The Geography of Injustice:

Redlining, Environment, and Biodiversity in Los Angeles

Jackson Coldiron

2025-10-18

Introduction

This assignment investigates how the legacy of 1930s HOLC "redlining" in Los Angeles County relates to present-day environmental and socioeconomic conditions. Using EJScreen block-group indicators alongside digitized HOLC grades from Mapping Inequality, I'll create maps and summaries to compare outcomes across grades. I'll also spatially join 2021–2023 GBIF bird observations to HOLC areas to examine whether biodiversity data collection mirrors historic inequities. Together, these analyses use R (tmap/ggplot) to visualize patterns, quantify disparities, and reflect on mechanisms linking past policy to current environmental justice.

Data Overview

This analysis draws on three primary spatial datasets. The **HOLC Mapping Inequality** dataset provides 1930s neighborhood grades (A–D) for Los Angeles, representing historic redlining boundaries. The **EPA EJScreen** dataset contributes present-day demographic and environmental indicators at the census block group level, including measures such as particulate matter and income. Finally, **GBIF bird observations** from 2021–2023 offer point data on species occurrences, allowing for assessment of biodiversity sampling patterns across historical HOLC zones.

Data Citation

HOLC Redlining Data: Nelson, R. K., Winling, L., Marciano, R., Connolly, N., et al. (2023). Mapping Inequality. American Panorama, ed. Robert K. Nelson and Edward L. Ayers. University of Richmond.

https://dsl.richmond.edu/panorama/redlining/

EJScreen Data: U.S. Environmental Protection Agency. (2023). EJSCREEN: Environmental Justice Screening and Mapping Tool. Data retrieved from EPA Environmental Justice datasets. https://www.epa.gov/ejscreen

GBIF Bird Observations: Global Biodiversity Information Facility. (2023). Bird observations in Los Angeles, 2021-2023 [Dataset]. GBIF.org. https://www.gbif.org/

Data Preparation

```
# Load libraries
library(tidyverse)
library(dplyr)
library(here)
library(tmap)
library(sf)
library(tmaptools)
library(ggplot2)
library(maptiles)
library(stars)
library(paletteer)
library(NatParksPalettes)
library(rnaturalearth)
library(knitr)
library(kableExtra)


# Bring in the data
birds <- suppressMessages(st_read(here("data", "gbif-birds-LA", "gbif-birds-LA.shp"))) # CRS: 
```

```
Reading layer `gbif-birds-LA' from data source
  `/Users/jacksoncoldiron/Documents/Bren/Fall 2025/EDS 223/EDS223-HW2/data/gbif-birds-LA/gbif-
  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
```

```
inequality <- suppressMessages(st_read(here("data", "mapping-inequality", "mapping-inequality-
```

```
Reading layer `mapping-inequality-los-angeles' from data source
  `/Users/jacksoncoldiron/Documents/Bren/Fall 2025/EDS 223/EDS223-HW2/data/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
```

```
ejscreen <- suppressMessages(st_read(here("data", "ejscreen", "EJSCREEN_2023_BG_StatePct_with_A
```

```
Reading layer `EJSCREEN_StatePctiles_with_AS_CNMI_GU_VI' from data source
  `/Users/jacksoncoldiron/Documents/Bren/Fall 2025/EDS 223/EDS223-HW2/data/ejscreen/EJSCREEN_20
  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
```

```
# Check CRS
#st_crs(birds)
#st_crs(inequality)
#st_crs(ejscreen)

# Verify CRS match
if (st_crs(birds) != st_crs(inequality)) {
  stop("CRS mismatch between birds and inequality data!")
}

# Transform ejscreen to match other datasets if needed
if (st_crs(ejscreen) != st_crs(birds)) {
  ejscreen <- st_transform(ejscreen, st_crs(birds))
  cat("EJScreen transformed to match bird and inequality CRS\n")
}
```

```
EJScreen transformed to match bird and inequality CRS
```

```
# Check the class
#class(birds) #sf data.frame
#class(inequality) #sf data.frame

# Peak at the geometry type
#st_geometry_type(birds) # point
#st_geometry_type(inequality) # multipolygon
```

```r
# Filter EJ Screen data to LA county
ejscreen_la <- ejscreen |>
  filter(CNTY_NAME == "Los Angeles County")
```

Part 1: Current Legacy of Redlining in EJ

i. Historical Redlining Map

First, we will create a map of the historical redlining neighborhoods colored by the HOLC grades.

```r
# Set static tmap mode
tmap_mode("plot")

# Factor the grades so there is a clear order distinction
# A is best...
ineq <- inequality |>
  mutate(grade = factor(grade, levels = c("A", "B", "C", "D")))

# Map of historical redlining neighborhoods
# Include:
  # neighborhoods colored by HOLC grade
  # an appopriate base map
  # a clear title and legend

# Get Acadia palette colors
acadia_colors <- c("#FED789FF", "#023743FF", "#72874EFF", "#A4BED5FF", "#A4BED5FF", "#453947FF

# Get U.S. state boundaries and filter for California
us_states <- ne_states(country = "United States of America", returnclass = "sf") # from rnatura
ca_ne <- us_states %>%
  filter(name == "California")

# Create custom border
neat <- st_as_sfc(tmaptools::bb(ineq, ext = 1.15))

# First create the main map that shows the different HOLC grades over LA
main_map <-
  tm_shape(ineq) +
  tm_polygons("grade", palette = acadia_colors[1:4], title = "HOLC Grade") +
  tm_compass(type = "arrow", position = tm_pos_in("left","bottom"), size = 2) +
  tm_scale_bar(position = c(0.08, 0.07), text.size = 0.8) +
  tm_layout(
```

```
  legend.outside = FALSE,
  legend.position = tm_pos_in("right","top"),
  legend.frame = TRUE,
  legend.bg.color = "transparent",
  legend.border.col = "transparent",
  legend.border.lwd = 0,
  legend.title.size = 1.2,
  legend.text.size = 0.8,
  inner.margins = c(0.08, 0.03, 0.03, 0.03)
) +
  tm_basemap("CartoDB.PositronNoLabels") +
  tm_shape(neat) + tm_borders(lwd = 3.2, col = "#2F2F2F") +
  tm_shape(neat) + tm_borders(lwd = 1.0, col = "#FFFFFF")

# Create a California bounding box for an inset map to add context
# Need to add additional space around California to make the boudning box look nicer
ca_bbox <- st_as_sfc(tmaptools::bb("California")) |> # make it a spatial object with st_as_sfc
  st_buffer(dist = 0.45) # add space around CA

# Create inset map
inset_map <-
  tm_shape(ca_bbox) +
  tm_shape(ca_ne) +
  tm_polygons(border.col = "black",
              lwd = 0.2) +
  tm_shape(ineq) +
  tm_polygons(fill = "#023743FF", border.col = "#023743FF", lwd = 3) +
  tm_layout(
    frame = TRUE,
    bg.color = "white",
    width = 2,
    height = 2,
    inner.margins = c(0, 0, 0, 0)) +
  tm_basemap("CartoDB.PositronNoLabels")
```

```
historical_map <- main_map +
  tm_inset(inset_map, height = 10, width = 10)

historical_map
```

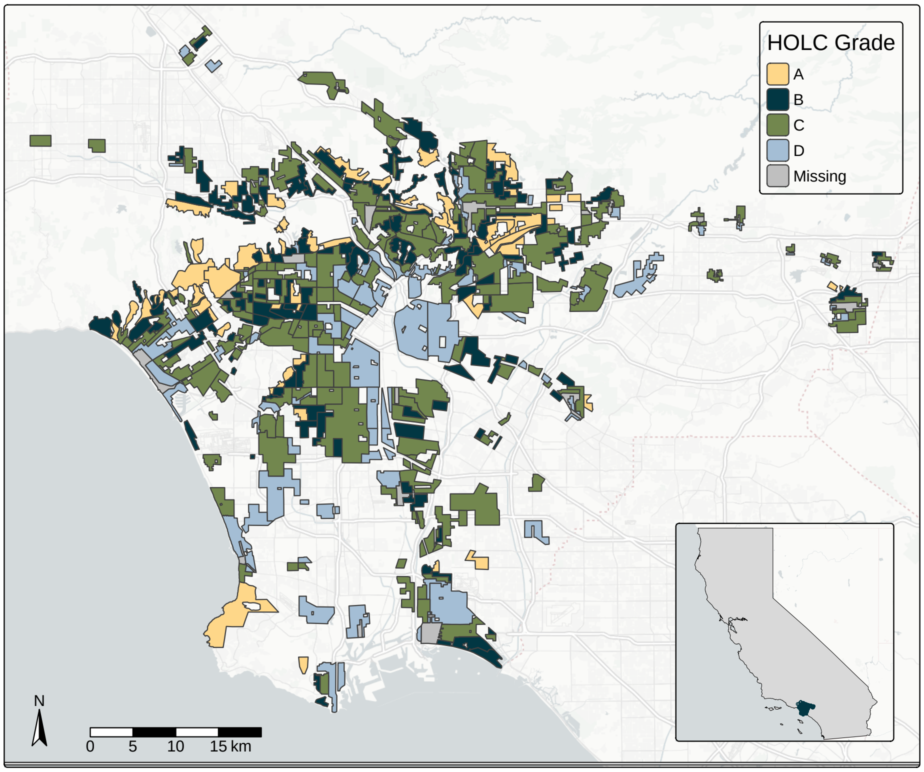

Figure 1: **Historic HOLC Redlining Patterns in Los Angeles.** Neighborhoods are shaded by 1930s HOLC investment grades—A (“Best”) shown in yellow, B (“Still Desirable”) in dark blue, C (“Definitely Declining”) in green, and D (“Hazardous”) in light blue—which reflect how lenders historically rated areas for mortgage access and perceived risk. The map presents the full spatial footprint of these designations atop a light basemap, with a north arrow and scale bar for reference, and serves as the geographic baseline for comparing present-day environmental and biodiversity patterns across legacy grades.

ii. Summary Table

Next, we want to investigate the percentage of census block groups that fall within each HOLC grade. We will also include the percent of census black groups that don’t fall within a HOLC grade to get a sense of the data gap extent.

```
# Clean up spatial objects before joining
ejscreen_la <- st_make_valid(ejscreen_la)
ineq <- st_make_valid(ineq)

# Use st_join() to combine the HOLC and EJ Screen data which contains the census block informa
ejscreen_holc <- st_join(ejscreen_la, ineq[, "grade"], join = st_within)

# Group_by() to obtain the percentage of block groups that fall within HOLC grade
grades <- ejscreen_holc |>
  st_drop_geometry() |># drop geometry for easier computation
  mutate(grade = ifelse(is.na(grade), "No Grade", grade)) |> # give NAs a different descriptio
```

```
  group_by(grade) |> # group the data by the HOLC grade
  summarise(n_bg = n()) |> # create variable for number of block groups in each grade
  mutate(pct = round(100 * (n_bg / sum(n_bg)), 2)) |># calculate percentage
  arrange(factor(grade, levels = c("A", "B", "C", "D", "No HOLC Grade")))

# Create a nice
grades_table <- kable(
  grades,
  caption = "Percentage of Census Block Groups within Each HOLC Grade",
  col.names = c("HOLC Grade", "Number of Block Groups", "Percent of Total"),
  booktabs = TRUE,
  align = c("c", "c", "c")
) %>%
  kable_styling(
    latex_options = c("hold_position", "scale_down"),
    font_size = 10
  )

grades_table
```

Table 1: Percentage of Census Block Groups within Each HOLC Grade

HOLC Grade	Number of Block Groups	Percent of Total
1	15	0.23
2	77	1.17
3	509	7.72
4	287	4.35
No Grade	5703	86.53

Note: The author is skeptical of the numbers reported in the No HOLC Grade columns. This is due to the fact that the EJ Screen data has been filtered to Los Angeles County while the HOLC Grade exists on the Los Angeles municipal scale. Because of the mismatching scopes, numbers may be distorted.

iii. Current Conditions Visualizations

Finally, to try to understand the relationship between current conditions (EJScreen data) and past redlining practices (HOLC Grades) we will create visualizations investigating three variables: % low income, percentile for Particulate Matter 2.5, percentile for low life expectancy .

```
# Create dataframe suitable for building the required plots
current_conditions <- ejscreen_holc |>
  st_drop_geometry() |> # drop spatial info
  drop_na() |> # don't include na values
```

```
  select(grade, LOWINCPCT, P_LIFEEXPPCT, P_PM25) |> # select for target variables
  group_by(grade) |> # pull together by the HOLC grades
  summarise(lowincome_avg = mean(LOWINCPCT),
            lifeexp_avg = mean(P_LIFEEXPPCT),
            pm25_avg = mean(P_PM25)) # Take averages at each grade for the target variables
```

```
# Create one bar graph to show the % of low income by HOLC grades: x axis will be % low income
lowincome_plot <- ggplot(current_conditions, aes(x = lowincome_avg * 100, y = grade)) +
  geom_col(fill = "#476F84FF") +
  scale_x_continuous(
    limits = c(0, 50),
    expand = c(0, 0),
    breaks = seq(0, 50, 5)) +
  labs(x = "Low Income (%)",
    y = "HOLC Grade"
  )  +
  theme_classic() +
   theme(
    # outer whitespace around the whole plot
    plot.margin = unit(c(12, 16, 12, 16), "pt"),
    # space between axis titles and tick labels
    axis.title.x = element_text(size = 15, face = "bold",
                                margin = margin(t = 10)),
    axis.title.y = element_text(size = 15, face = "bold",
                                margin = margin(r = 10)),
    # tick label size and spacing from the axis
    axis.text.x  = element_text(size = 11, margin = margin(t = 6)),
    axis.text.y  = element_text(size = 13, margin = margin(r = 6)),
    # optional: title spacing
    plot.title = element_text(hjust = 0.5, face = "bold",
                              margin = margin(b = 8))
  )

lowincome_plot
```

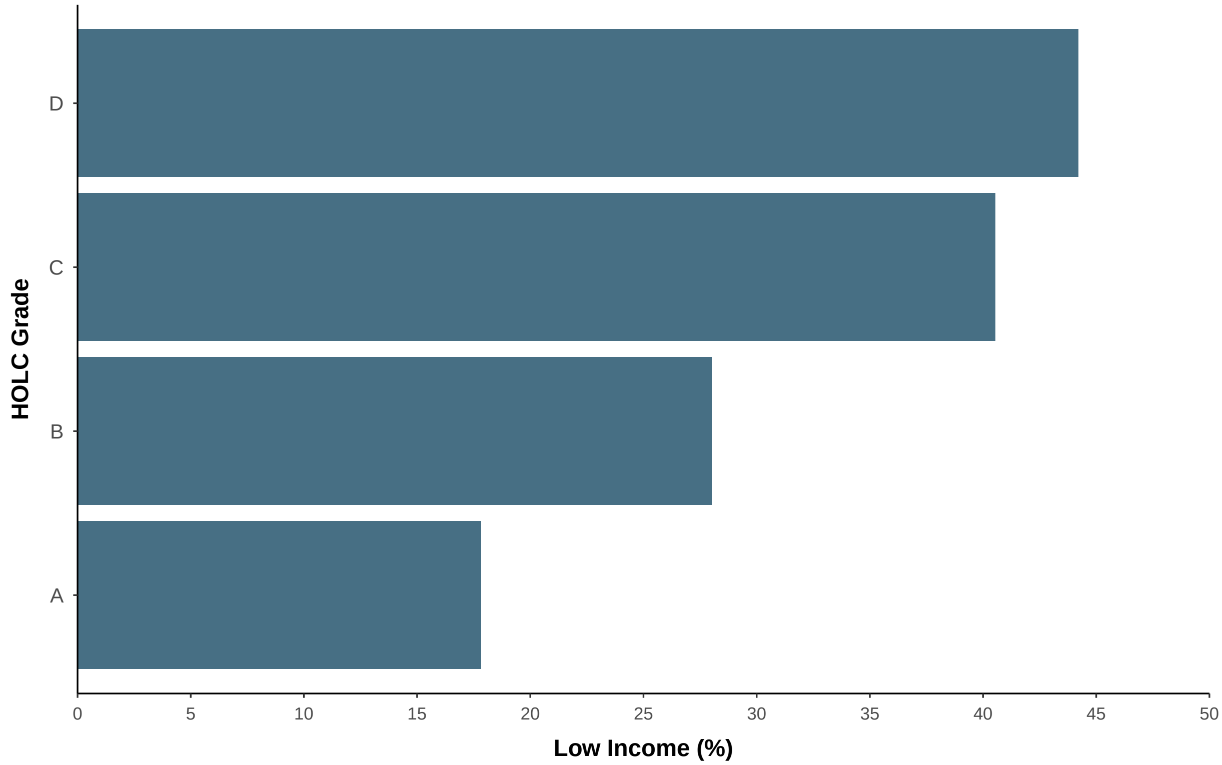

Figure 2: **Historical low HOLC grades match with a higher share of Low-income in Los Angeles.** Horizontal bars show the mean percentage of residents with household income below 2× the federal poverty level (EJScreen "low income") for block groups assigned to each HOLC grade (A–D). The pattern increases from A to D (≈18% → 44%), indicating that neighborhoods historically rated worse for mortgage access today have higher concentrations of low-income households. This begins to unveil the lasting impacts of these HOLC gradings on Los Angeles communities.

```
# Create one bar graph that shows both the percentile for low life expectancy and PM 2.5: x axi

# Pivot data into tidy form to be plotted together
current_long <- current_conditions |>
  select(grade, lifeexp_avg, pm25_avg) |>
  pivot_longer(cols = c(lifeexp_avg, pm25_avg), # Take the names of the columns and give them c
               names_to = "variable",
               values_to = "percentile") # Give the percentile values another column

# Plot both variables together
ggplot(current_long, aes(x = percentile, y = grade, fill = variable)) +
  geom_col(position = "dodge") +
  scale_x_continuous(limits = c(0, 100), breaks = seq(0, 100, 10), expand = c(0, 0)) +
  scale_fill_manual(
    values = c("lifeexp_avg" = "#A4BED5FF", "pm25_avg" = "#453947FF"),
    labels = c("Low Life Expectancy", "PM2.5 Exposure")
  ) +
```

```r
  labs(x = "Percentile",
    y = "HOLC Grade",
    fill = ""
  ) +
  theme_classic() +
   theme(
    # adjust legend sizing
     legend.position = c(0.87, 0.23),
     legend.text  = element_text(size = 14),
    legend.key.size = unit(20, "pt") ,
    # outer whitespace around the whole plot
    plot.margin = unit(c(12, 16, 12, 16), "pt"),
    # space between axis titles and tick labels
    axis.title.x = element_text(size = 15, face = "bold",
                               margin = margin(t = 10)),
    axis.title.y = element_text(size = 15, face = "bold",
                               margin = margin(r = 10)),
    # tick label size and spacing from the axis
    axis.text.x  = element_text(size = 11, margin = margin(t = 6)),
    axis.text.y  = element_text(size = 13, margin = margin(r = 6)),
    # optional: title spacing
    plot.title = element_text(hjust = 0.5, face = "bold",
                             margin = margin(b = 8))
  )
```

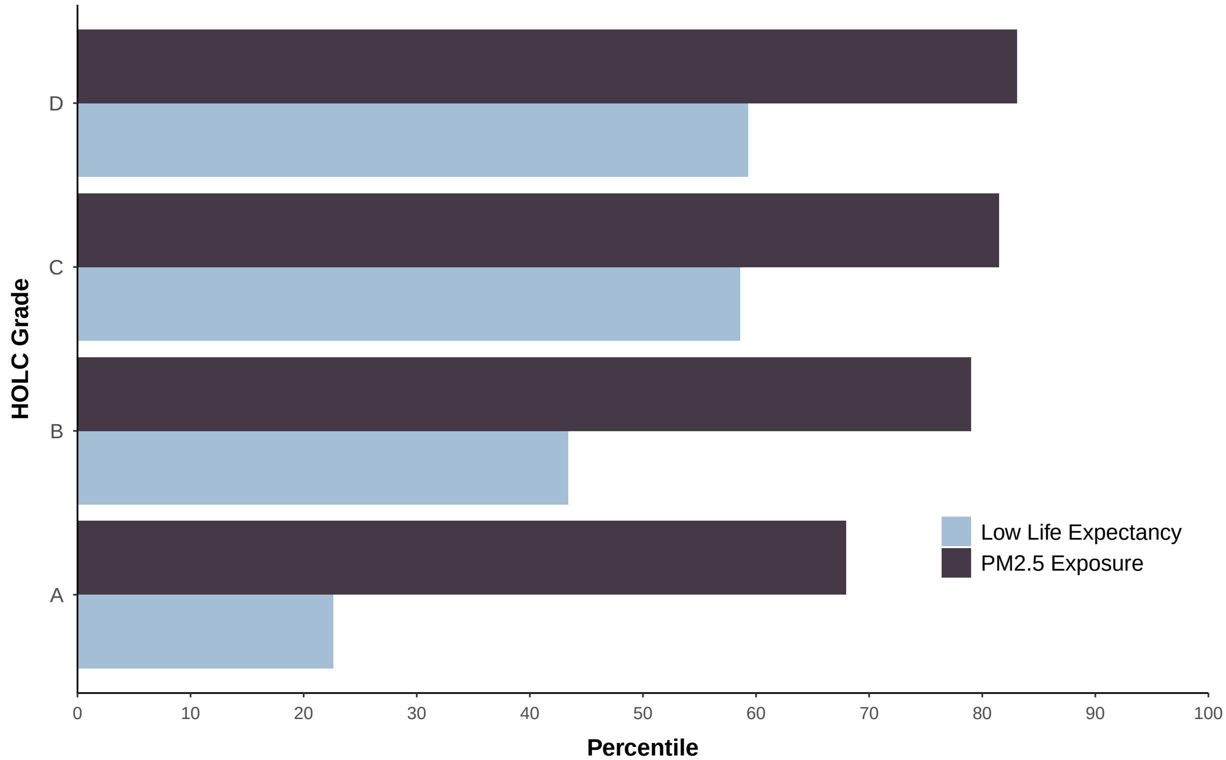

Figure 3: **Health and pollution burdens increase with lower historical HOLC grade in Los Angeles.** Bars show the mean percentile (higher = relatively worse compared to U.S. block groups) for low life expectancy (light blue) and $PM_{2.5}$ exposure (burgundy) across grades A–D. Both indicators rise from A to D—roughly mid-upper percentiles in D areas—indicating that neighborhoods historically rated as "hazardous" by the HOLC grades face systematically greater health and air-quality burdens today. This illustrates the persisting disadvantages faced by these HOLC scores.

iv. Discussion

The analysis reveals a strong and consistent relationship between historical HOLC grades and present-day socioeconomic and environmental conditions across Los Angeles County census block groups. Neighborhoods that were rated most favorably in the 1930s ("A" areas) currently exhibit the lowest percentage of low-income residents—about 18%—while the share increases steadily through grades "B" and "C" to roughly 44% in the historically redlined "D" neighborhoods. This disparity demonstrates how the spatial pattern of investment and disinvestment initiated by redlining continues to shape the economic composition of communities nearly a century later. The persistence of these disparities underscores how past exclusionary lending practices translated into lasting inequalities in household wealth and opportunity.

Environmental and health indicators follow the same pattern. The percentile for low life expectancy, a measure of relative disadvantage in health outcomes, rises from about 23 in "A" areas to nearly 59 in "D" areas, suggesting that residents in historically redlined neighborhoods experience markedly shorter lifespans on average. Similarly, exposure to fine particulate matter ($PM_{2.5}$) increases from the 68th to the 83rd percentile across the same gradient, indicating that poorer air quality is concentrated in areas once deemed "hazardous" for investment. Taken together, these results illustrate a clear legacy of redlining: neighborhoods denied investment decades ago continue to bear disproportionate economic, health, and environmental burdens today, reinforcing the interconnection between social policy, environmental justice, and public health.

Part 2: Legacy of Redlining in Biodiversity Observations

Having briefly examined the relationships between historical redlining and current environmental and socioeconomic factors, Part 2 investigates whether historical redlining is reflected in today's biodiversity data by spatially joining 2021–2023 GBIF bird observations to HOLC grades in Los Angeles. By comparing observation densities across grades, we assess potential sampling bias and the representativeness of community-science data in communities that have faced historic disinvestment.

i. Joining Bird Observations within Redlined Neighborhoods

```
# Use the ejscreen_holc data because it has filtered the HOLC grades to Los Angeles County
birds_holc <- st_join(birds, ejscreen_holc[, "grade"], join = st_within)

# Drop geometry for faster computation and NA values
birds_holc_group <- birds_holc |>
  st_drop_geometry() |>
  drop_na() |>
  group_by(grade) |>
  summarise(n_birds = n()) |>
  mutate(pct = round(100 * (n_birds / sum(n_birds)), 2)) |>
  arrange(grade)
```

ii. Plotting

```
# Create a plot showing the breakdown of bird observation in percent for each HOLC grade
birds_plot <- ggplot(birds_holc_group, aes(x = pct, y = grade)) +
  geom_col(fill = "#A4BED5FF") +
  scale_x_continuous(expand = c(0, 0), limits = c(0, NA)) +
   labs(x = "Bird Observations (% of total)",
    y = "HOLC Grade"
  )  +
  theme_classic() +
   theme(
    # outer whitespace around the whole plot
    plot.margin = unit(c(12, 16, 12, 16), "pt"),
    # space between axis titles and tick labels
    axis.title.x = element_text(size = 15, face = "bold",
```

```
                                 margin = margin(t = 10)),
    axis.title.y = element_text(size = 15, face = "bold",
                                 margin = margin(r = 10)),
    # tick label size and spacing from the axis
    axis.text.x  = element_text(size = 11, margin = margin(t = 6)),
    axis.text.y  = element_text(size = 13, margin = margin(r = 6)),
    # optional: title spacing
    plot.title = element_text(hjust = 0.5, face = "bold",
                               margin = margin(b = 8))
  )

birds_plot
```

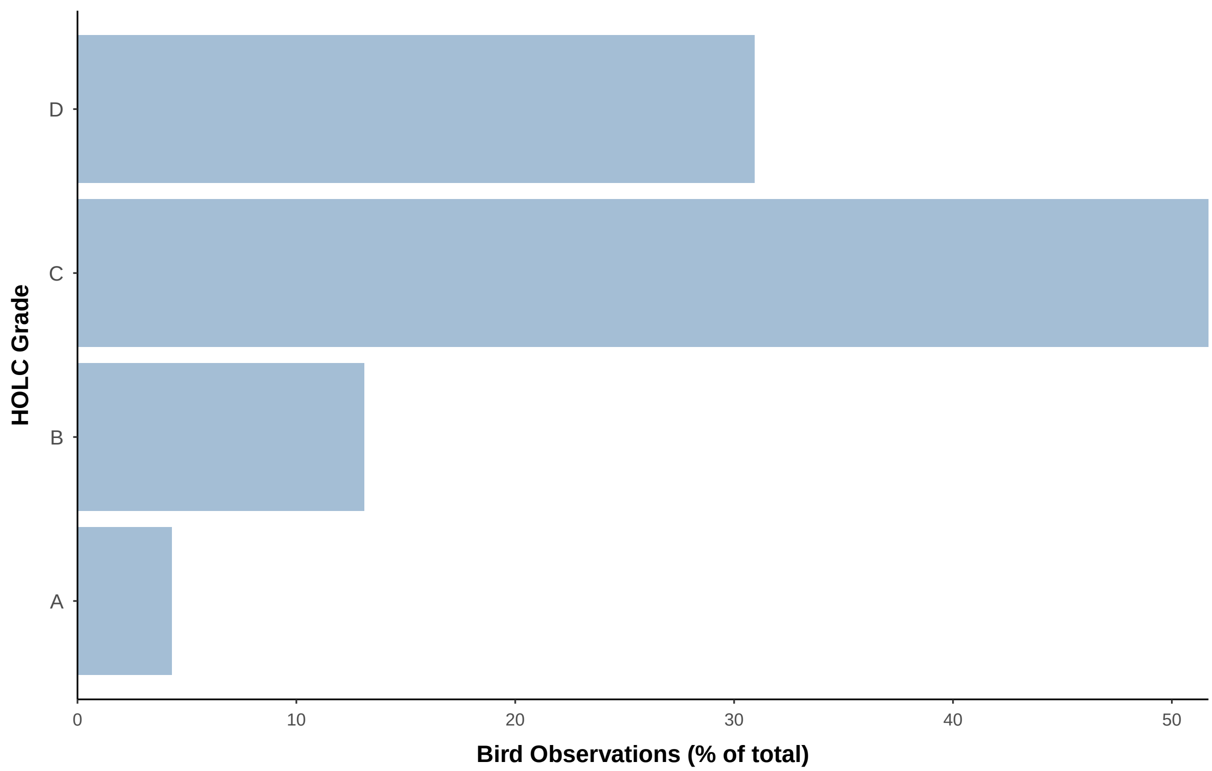

Figure 4: **Unequal Distribution of Bird Observations Across Historical Redlining Grades in Los Angeles.** Bird observations from 2021-2023 are heavily concentrated in neighborhoods with higher HOLC grades (C and D), which were historically rated as less desirable for investment. Grade D neighborhoods (historically redlined) account for 30.77% of observations, while Grade A neighborhoods (historically considered most desirable) represent only 4.29% of observations. This distribution suggests potential spatial bias in citizen science biodiversity data collection, with underrepresentation in the most historically marginalized neighborhoods.

ii. Discussion

Our analysis of bird observations across historical HOLC grades in Los Angeles reveals a pattern of sampling that broadly follows the city's legacy of redlining, though not in the direction reported by national studies. We found that over half of all recorded bird observations (≈51%) occurred in areas graded C, followed by about 31% in D ("hazardous") neighborhoods, 13% in B, and only 4% in A areas. This distribution suggests that, within Los Angeles, the majority of bird observations occur in neighborhoods historically deemed less desirable for investment. This could reflect the city's spatial geography—many C and D zones encompass large, open, or industrial areas with more observation records—or differences in where active birders live or where public parks and observation sites are located today.

These results differ from the national findings of Ellis-Soto et al. (2023), who reported that historically redlined neighborhoods are under-sampled compared to higher-graded areas. The discrepancy likely arises from differences in scale, data scope, and contextual factors. Our analysis focused solely on Los Angeles over a limited time window (2021–2023), while the published study examined nearly 200 metropolitan regions across decades of citizen-science data. The Los Angeles birding community may also have distinct spatial patterns of participation or habitat access that are not representative of broader U.S. trends. Furthermore, we did not control for environmental variables such as vegetation, land cover, or park distribution, which the national study explicitly adjusted for. As a result, while our city-level findings do not show the same direction of sampling bias, they still highlight how spatial and social context—including historical redlining—can influence where biodiversity data are collected and which communities' environments are most visible in ecological research.