPPCT, na.rm = TRUE))

# Change dataset to long dataset for visualization
la_ej_long <- la_ej_mean %>%
  pivot_longer(cols = c(lowincome_mean, pm_mean, lifeexp_mean),
               names_to = 'variable',
               values_to = 'mean_value') %>%
  mutate(variable = case_when(
    variable == 'lowincome_mean' ~ '% low income',
    variable == 'pm_mean' ~ 'PM2.5 percentile',
    variable == 'lifeexp_mean' ~ 'Low life expectancy percentile'))
```

```r
# Visualization of summary of EJScreen data in HOLC grade of LA county
summary_ejscreen <- ggplot(la_ej_long, aes(x = grade, y = mean_value, fill = grade)) +
  geom_col(position = 'dodge', linewidth = 0.3) +
  facet_wrap(~variable, scales = 'free_y', ncol = 3) +
  scale_fill_manual(values = c('A' = 'forestgreen',
                               'B' = 'blue',
                               'C' = 'yellow1',
                               'D' = 'red3')) +
  labs(title = 'Summary of EJScreen data in HOLC grade of LA county',
       x = 'HOLC grade',
       y = 'Mean value (%)',
       fill = 'HOLC Grade') +
       theme_minimal() +
  theme(legend.position = 'none',
        axis.text.x = element_text(angle = 45, hjust = 1))

summary_ejscreen
```




```r
# Save the plot of summary of EJScreen data
ggsave(here::here('figs', 'summary-ejscreen.jpg'),
                       width = 10,
                       height = 8,
                       dpi = 300)
```

4. Write a brief paragraph reflecting on these results.

a. Interpret the patterns you observe in your results. **Answer** The ***'Summary of EJScreen data in HOLC grade of LA county'*** figure reveals clear disparities across historically redlined neighborhoods

The low income graph shows that block D neighborhoods, historically redlined areas, have the highest percentage of low income population compared to Grades A, B, and C. This indicates that redlined areas contain higher concentrations of economically disadvantaged populations. The life expectancy graph demonstrates that block D neighborhoods have the highest percentile for low life expectancy. Residents in redlined areas experience shorter lifespans compared to those in better-graded neighborhoods. The PM2.5 graph shows that block D neighborhoods have the highest particulate matter pollution levels. The population in redlined areas face greater exposure to air pollution than those in other neighborhoods.

These patterns align with existing research showing that redlined neighborhoods typically have less green space and higher temperatures than other areas.

b. Discuss potential relationships between historical redlining grades and current environmental/socioeconomic conditions. **Answer** The analysis reveals a strong relationship between historical redlining practices and present-day environmental and socioeconomic conditions. Communities in formerly redlined neighborhoods (Grade D) experience multiple disadvantages: higher rates of poverty, reduced life expectancy, and increased exposure to air pollution. These areas often lack adequate infrastructure, healthcare facilities, and green spaces that are more common in higher-graded neighborhoods. This pattern demonstrates how discriminatory housing policies from the 1930s continue to affect environmental and health outcomes today. The legacy of redlining has created lasting inequalities in access to environmental quality and economic opportunity, disproportionately impacting low-income communities.

Part 2: Legacy of redlining in biodiversity observations

Create a visualizations that shows: a. The percentage of bird observations within each HOLC grade b. Include an appropiate title, axis labels, and legend ***Hints:*** Ensure the bird observations and HOLC dataset have matching CRS', then perform a spatial join to assign each bird observations to a corresponding HOLC grade.

Spolier alert!! 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?

```
# Check CRS of the dataset
if(st_crs(map_ineq_transform) == st_crs(la_birds)){
  print("It's a match!")
} else {
  print("It's not a match!!")
}
```

```
[1] "It's not a match!!"
```

```
# Transform the CRS of a dataset
la_birds_transform <- st_transform(la_birds, crs = st_crs(map_ineq_transform))
la_birds_transform
```

```
# Spatial join between bird observation and HOLC grade
holc_birds <- st_join(map_ineq_transform, la_birds_transform, join = st_intersects) %>%
# Drop geometry
  st_drop_geometry()
```

```r
# Calculate the percentage of bird observations within each HOLC grade
holc_birds_percent <- holc_birds %>%
  filter(grade != 'NA', # Filter out the NA
         year >= 2021 & year <= 2023) %>%  # Filter for the year from 2021 - 2023
  group_by(grade, year) %>%
  summarise(count = n()) %>%
  ungroup() %>%
  group_by(year) %>%
  mutate(percentage = (count / sum(count)) * 100) %>%
  ungroup()

holc_birds_percent
```

```
# A tibble: 12 x 4
   grade  year count percentage
   <chr> <int> <int>      <dbl>
 1 A      2021 29230       23.3
 2 A      2022  1073       15.7
 3 A      2023    42       12.9
 4 B      2021 22832       18.2
 5 B      2022  1329       19.4
 6 B      2023    37       11.3
 7 C      2021 45380       36.1
 8 C      2022  2439       35.6
 9 C      2023   154       47.2
10 D      2021 28149       22.4
11 D      2022  2004       29.3
12 D      2023    93       28.5
```

```r
# Visualization of percentage of bird observations within each HOLC grade
ggplot(holc_birds_percent, aes(x = year,
                               y = percentage,
                               color = grade,
                               group = grade)) +
  geom_line(linewidth = 1, stat = "smooth", method = "loess", se = FALSE) +
  geom_point(size = 2) +
  scale_color_manual(values = c('A' = 'forestgreen',
                                'B' = 'blue',
                                'C' = 'yellow1',
                                'D' = 'red3')) +
  scale_x_continuous(breaks = c(2021, 2022, 2023)) +
```

```r
labs(title = 'Percentage of Bird Observations within each HOLC Grade',
     subtitle = 'Year: 2021-2023',
     x = 'Year of observation',
     y = 'Percentage of bird observations (%)',
     color = 'HOLC Grade') +
theme_minimal(base_size = 8)
```