

Exploring Patterns of Environmental Justice

Melannie Moreno Rolón

2025-10-18

Table of contents

1	Introduction	1
2	Part 1: Legacy of redlining in current environmental (in)justice	1
2.1	Interpretation	10
3	Part 2: Legacy of redlining in biodiversity observations	10
3.1	Interpretation	12

1 Introduction

This assignment explores the legacy of redlining in Los Angeles, California, and its relationship to present-day environmental and ecological conditions. Using spatial data from the EPA's EJScreen (2023), historical HOLC neighborhood grades, and bird biodiversity observations from GBIF, we analyze how historically marginalized areas compare in terms of socioeconomic vulnerability, pollution exposure, and ecological sampling. This work reveals the impact of discriminatory housing policies and highlights the importance of equitable data representation in conservation and public health.

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

Our first task is to explore historical redlining in Los Angeles and its legacy on present-day environmental justice.

1. Create a map of historical redlining neighborhoods

```

# Read in data
ejscreen <- st_read(here::here("data", "ejscreen",
                               "EJSCREEN_2023_BG_StatePct_with_AS_CNMI_GU_VI.gdb"))

Reading layer `EJSCREEN_StatePctiles_with_AS_CNMI_GU_VI` from data source
  `C:\Documents\MEDS\EDS223\EDS223-2025-Homework-2\data\ejscreen\EJSCREEN_2023_BG_StatePct_w
  using driver `OpenFileGDB'
Simple feature collection with 243021 features and 223 fields
Geometry type: MULTIPOLYGON
Dimension:      XY
Bounding box:  xmin: -19951910 ymin: -1617130 xmax: 16259830 ymax: 11554350
Projected CRS: WGS 84 / Pseudo-Mercator

holc <- st_read(here::here("data", "mapping-inequality",
                           "mapping-inequality-los-angeles.json"))

Reading layer `mapping-inequality-los-angeles` from data source
  `C:\Documents\MEDS\EDS223\EDS223-2025-Homework-2\data\mapping-inequality\mapping-inequality-
  using driver `GeoJSON'
Simple feature collection with 417 features and 14 fields
Geometry type: MULTIPOLYGON
Dimension:      XY
Bounding box:  xmin: -118.6104 ymin: 33.70563 xmax: -117.7028 ymax: 34.30388
Geodetic CRS:  WGS 84

birds_la <- st_read(here::here("data", "gbif-birds-LA","gbif-birds-LA.shp"))

Reading layer `gbif-birds-LA` from data source
  `C:\Documents\MEDS\EDS223\EDS223-2025-Homework-2\data\gbif-birds-LA\gbif-birds-LA.shp'
  using driver `ESRI Shapefile'
Simple feature collection with 1288865 features and 1 field
Geometry type: POINT
Dimension:      XY
Bounding box:  xmin: -118.6099 ymin: 33.70563 xmax: -117.7028 ymax: 34.30385
Geodetic CRS:  WGS 84

#Make sure that holc crs matches basemap crs
holc <- st_transform(holc, 3857)
st_crs(holc)

```

```

Coordinate Reference System:
User input: EPSG:3857
wkt:
PROJCRS["WGS 84 / Pseudo-Mercator",
    BASEGEOGCRS["WGS 84",
        ENSEMBLE["World Geodetic System 1984 ensemble",
            MEMBER["World Geodetic System 1984 (Transit)"],
            MEMBER["World Geodetic System 1984 (G730)"],
            MEMBER["World Geodetic System 1984 (G873)"],
            MEMBER["World Geodetic System 1984 (G1150)"],
            MEMBER["World Geodetic System 1984 (G1674)"],
            MEMBER["World Geodetic System 1984 (G1762)"],
            MEMBER["World Geodetic System 1984 (G2139)"],
            MEMBER["World Geodetic System 1984 (G2296)"],
            ELLIPSOID["WGS 84",6378137,298.257223563,
                LENGTHUNIT["metre",1]],
            ENSEMBLEACCURACY[2.0]],
        PRIMEM["Greenwich",0,
            ANGLEUNIT["degree",0.0174532925199433]],
        ID["EPSG",4326]],
    CONVERSION["Popular Visualisation Pseudo-Mercator",
        METHOD["Popular Visualisation Pseudo Mercator",
            ID["EPSG",1024]],
        PARAMETER["Latitude of natural origin",0,
            ANGLEUNIT["degree",0.0174532925199433],
            ID["EPSG",8801]],
        PARAMETER["Longitude of natural origin",0,
            ANGLEUNIT["degree",0.0174532925199433],
            ID["EPSG",8802]],
        PARAMETER["False easting",0,
            LENGTHUNIT["metre",1],
            ID["EPSG",8806]],
        PARAMETER["False northing",0,
            LENGTHUNIT["metre",1],
            ID["EPSG",8807]]],
    CS[Cartesian,2,
        AXIS["easting (X)",east,
            ORDER[1],
            LENGTHUNIT["metre",1]],
        AXIS["northing (Y)",north,
            ORDER[2],
            LENGTHUNIT["metre",1]]],
    USAGE[

```

```

SCOPE["Web mapping and visualisation."],
AREA["World between 85.06°S and 85.06°N."],
BBOX[-85.06,-180,85.06,180]],
ID["EPSG",3857]

# Map of historical redlining neighborhoods
tmap_mode("plot")

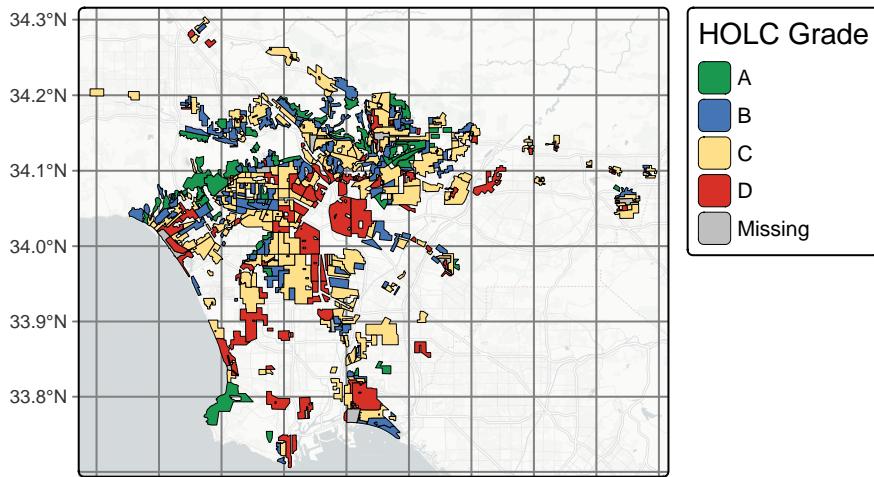
tm_shape(holc) +
  tm_graticules() +
  tm_polygons(fill = "grade",
               title = "HOLC Grade",
               palette = c("A" = "#1a9850", # green
                           "B" = "#4575b4", # blue
                           "C" = "#fee08b", # yellow
                           "D" = "#d73027"), # red
               border.col = "black",
               lwd = 0.4) +
  tm_basemap("CartoDB.PositronNoLabels") +
  tm_title("Los Angeles: Historic Redlining Neighborhoods", size = 2)

```

```

|-----|-----|-----|-----|
=====
```

Los Angeles: Historic Redlining Neighborhoods



2. Create a table summarizing the percentage of census block groups that fall (and don't fall) within each HOLC grade

```
# Custom function to check CRS and transform if needed
check_transform_crs <- function(sf1, sf2) {
  if (st_crs(sf1) != st_crs(sf2)) {

    warning("CRS does not match. Transforming sf1 to match sf2...")

    # Attempt transformation
    transformed <- st_transform(sf1, st_crs(sf2))

    # Verify transformation
    if (st_crs(transformed) != st_crs(sf2)) {
      stop("Error: CRS transformation failed. Please check CRS definitions.")
    } else {
      message("CRS transformation successful.")
      sf1 <- transformed
    }
  } else {
    message("CRS already matches. No transformation needed.")
  }
}
```

```

    return(sf1)
}

# Make sure crs for holc and ejscreen matches
holc <- check_transform_crs(sf1 = holc, sf2 = ejscreen)

# Filter EJ Screen data to LA
ejscreen_la <- ejscreen %>%
  st_filter(y = holc, .predicate = st_intersects)

# Join data
ej_holc <- st_join(ejscreen_la, holc, left = TRUE)

# Create summary table of data
summ_table <- ej_holc %>%
  # Replace NA values
  mutate(grade = ifelse(is.na(grade), "No grade", grade)) %>%
  # Group by HOLC grade
  group_by(grade) %>%
  # Count block groups in each grade
  summarise(count = n()) %>%
  # Calculate percentage
  mutate(percent = round(100 * count / sum(count), 1)) %>%
  # Display grades in order (A, B, C, D, No grade)
  mutate(grade = factor(grade, levels = c("A", "B", "C", "D", "No grade"))) %>%
  arrange(grade) %>%
  # Remove geometries
  st_drop_geometry()

# Format table
ej_holc_gt <- summ_table %>%
  #remove count column
  select(-count) %>%
  gt() %>%
  tab_header(title = "HOLC Grades and Census Block Groups in LA") %>%
  cols_label(grade = "HOLC Grade",
             percent = "Percent of Census Block Groups (%)")

ej_holc_gt

```

[
HOLC Grades and Census Block Groups in LA]
HOLC Grades and Census Block Groups in LA

HOLC Grade	Percent of Census Block Groups (%)
A	7.0
B	19.4
C	47.9
D	21.1
No grade	4.6

3. Create at least two visualizations summarizing current conditions (from the EJScreen data) within HOLC grades using the mean of the percent low income, the percentile for Particulate Matter 2.5, and the percentile for low life expectancy.

```
# Summarize variables
ej_holc_avgs <- ej_holc %>%
  # Replace NA values in 'grade'
  mutate(grade = ifelse(is.na(grade), "No grade", grade)) %>%
  # Group by grade
  group_by(grade) %>%
  # Calculate the mean for each variable
  summarise(
    mean_lowincpct = round(mean(LOWINCPCT, na.rm = TRUE), 2),
    mean_pm25 = round(mean(P_PM25, na.rm = TRUE), 2),
    mean_lifeexpct = round(mean(P_LIFEEXPCT, na.rm = TRUE)), 2) %>%
  # Display HOLC grades in order
  mutate(grade = factor(grade, levels = c("A", "B", "C", "D", "No grade"))) %>%
  arrange(grade) %>%
  # Remove geometries
  st_drop_geometry()

# View table
ej_holc_avgs
```

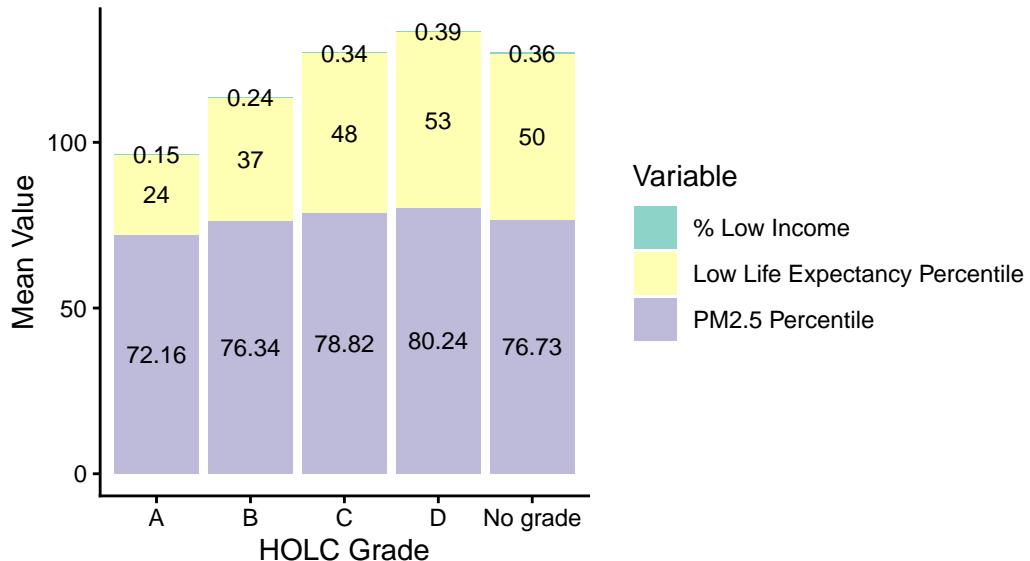
```
# A tibble: 5 x 5
  grade    mean_lowincpct mean_pm25 mean_lifeexpct `2`
* <fct>      <dbl>     <dbl>        <dbl> <dbl>
1 A            0.15      72.2         24     2
2 B            0.24      76.3         37     2
```

3 C	0.34	78.8	48	2
4 D	0.39	80.2	53	2
5 No grade	0.36	76.7	50	2

```
# Reshape data frame to long format
ej_holc_avgs_long <- ej_holc_avgs %>%
  # Pivot data frame to long format
  pivot_longer(cols = starts_with("mean_"),
               names_to = "variable",
               values_to = "mean_val") %>%
  # Rename variables for the 'variable' column
  mutate(variable = recode(variable,
                           "mean_lowincpct" = "% Low Income",
                           "mean_pm25" = "PM2.5 Percentile",
                           "mean_lifeexpct" = "Low Life Expectancy Percentile"))
```

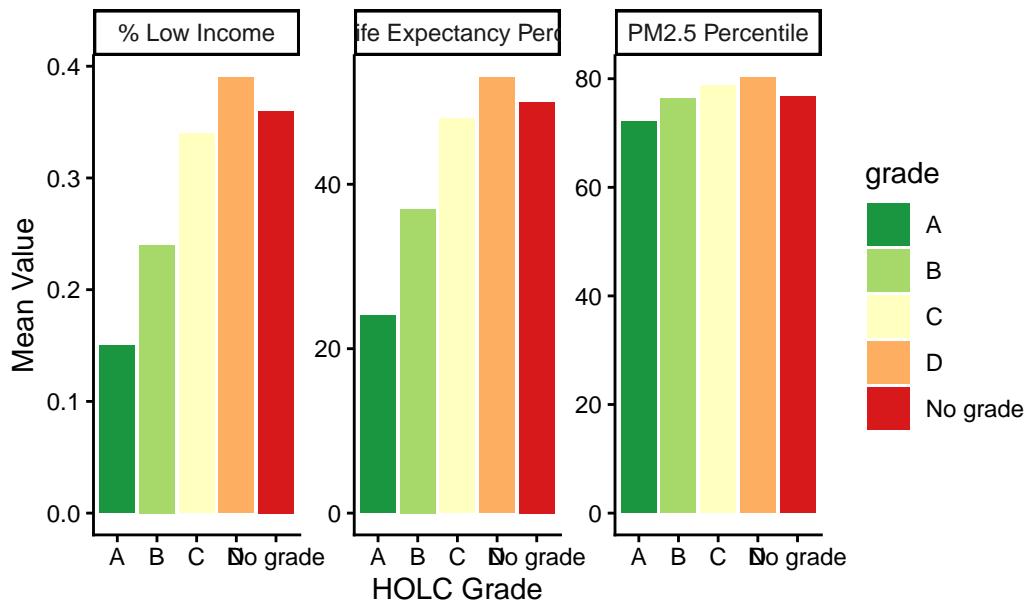
```
#Stacked Bar Chart
ggplot(ej_holc_avgs_long, aes(x = grade,
                                y = mean_val,
                                fill = variable,
                                label = mean_val)) +
  geom_col() +
  # Add labels to our plot
  labs(title = "Average Environmental \nand Socioeconomic Indicators by HOLC Grade",
       x = "HOLC Grade",
       y = "Mean Value",
       fill = "Variable") +
  # Select a theme
  theme_classic() +
  # Show values for the stacked bar chart
  geom_text(size = 3, position = position_stack(vjust = 0.5)) +
  # Select color theme
  scale_fill_brewer(palette = "Set3")
```

Average Environmental and Socioeconomic Indicators by HOLC Grade



```
# Faceted Bar Chart
ggplot(ej_holc_avgs_long, aes(x = grade, y = mean_val, fill = grade)) +
  geom_col() +
  # Facet the plot by variable and have different y-axis ranges
  facet_wrap(~variable, scales = "free_y") +
  # Add labels
  labs(title = "EJScreen Conditions by HOLC Grade in Los Angeles",
       x = "HOLC Grade",
       y = "Mean Value") +
  # Choose a theme
  theme_classic() +
  # Color scheme for HOLC grades
  scale_fill_brewer(palette = "RdYlGn", direction = -1)
```

EJScreen Conditions by HOLC Grade in Los Angeles



2.1 Interpretation

The plots show a clear relationship between historical redlining grades and present-day environmental and socioeconomic conditions in Los Angeles. Neighborhoods with worse HOLC grades (C and D) tend to have higher percentages of low-income residents, higher percentiles of low life expectancy, and higher PM2.5 concentrations compared to areas graded A or B. These patterns suggest that the effects of redlining persist today, with some areas experiencing greater environmental burdens and poorer health and economic outcomes. The upward trend across variables highlights how historical housing discrimination has contributed to lasting environmental injustice in the city.

3 Part 2: Legacy of redlining in biodiversity observations

```
# Check crs for birds_la and holc data frames
birds_la <- check_transform_crs(sf1 = birds_la, sf2 = holc)

# Use st_within to ensure that bird observations are located inside HOLC polygons
birds_within_holc <- st_join(birds_la, holc, join = st_within)
```

```
# Check whether NAs are present in data frame
if (anyNA(birds_within_holc$grade)) {
  warning("Some bird observations fall outside HOLC polygons")
}
```

```
# Calculate percentage of birds within holc grades
birds_summ <- birds_within_holc %>%
  st_drop_geometry() %>%
  # Exclude bird observations with NA HOLC grades to focus on areas with historical redlining
  filter(!is.na(grade)) %>%
  # Group by grade
  group_by(grade) %>%
  # Count birds in each HOLC grade
  summarise(count = n()) %>%
  # Calculate percentage of counts
  mutate(percent = 100 * count / sum(count))
```

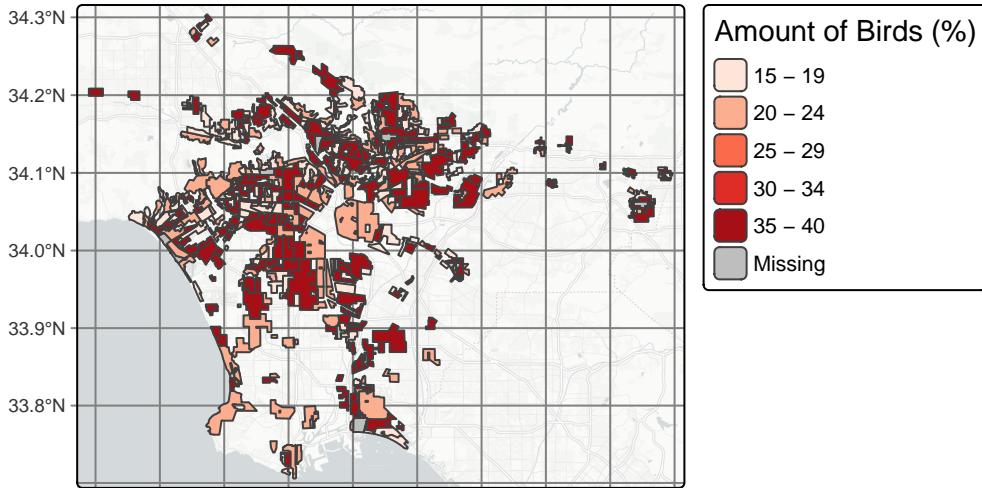
```
birds_summ
```

	grade	count	percent
1	A	30344	22.9
2	B	24197	18.2
3	C	47973	36.1
4	D	30246	22.8

```
# Join percentages back into HOLC polygons
holc_birds_summ <- holc %>%
  left_join(birds_summ, by = c("grade"))
```

```
# Plot our result
tm_shape(holc_birds_summ) +
  tm_graticules() +
  tm_polygons("percent",
    palette = "brewer.reds",
    title = "Amount of Birds (%)") +
  tm_basemap("CartoDB.PositronNoLabels") +
  tm_title(text = "Percentage of Bird Observations by HOLC Grade (2021-2023)")
```

Percentage of Bird Observations by HOLC Grade (2021–2023)



3.1 Interpretation

Our results don't match the findings from Ellis-Soto et al. 2023! Why might we have obtained different results in our analysis? What did the paper consider that we did not?

Our analysis has a few differences from the methodology used by Ellis-Soto (2023). First, the paper notes that bird observations are disproportionately concentrated in historically white, “desirable” neighborhoods. Our map only includes birds within HOLC polygons, excluding a large portion of data (around 90%) that falls outside these areas. This created the misleading assumption that all neighborhoods were equally sampled, whereas the paper explicitly adjusts for uneven sampling and density. Second, our analysis lacks important control variables such as vegetation, open space, population density, and climate, all of which were accounted for in the original study. These factors significantly influence where birds are observed and must be considered to avoid confounding results. Third, our analysis focused solely on Los Angeles, while the paper analyzes data from 38 metro areas across the U.S. This difference in scale means that our findings may not reflect national patterns. Fourth, the metrics used in this study, that is, the percentage of total birds per HOLC grade, differs from the paper's focus on bird observation density and sampling completeness. Thus, our map shows raw counts rather than adjusted measures, which may have obscured under-sampled areas. Lastly, the presence of many bird records with no HOLC grade (NA) suggests spatial matching issues. Ignoring these unmatched points could have skewed our results and misrepresented the true distribution of bird observations.