

EDS 223 Hw2

Richard Montes Lemus

2025-10-18

```
# Load in libraries
library(tidyverse)
library(sf)
library(here)
library(dplyr)
library(tmap)
library(gt)
```

```
# Read in data
```

```
suppressMessages(ej <- st_read(here("data", "ejscreen", "EJSCREEN_2023_BG_StatePct_with_AS_CNI
```

Reading layer `EJSCREEN_StatePctiles_with_AS_CNMI_GU_VI' from data source

`/Users/richardmonteslemus/Documents/MEDS/EDS-223/eds223-hw2/data/ejscreen/EJSCREEN_2023_BG

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

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

Reading layer `mapping-inequality-los-angeles' from data source

`/Users/richardmonteslemus/Documents/MEDS/EDS-223/eds223-hw2/data/mapping-inequality/mappin

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

```
suppressMessages(bird <- st_read(here("data", "gbif-birds-LA", "gbif-birds-LA.shp")))
```

Reading layer `gbif-birds-LA' from data source

`/Users/richardmonteslemus/Documents/MEDS/EDS-223/eds223-hw2/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

```
# Check if CRS match
```

```
st_crs(bird) == st_crs(ej)
```

```
[1] FALSE
```

```
st_crs(holc) == st_crs(ej)
```

```
[1] FALSE
```

```
# If crs do not match transform them
```

```
if (st_crs(bird) != st_crs(ej)){ # ej has coordinate projection needed for maps
  warning("coordinate reference systems do not match, transforming them to match")
  bird <- st_transform(bird, crs = st_crs(ej)) # Transform bird geo proj to ej proj
}
```

Warning: coordinate reference systems do not match, transforming them to match

```
if (st_crs(holc) != st_crs(ej)){ # ej has coordinate projection needed for maps
  warning("coordinate reference systems do not match, transforming them to match")
  holc <- st_transform(holc, crs = st_crs(ej)) # Transform holc geo proj to ej proj
}
```

Warning: coordinate reference systems do not match, transforming them to match

```
# Check if all data match crs
```

```
st_crs(ej) == st_crs(holc) & st_crs(ej) == st_crs(bird)
```

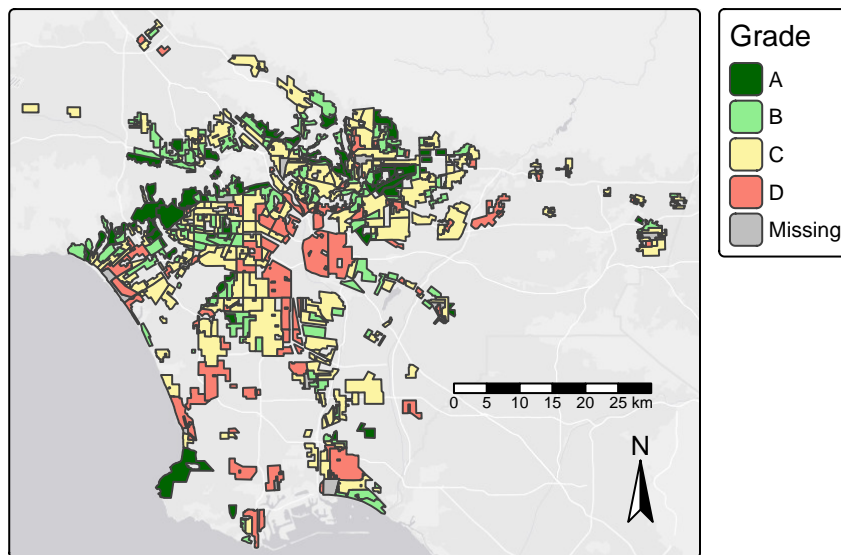
```
[1] TRUE
```

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

1. Create a map of historical redlining neighborhoods

```
# Create map with HOLC grade for LA neighborhoods
tm_shape(holc) +
  tm_polygons(fill = "grade",
              fill.scale = tm_scale(values = c("darkgreen",
                                                "lightgreen",
                                                "#FCF4A3",
                                                "#FA8072")),
              fill.legend = tm_legend(title = "Grade")) +
  tm_title(text = "LA County Home Owners' Loan Corporation Grade") +
  tm_basemap() +
  tm_scalebar() +
  tm_compass()
```

LA County Home Owners' Loan Corporation Grade



2. Create summary table

Census Block Groups by HOLC Grades

```
# Filter ej to LA to avoid working with large data
ej_la <- ej %>%
  filter(CNTY_NAME == "Los Angeles County")

# Clip ej groups based on holc neighborhood
ej_holc <- st_intersection(holc, ej_la) %>%
  st_drop_geometry()

# Get percentage of census groups in each grade
ej_holc_per <- ej_holc %>%
  group_by(grade) %>%
  summarise(group_percent = n() / nrow(ej_holc) * 100) %>%
  rename("HOLC Grade" = grade,
         "Percent of Group" = group_percent)

# Create table for percent of census groups in each grade

ej_holc_per_table <- knitr::kable(
  ej_holc_per,
  col.names = c("HOLC Grade", "Percent of Group"),
  caption = "Los Angeles Census Group by HOLC Grade Percent by ",
  align = "c"
)
ej_holc_per_table
```

Table 1: Los Angeles Census Group by HOLC Grade Percent by

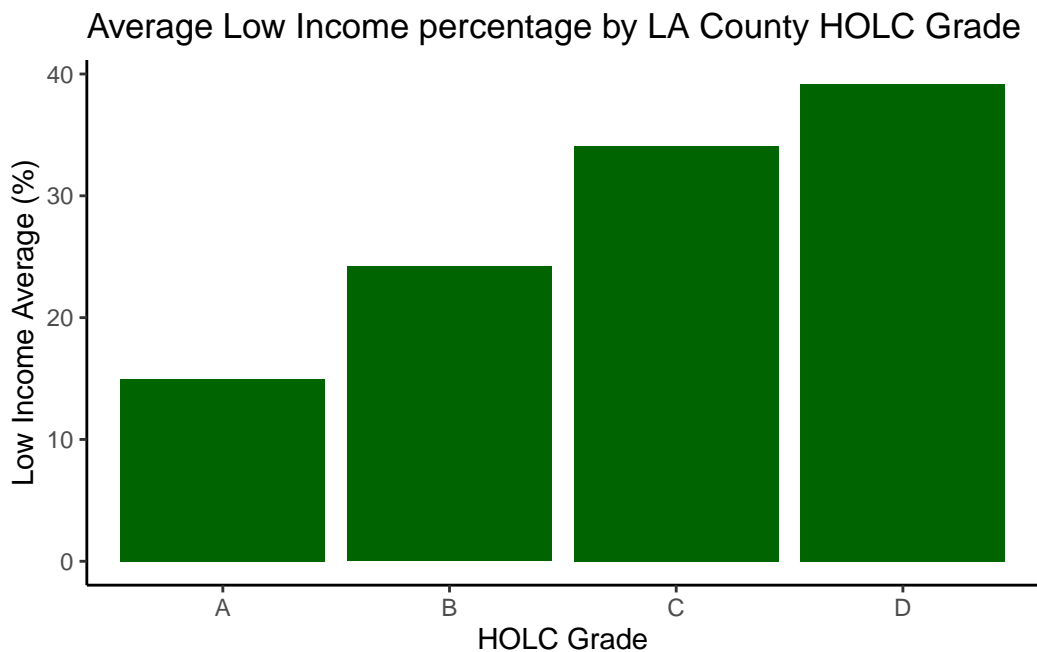
HOLC Grade	Percent of Group
A	7.028804
B	19.395742
C	47.871008
D	21.070758
NA	4.633688

3. Create visualizations for conditions

```
# Calculate income average per group
holc_low_inc <- ej_holc %>%
  filter(!is.na(grade)) %>% # Remove NA row
  group_by(grade) %>%
  summarise(inc_avg = mean(LOWINCPCT) * 100)

if (any(holc_low_inc$inc_avg < 0 | holc_low_inc$inc_avg > 100)) {
  stop("Low income average percentage must be between 0 and 100.")
}
```

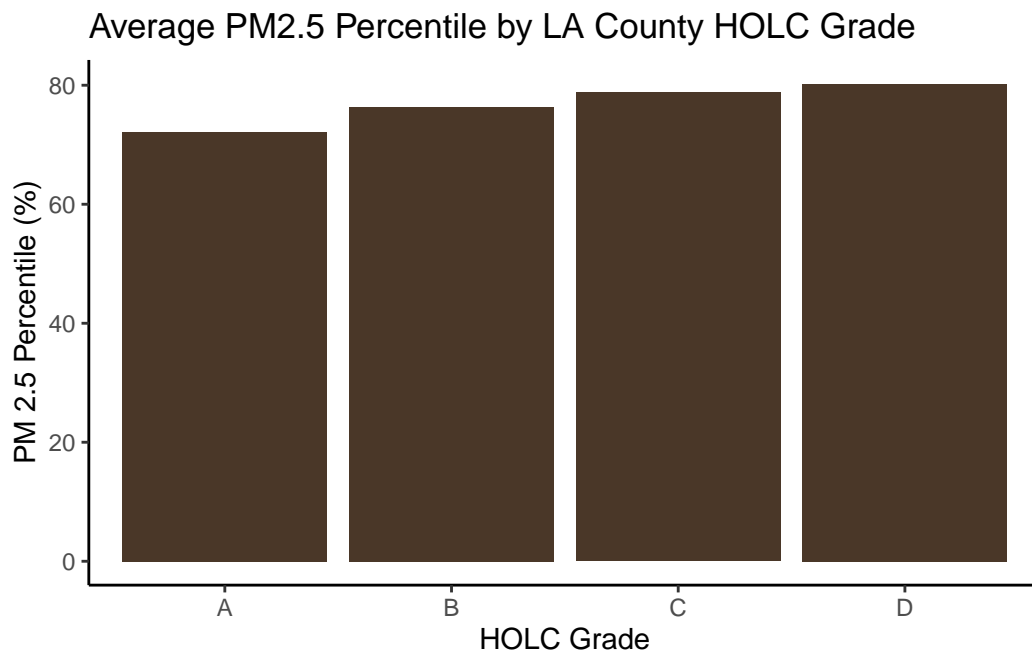
```
# Observe income average difference between grades
holc_low_inc %>%
  ggplot(aes(x = grade, y = inc_avg)) +
  geom_col(fill = "darkgreen") +
  labs(x = "HOLC Grade",
       y = "Low Income Average (%)",
       title = "Average Low Income percentage by LA County HOLC Grade") +
  theme_classic()
```



```
# Calculate percentile PM average per group
holc_pm <- ej_holc %>%
  filter(!is.na(grade)) %>% # Remove NA row
  group_by(grade) %>%
  summarise(pm_avg = mean(P_PM25))

if (any(holc_pm$pm_avg < 0 | holc_pm$pm_avg > 100)) {
  stop("HOLC Grade PM Average must be a percentile between 0 and 100.")
}
```

```
# Observe percentile PM average difference between grades
holc_pm %>%
  ggplot(aes(x = grade, y = pm_avg)) +
  geom_col(fill = "#4A3728") +
  labs(x = "HOLC Grade",
       y = "PM 2.5 Percentile (%)",
       title = "Average PM2.5 Percentile by LA County HOLC Grade") +
  theme_classic()
```



```
# Calculate percentile low life expectancy average per group
holc_age <- ej_holc %>%
  filter(!is.na(grade)) %>% # Remove NA row
```

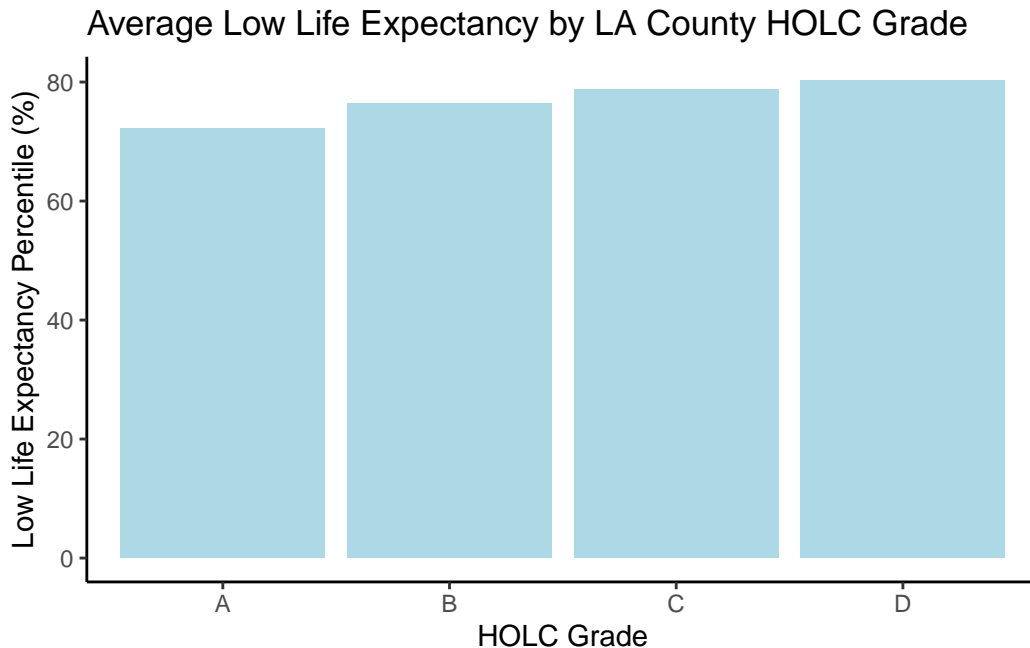
```

group_by(grade) %>%
  summarise(age_avg = mean(P_LIFEEXPPCT, na.rm = TRUE)) # remove na from mean calculation

if (any(holc_age$age_avg < 0 | holc_age$age_avg > 100)) {
  stop("Low life expectancy average must be a percentile between 0 and 100.")
}

# Observe percentile low life expectancy difference between grades
holc_pm %>%
  ggplot(aes(x = grade, y = pm_avg)) +
    geom_col(fill = "lightblue") +
    labs(x = "HOLC Grade",
         y = "Low Life Expectancy Percentile (%)",
         title = "Average Low Life Expectancy by LA County HOLC Grade") +
    theme_classic()

```



The damaging legacy of redlining is apparent across LA County communities. Low-grade neighborhoods, compared to high-grade neighborhoods, tend to have lower income averages, lower life expectancy averages, and are exposed to more particulate matter ≤ 2.5 microns on average.

These trends can be attributed to the fact that low grades were often assigned to poor communities of color. This contributed to neighborhood devaluation and mortgage denial. The lack

of investment in these neighborhoods created a self-reinforcing cycle of poverty in areas that were already low-income. And poverty is correlated with low life expectancy and increased exposure to environmental hazards. This explains why we see this relationship between HOLC grades and the socioeconomic and environmental conditions we observed.

Part 2. Legacy of redlining in biodiversity observations

1. Figure representing percent observations in HOLC grades

```
# Join bird and holc data with st_intersect default
holc_bird <- st_join(bird, holc)
```

```
# Warning tells me a point likely fell on a neighborhood boundary and was duplicated
```

```
if(nrow(bird) == nrow(holc_bird)) {
  print("join matches original data dimensions")
} else {
  warning("join does not match original data dimensions")
  print(paste("bird has", nrow(bird), "rows"))
  print(paste("holc bird has", nrow(holc_bird), "rows"))
}
```

```
Warning: join does not match original data dimensions
```

```
[1] "bird has 1288865 rows"
[1] "holc bird has 1288866 rows"
```

```
# Use stricter join to prevent duplicates
holc_bird_within <- st_join(bird, holc, join = st_within)
```

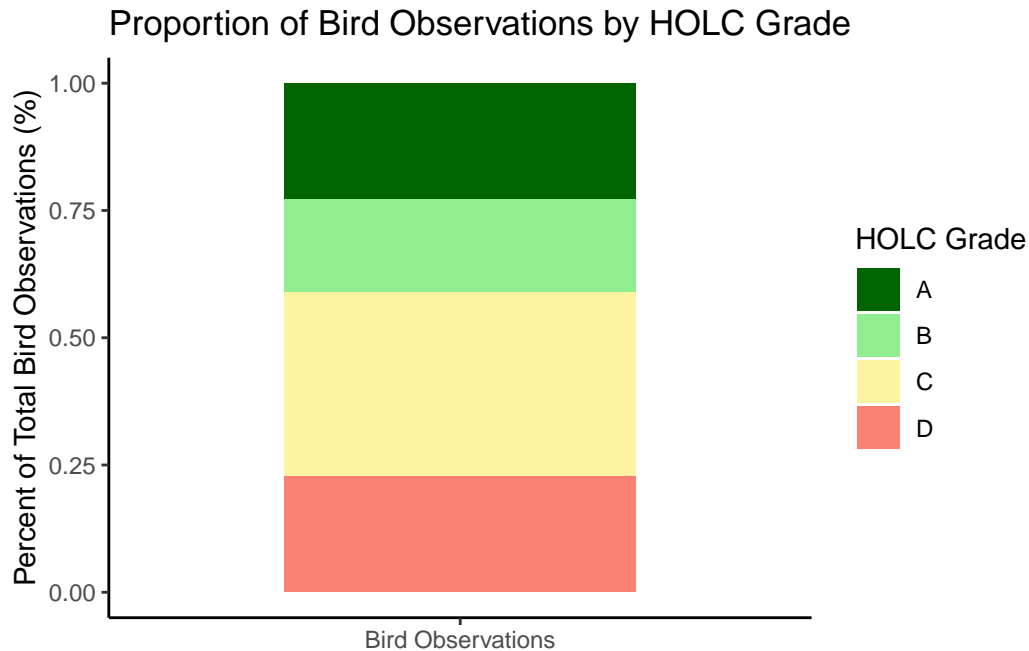
```
# Rows match after join with st_within
if(nrow(bird) == nrow(holc_bird_within)) {
  print("join matches original data dimensions")
} else {
  warning("join does not match original data dimensions")
  print(paste("bird has", nrow(bird), "rows"))
  print(paste("holc bird has", nrow(holc_bird_within), "rows"))
}
```



```
[1] "join matches original data dimensions"
```

```
# Calculate bird observations by grade
bird_obs_grade <- holc_bird %>%
  group_by(grade) %>%
  summarise(bird_obs = n()) %>%
  st_drop_geometry() %>%
  filter(!is.na(grade)) %>%
  mutate(bird_obs_per = bird_obs/sum(bird_obs)) # Get percentage
```

```
# Observe bird observation percentages by grades
ggplot(bird_obs_grade, aes(x = "Bird Observations",
                           y = bird_obs_per,
                           fill = grade)) +
  geom_bar(stat = "identity", width = 0.6) +
  labs(
    title = "Proportion of Bird Observations by HOLC Grade",
    x = NULL, # Remove x from axis label
    y = "Percent of Total Bird Observations (%)",
    fill = "HOLC Grade") +
  scale_fill_manual(values = c("darkgreen",
                                "lightgreen",
                                "#FCF4A3",
                                "#FA8072")) +
  theme_classic()
```



2. Explain Results

The results of this bird observation analysis by grade show that grade C neighborhoods have the most bird observations, while grade D and A neighborhoods have similar amounts. At a glance, these results conflict with the findings from Elis-Soto et al. 2023, which suggest wealthy grade A neighborhoods have the highest bird biodiversity. This is not the case, however, in the study, they use sample density to compare grade neighborhoods, which accounts for land area. My analysis does not account for land; therefore, the higher bird observations in lower-grade neighborhoods may simply be because they cover more land and therefore have more birds than high-grade neighborhoods.