EDS 223 HW2 - Exploring patterns of environmental justice

Lucian Scher

Load Librarys

```r
# Library Required packages
library(sf) # For reading spatial data
library(stars) # For reading raster data
library(tmap) # For mapping
library(tidyverse)
library(here)
library(ggplot2)
library(dplyr)
library(patchwork) # for putting plots side by side
library(gt) # for pretty
```

Read in data

```r
ejscreen_raw <- st_read("data/ejscreen/EJSCREEN_2023_BG_StatePct_with_AS_CNMI_GU_VI.gdb")
```

```
Reading layer `EJSCREEN_StatePctiles_with_AS_CNMI_GU_VI' from data source
  `/Users/lucianscher/Desktop/MEDS/EDS-223/EDS223-HW2/data/ejscreen/EJSCREEN_2023_BG_StatePct
  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
```

```r
mapping_inequality_raw <- st_read('data/mapping-inequality/mapping-inequality-los-angeles.jso
  st_transform(mapping_inequality_raw, crs = (st_crs(ejscreen_raw)))
```

```
Reading layer `mapping-inequality-los-angeles' from data source
  `/Users/lucianscher/Desktop/MEDS/EDS-223/EDS223-HW2/data/mapping-inequality/mapping-inequal
  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
```

```r
gbif <- st_read('data/gbif-birds-LA') |>
   st_transform(gbif, crs = (st_crs(ejscreen_raw)))
```

```
Reading layer `gbif-birds-LA' from data source
  `/Users/lucianscher/Desktop/MEDS/EDS-223/EDS223-HW2/data/gbif-birds-LA'
  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

```r
st_crs(mapping_inequality_raw) == st_crs(gbif)  # See if CRS match
```

```
[1] TRUE
```

```r
st_crs(gbif) == st_crs(ejscreen_raw)
```

```
[1] TRUE
```

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

Data Exploration

```r
head(mapping_inequality_raw) # Show first 10 rows
```

```
Simple feature collection with 6 features and 14 fields
Geometry type: MULTIPOLYGON
Dimension:     XY
Bounding box:  xmin: -13187810 ymin: 4042916 xmax: -13143940 ymax: 4054200
Projected CRS: WGS 84 / Pseudo-Mercator
  area_id city_id grade    fill label name category_id sheets         area
1    7761      16     A #76a865    A1                1      1 3.359915e-04
2    7775      16     A #76a865   A10                1      1 1.814147e-04
3    7808      16     A #76a865   A11                1      1 5.978184e-05
4    8025      16     A #76a865   A12                1      1 2.587288e-04
5    7608      16     A #76a865   A13                1      1 1.326238e-04
6    7797      16     A #76a865   A14                1      1 2.629402e-04
                                                                              bound
1                    [ [ 34.136969999999998, -118.46807 ], [ 34.153350000000003, -118.42031 ]
2                                [ [ 34.1691, -118.11198 ], [ 34.188940000000002, -118.0979 ]
3                    [ [ 34.155970000000003, -118.11301 ], [ 34.163069999999998, -118.09853 ]
4 [ [ 34.122999999999998, -118.18331000000001 ], [ 34.159080000000003, -118.15949999999999 ]
5         [ [ 34.127760000000002, -118.18619 ], [ 34.146239999999999, -118.16952000000001 ]
6 [ [ 34.105049999999999, -118.12802000000001 ], [ 34.130420000000001, -118.07402999999999 ]
  residential commercial industrial     label_coords
1        TRUE      FALSE      FALSE 34.147, -118.452
2        TRUE      FALSE      FALSE 34.177, -118.104
3        TRUE      FALSE      FALSE 34.159, -118.102
4        TRUE      FALSE      FALSE 34.148, -118.171
5        TRUE      FALSE      FALSE 34.133, -118.175
6        TRUE      FALSE      FALSE 34.119, -118.105
                        geometry
1 MULTIPOLYGON (((-13186622 4...
2 MULTIPOLYGON (((-13148113 4...
3 MULTIPOLYGON (((-13148279 4...
4 MULTIPOLYGON (((-13154761 4...
5 MULTIPOLYGON (((-13155134 4...
6 MULTIPOLYGON (((-13149259 4...
```

```r
dim(mapping_inequality_raw) # Check size of dataframe
```

```
[1] 417  15
```

```r
colnames(mapping_inequality_raw) # See column names
```

```
 [1] "area_id"      "city_id"      "grade"        "fill"         "label"
 [6] "name"         "category_id"  "sheets"       "area"         "bounds"
[11] "residential"  "commercial"   "industrial"   "label_coords" "geometry"
```

1. Neighborhoods by HOLC Grade

```r
# Create grade map
tm_shape(mapping_inequality_raw) + # Map data
  tm_graticules() + # With gridlines
  tm_polygons(fill = 'grade', # Color by grade
              fill.legend = tm_legend(title = "HOLC Grade")) + # Legend tile
  tm_title(text = "Neighborhoods by HOLC Grade") # Title
```

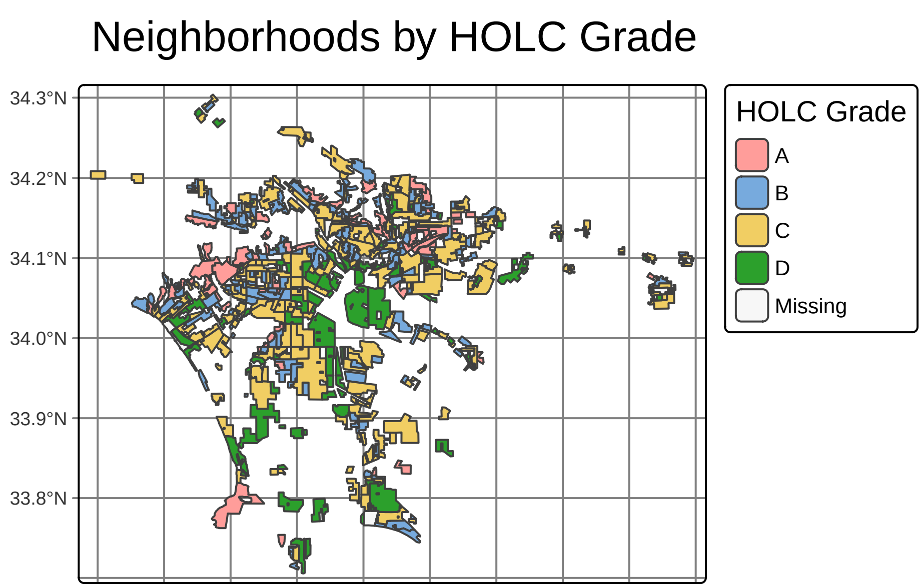


Table Summary

```
# Attach HOLC grades to each EJScreen census block group

ej_holc <- st_join(ejscreen_raw, mapping_inequality_raw) |>
  st_drop_geometry() # Drop unneeded geometries

#glimpse(ej_holc)
#sum(is.na(ej_holc$grade))
#colnames(ej_holc)
```

```
# Create table summarizing grades
ej_holc_grade_sum <-
  ej_holc |>
  group_by(grade) |> # Group by grade
  summarise(n_block_groups = n()) |> # Summaries by number in each group
  mutate(grade = ifelse(is.na(grade), "No Grade", grade)) |> # Change NA to None
  ungroup() |>
  mutate( # Add column of group number out of 100%
    percent = 100 * n_block_groups / sum(n_block_groups))

ej_holc_grade_sum |> gt() |>
  cols_hide(columns = n_block_groups) |> # Hide summarization number
  tab_header(
    title = "Grade Percentage summary table") |>
  fmt_percent(
    columns = percent,
    scale_values = FALSE) # tell gt the values are 0-100
```

Visualizations

```
# Calculate mean of each variable grouped by HOLC grade.
ej_holc_means <- ej_holc |>
  mutate(grade = ifelse(is.na(grade), "None", grade)) |>   # Replace NA with "No Grade"
  group_by(grade) |>
  summarise(
    mean_low_income = mean(LOWINCPCT, na.rm = TRUE), # Add new columns with means
    mean_pm25 = mean(P_PM25, na.rm = TRUE),
    mean_low_life_expectancy = mean(P_LIFEEXPPCT, na.rm = TRUE)
```

[

Grade Percentage summary table]

Grade Percentage summary table	
grade	percent
A	0.18%
B	0.50%
C	1.25%
D	0.55%
No Grade	97.52%

```
  ) |>
  ungroup()
```

Plot 1: Mean of % low income

```
# Plot mean of % low income by grade
plot1 <- ggplot(ej_holc_means, aes(x = grade, y = mean_low_income, fill = grade)) +
  geom_col() +
  labs(
    title = "Mean % Low Income
    by HOLC Grade",
    x = "HOLC Grade",
    y = "Mean % Low Income"
  ) +
  theme_minimal() +
    scale_fill_brewer(palette = "Reds")
```

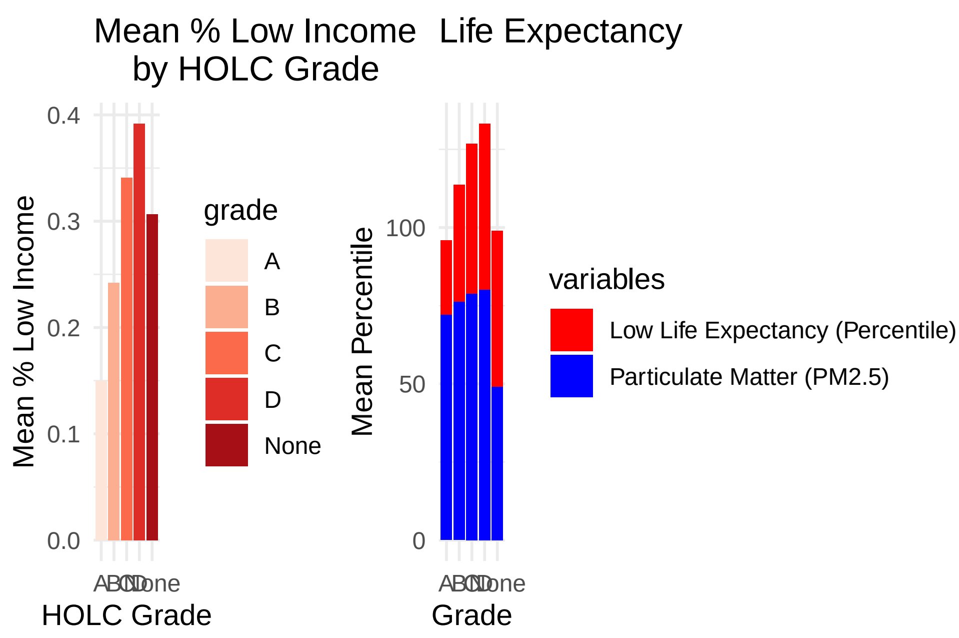
Plot 2: Low life expectancy percentiles on top of Particulate matter percentiles

```
# Pivot long version for plotting
ej_holc_long <- ej_holc_means |>
  pivot_longer(
    cols = c(mean_pm25, mean_low_life_expectancy), # The columns you want to plot
    names_to = "variables",                        # New column for the variables
```

```r
    values_to = "percentile_means"                  # New column for the values
  )
```

```r
#Plot percentile of Particulate matter next to percentile of low life expectancy
plot2 <- ggplot(data = ej_holc_long, aes(x = grade, y = percentile_means, fill = variables))
  geom_col() + #(position = "dodge") + # geom_col for bar chart, "dodge" for side-by-side ba
  labs(
    title = "Life Expectancy",
    by = "Grade",
    x = "Grade",
    y = "Mean Percentile",
    fill = "variables"
  ) +
  scale_fill_manual(
    values = c("mean_pm25" = "blue", "mean_low_life_expectancy" = "red"),
    labels = c("mean_pm25" = "Particulate Matter (PM2.5)", "mean_low_life_expectancy" = "Low
  ) +
  theme_minimal()
```

```r
# Show plots together using patchwork
(plot1 + plot2) +
  theme(plot.margin = margin(r = 20))
```

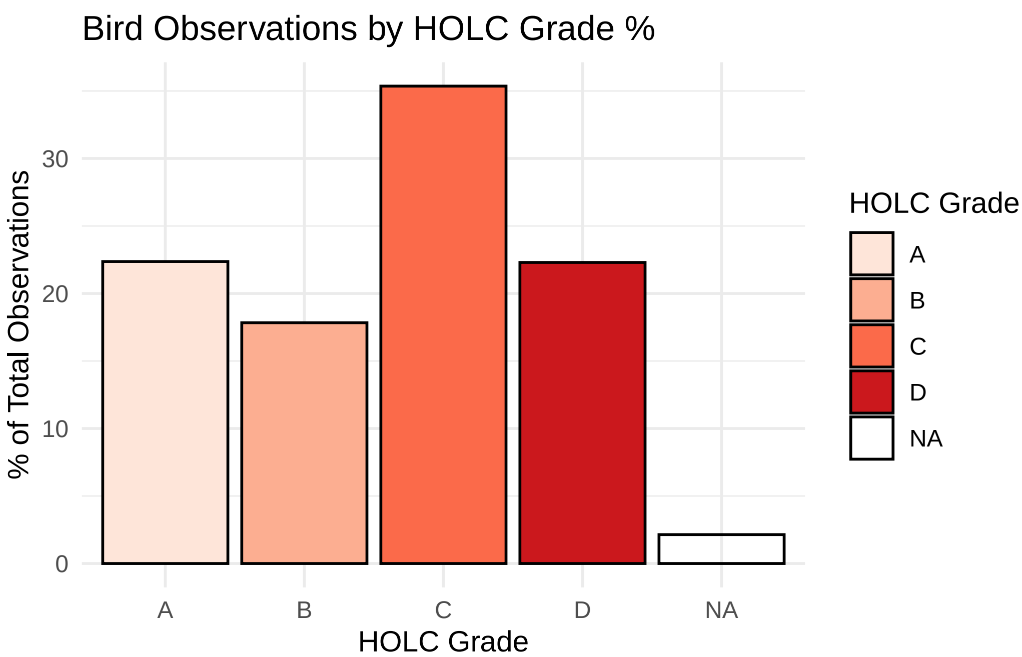


Both plots show a clear correlation between HOLC grade and environmental injustices. While it is not surprising that lower grades of “residential security” are related to lower income, it does set a precedent for unfair environmental and socioeconomic conditions. When compared directly next to low life expectancy and particulate matter we can see that the historical redlining grades represent lower quality of life and even shorter life for people in worse graded areas. We saw how red lined districts had less trees which is likely a direct causality of worse particulate matter and therefore reasonably can be presumed to have an effect on low life expectancy.

Part 2: Legacy of redlining in biodiversity observations

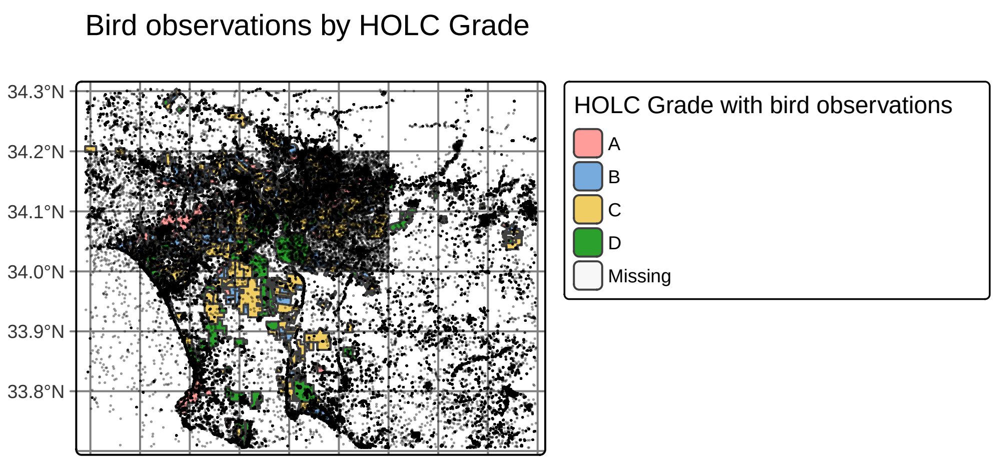
```
# Join gbif (Bird Observation data) and mapping_inequality_raw (HOLC grade data)

HOLC_birds <- st_join(mapping_inequality_raw, gbif)
```

```
# Plot percentage of bird observations within each HOLC grade
ggplot(HOLC_birds, aes(x = grade, fill = grade)) +
  # Calculate percentages and ensure plotting happens after the transformation
  geom_bar(aes(y = after_stat(count / sum(count) * 100)), color = "black") +
 scale_fill_brewer(palette = "Reds") +
  labs(
    title = "Bird Observations by HOLC Grade %",
    x = "HOLC Grade",
    y = "% of Total Observations",
    fill = "HOLC Grade"
  ) +
  theme_minimal()
```




```
# Create map of bird observations over redlining data map
tm_shape(HOLC_birds) + # Map data
  tm_graticules() + # With gridlines
  tm_polygons(fill = 'grade', # Color by grade
              fill.legend = tm_legend(title = "HOLC Grade with bird observations")) +
tm_shape(gbif) +
  tm_dots(size = 0.02,
          fill_alpha = 0.4,
          col = "darkblue") +
  tm_title(text = "Bird observations by HOLC Grade") # Title
```

```
[plot mode] fit legend/component: Some legend items or map compoments do not
fit well, and are therefore rescaled.
i Set the tmap option `component.autoscale = FALSE` to disable rescaling.
```



Reflection

Here we can see a clear lack of either birds or bird observation data in red lined zones. This is clear by the appearance of yellow and blue zones that are hardly covered by bird observations in contrast with the blue and pink zones we can barely see because they are swarmed. The Diego Ellis Soto paper considers how our result may not be simply less birds, but instead less data collection going on in red lined zones in general. Our bird observation percentage plot shows a more balanced reporting of bird observations which may be a result of us not considering the unreliability of citizen science.