

# Correlates and Policy Remedies of DC's Urban Heat Island Effect

Evan Bowman

December 06, 2022

# Introduction

A major discussion in today's politics is the consequences of climate change and what is to come if no meaningful attempt at mitigating atmospheric greenhouse gasses occurs. These consequences, on a meta level, range from higher mean temperatures, sea level rise, and changes in the ability to grow sustainable food. These potential consequences poses large public health risks such as increases in heat related deaths and illness. This is cause for concern as about 25 percent of natural hazard mortality in the country is caused by heat exposure (Benz and Burney 2021). Furthermore, heat-related mortality in the USA causes roughly 1,500 more deaths annually than other severe weather events, as well as other outcomes such as heat strokes, dehydration, loss of labor productivity, and decreased learning (Hsu et al. 2021). The potential increase in heat related health incidents requires a substantial need for policy to combat increases in mean temperatures.

Little discussion looks at the discrepancies these effects might have on urban populations. On the topic of increased global temperature averages, there is a need to look at the intersectionality of an already occurring phenomenon: urban heat island effects. This concept revolves around the key fact that urbanized city centers already create higher mean temperatures compared to their rural counterparts. Not only are these urban centers more likely to have higher mean temperatures than rural areas, there are clear sociodemographic discrepancies between populations within these cities. Few studies have been published analyzing these demographic discrepancies. However, in one report, research has found that roughly 77 percent of counties studied show that there are large differences in change of daytime temperature between census tracts with the highest and lowest Black population proportions. Furthermore, 81 percent of counties studied with lower difference in daytime temperature are majority non-Hispanic White. (Benz and Burney 2021). There is also empirical evidence that shows discrepancies in urban heat island effects targeting communities that are more impoverished. For example, a 2020 study found that there is a negative correlation of .88 between cities' urban heat island effects and income (Chakraborty et al. 2020). This is further cause for concern when looking at potential correlations between poverty and race in city census tracts.

These findings are consistent with research that shows majority minority populations in urban centers have less tree cover and larger proportions of impervious surfaces than majority non-Hispanic White populations. This shows a clear need for social policy that protects majority minority census tracts from being disproportionately affected by increased heat island effects. Research regarding the effects of urban heat island effects must be on a local level to adequately identify meaningful social demographic trends for areas within urban centers. These could be specific neighborhoods, districts, census tracts, etc. In the context of racial inequities in Washington DC, most research is done at the ward level. For this study, census tracts are chosen as sample observations to be able to identify which specific areas of DC's eight wards are most at risk of the urban heat island effect, and the demographic patterns that coincide with these risks.

## Research Questions

**What are reoccurring sociodemographic trends that are associated with higher heat sensitivity indexes?** Who are the communities at a larger risk of heat-induced health events due to higher heat sensitivity index (HSI)? HSI is a metric that describes an individuals ability to adapt, cope, or recover from extreme heat and is explained by specific demographic and health variables of DC's census tracts. Identifying socioeconomic and health trends can assist policy makers in identifying the communities who would best benefit from policy implementation.

**What are the environmental metrics that are correlated with the census tracts with the highest HSI?** What environmental characteristics of certain census tracts create larger heat exposure indexes (HEI)? In the data, HEI is a metric that takes a census tract's mean ambient air temperature as the exposure variable (50 percent) and two physical variables (impervious surfaces: 25 percent, lack of tree cover: 25 percent) that describe a tract's heat retention. Understanding the environmental factors would help in identifying policy approaches that would mitigate the consequences of the urban heat island effect by assisting at risk census tracts with solutions to heat retention.

**Do these findings highlight environmental justice issues that exist within Washington DC?** Do the findings of this study concur with previous research regarding race inequities for heat exposure and sensitivity in urban centers? The ultimate goal in answering these two research questions is to find the intersectionality of environmental that exacerbate the urban heat island effect and the populations of urban centers most at risk.

## Heat Sensitivity Index and Socioeconomic Patterns

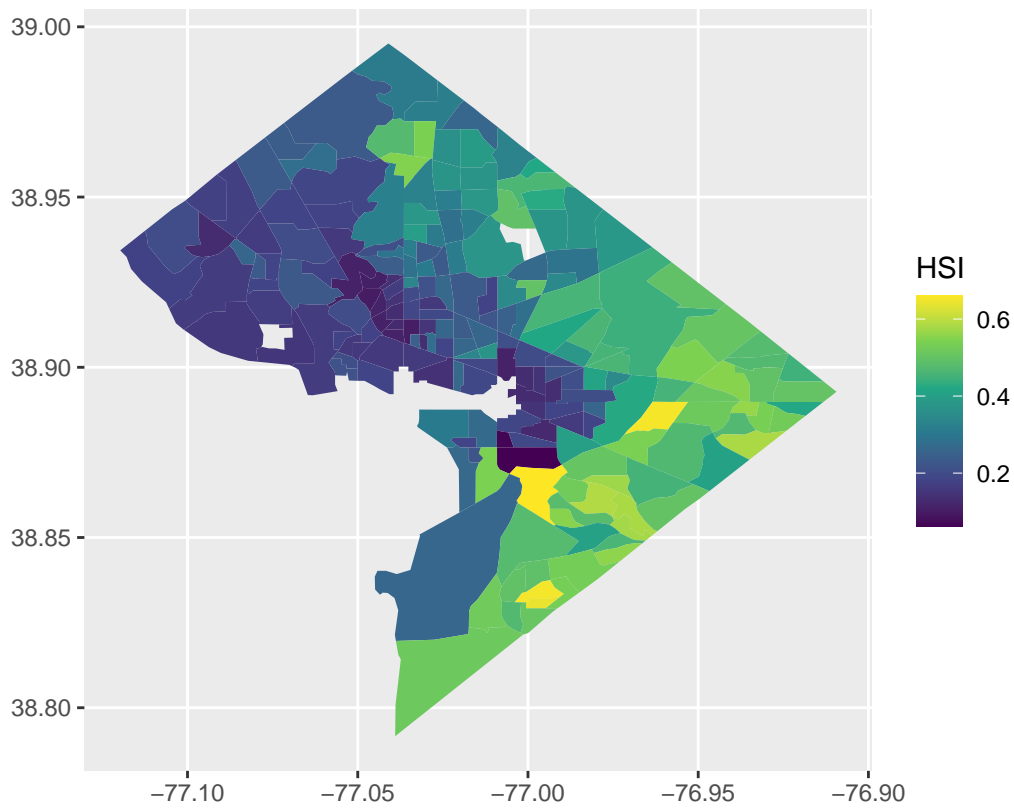


Fig. 1

As seen from the Figure 1, the census tracts with the highest HSI levels fall within areas along the Anacostia River, also found within Wards 7 and 8. These census tracts are a vast contrast when compared to census tracts with low HSI levels found in the Northwest portion of the city. Analysis conducted in this section will look to define what specific patterns exist within tracts with high HSI levels.

## Visualizing Patterns

Areas along the Anacostia River are typically known for being predominately majority minority as well as lower income compared to census tracts in the Northwest. This further contributes to the hypothesis of racial inequity in DC's Urban Heat Island Effect as the intersectionality of poverty and race exacerbates individuals ability to cope with extreme heat (Fig. 2).

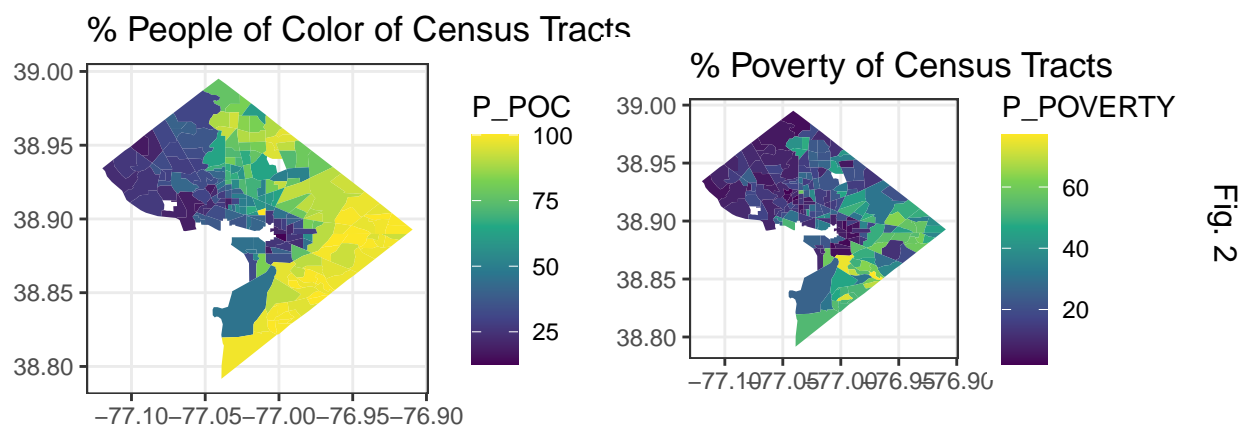


Fig. 2

Other interesting patterns found within the data are health trends and specific discrepancies between predominately non-Hispanic White census tracts and predominately Black census tracts. For all four health variables found within the data set, asthma, obesity, coronary heart disease (CHD), and disability, majority minority census tracts exhibit higher prevalence than predominately White tracts (Fig. 3). This highlights the increased risk of health related events that are exacerbated by a higher sensitivity to heat and the inability to cope or adapt to it.

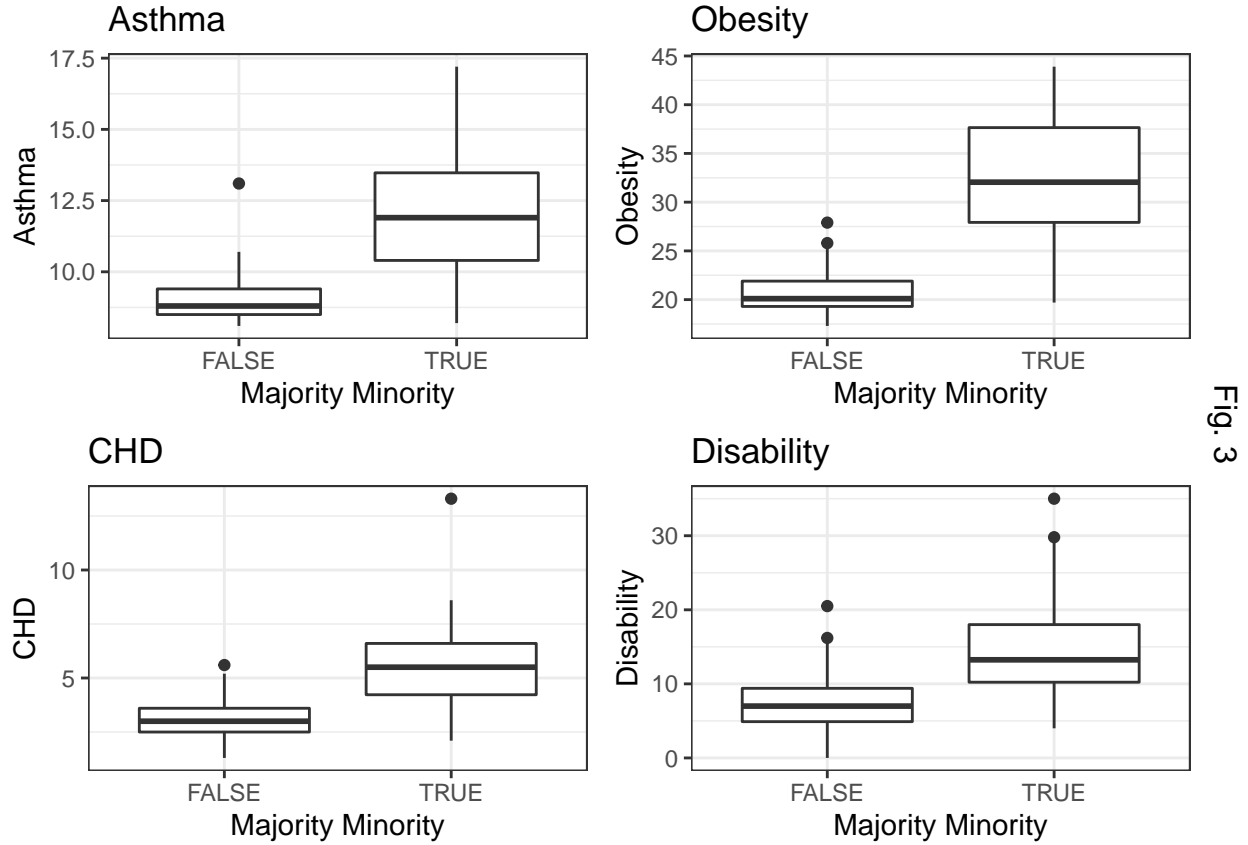


Fig. 3

With preliminary patterns relating to a census tract's Heat Sensitivity Index established, this paper will shift to creating a preliminary model that best exemplifies which variables contribute most to HSI metrics for a given census tract.

## Regression Definition

### Preliminary Regression Identification

To best understand to what magnitude socioeconomic factors affect HSI levels, preliminary regression models will be created to identify variables of statistical significance. To achieve the best possible model, best subsets regression, forward stepwise, and backward stepwise will be conducted and the best model will be taken following the analysis of adjusted r-squared, Mallows CP, and BIC.

The first model will be utilizing the best subsets regression strategy through the leaps package. The model was fitted and the best eight models are shown below. Summary statistics of each model will be analyzed to figure which model will be chosen (Table 1).

Table 1: Best Subsets Regression

model	adjr2	cp	bic
1	0.8979830	4521756.42	-453.7521
2	0.9350345	2865107.80	-541.0617
3	0.9555091	1952266.17	-613.6160
4	0.9727876	1188009.46	-709.1226
5	0.9859260	611231.14	-838.6844

Table 2: Forward and Backward Stepwise

model	adjr2	cp	bic	model	adjr2	cp	bic
1	0.8979830	4521756.42	-453.7521	1	0.8141234	8238877.04	-331.9635
2	0.9350345	2865107.80	-541.0617	2	0.9095647	3988449.75	-473.9139
3	0.9555091	1952266.17	-613.6160	3	0.9351507	2845681.30	-537.1296
4	0.9719987	1222453.67	-703.3216	4	0.9539562	2010260.43	-602.3611
5	0.9859260	611231.14	-838.6844	5	0.9800948	864557.15	-768.3144
6	0.9911926	380491.58	-929.5581	6	0.9896876	445543.48	-897.5332
7	0.9960067	171535.88	-1085.8489	7	0.9950826	211270.99	-1043.5956
8	0.9977593	95677.64	-1198.8777	8	0.9979555	87284.43	-1217.4780

model	adjr2	cp	bic
6	0.9911926	380491.58	-929.5581
7	0.9960067	171535.88	-1085.8489
8	0.9979555	87284.43	-1217.4780

Model 5 returned the highest adjusted r-squared value, lowest Mallows' Cp value, and second lowest BIC value. This shows that the model captures the highest variation in HSI, the best predicting power for all possible models, and the second lowest bias for all possible models using the Best Subsets method. Forward and Backwards Stepwise regression returned the same results (Table 3). Thus, the model selected will include the variables percent people of color (P\_POC), percent elderly (P\_ELDERLY), percent poverty (P\_POVERTY), percent limited English (P\_LIMENG), and asthma (ASTHMA).

The fitted regression model is shown below

```
## # A tibble: 6 x 5
##   term          estimate std.error statistic  p.value
##   <chr>          <dbl>    <dbl>    <dbl>    <dbl>
## 1 (Intercept) -0.190    0.0228     -8.33 4.00e-14
## 2 P_POC        0.00271  0.000163    16.6 3.57e-36
## 3 P_ELDERLY    0.00461  0.000313    14.7 3.00e-31
## 4 P_POVERTY    0.00152  0.000270     5.61 8.79e- 8
## 5 P_LIMENG     0.00160  0.000524     3.04 2.75e- 3
## 6 ASTHMA      0.0218   0.00303     7.20 2.43e-11
```

## Shrinkage Regression

While best subsets regression identified the best model as including five variables, there is clear covariation that affects the magnitude of the coefficients (Appendix A). As a result, a LASSO regression model will be fitted and compared to the original regression model. LASSO takes into account the covariation of the predictors, and adjusts the coefficients accordingly.

```
## # A tibble: 6 x 5
##   term          estimate std.error statistic  p.value
##   <chr>          <dbl>    <dbl>    <dbl>    <dbl>
## 1 (Intercept) -0.190    0.0228     -8.33 4.00e-14
## 2 P_POC        0.00271  0.000163    16.6 3.57e-36
## 3 P_ELDERLY    0.00461  0.000313    14.7 3.00e-31
## 4 P_POVERTY    0.00152  0.000270     5.61 8.79e- 8
## 5 P_LIMENG     0.00160  0.000524     3.04 2.75e- 3
## 6 ASTHMA      0.0218   0.00303     7.20 2.43e-11
```

# Environmental Factors of Heat Exposure Indexes

It is important to understand not only socioeconomic factors that contribute to a census tract's HSI statistic, but also the environmental factors that contribute to their Heat Exposure Index. However, unlike HSI, HEI is defined in for each census tract as 50 percent mean air temperature, 25 percent lack of tree canopy, and 25 percent impervious surface. As a result, analysis will be centered around patterns of three variables as opposed to any model building to predict HEI values.

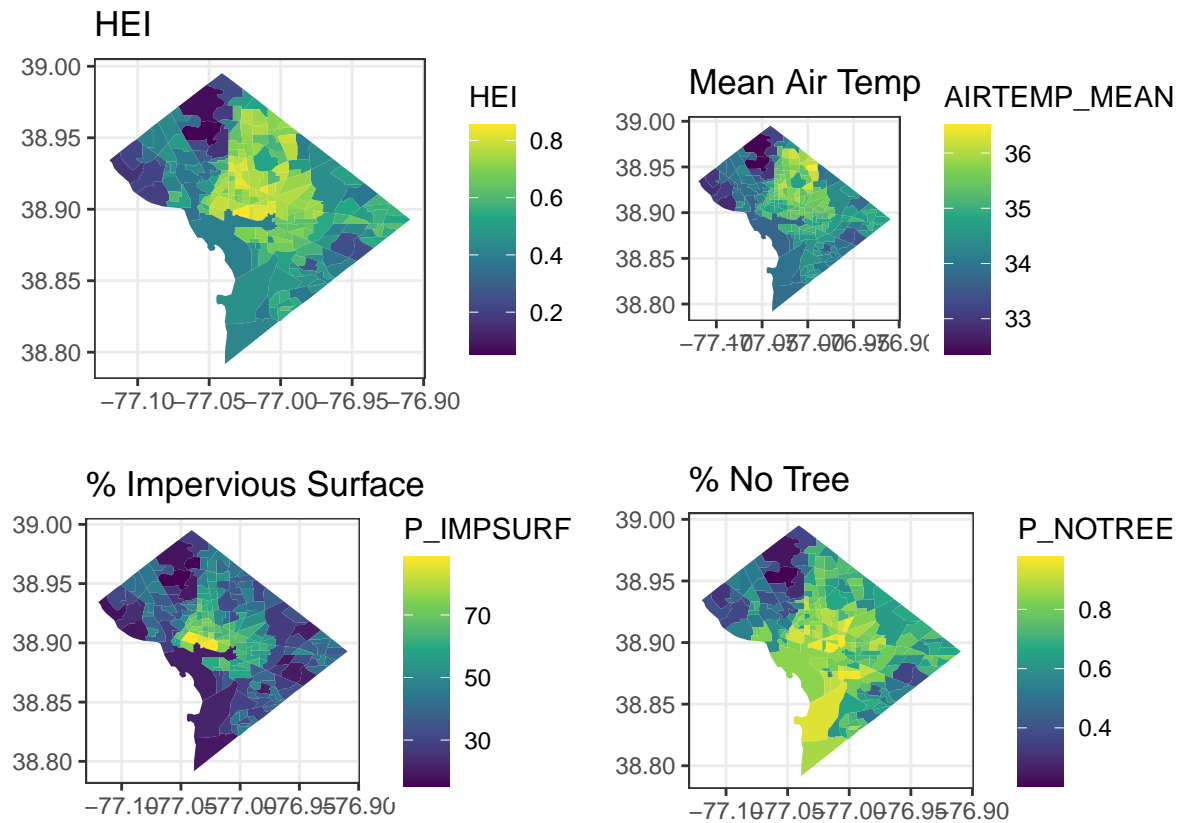


Fig. 7

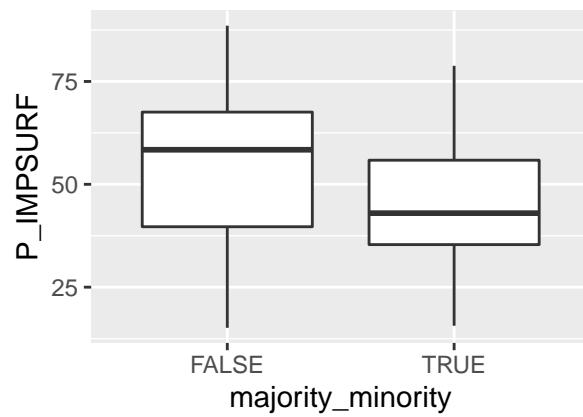
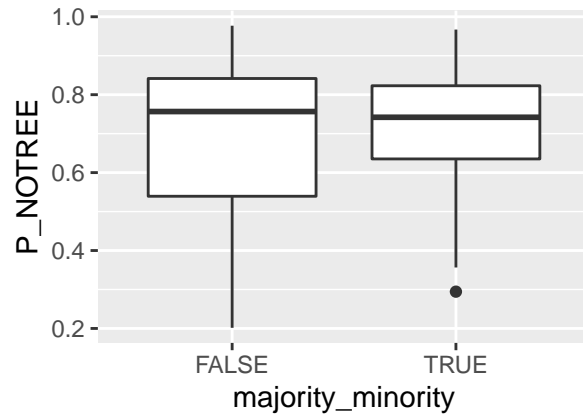
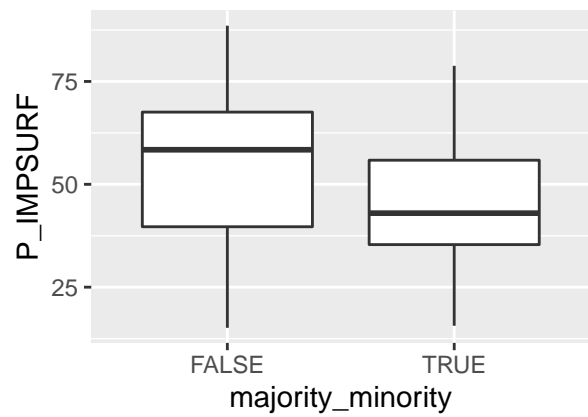
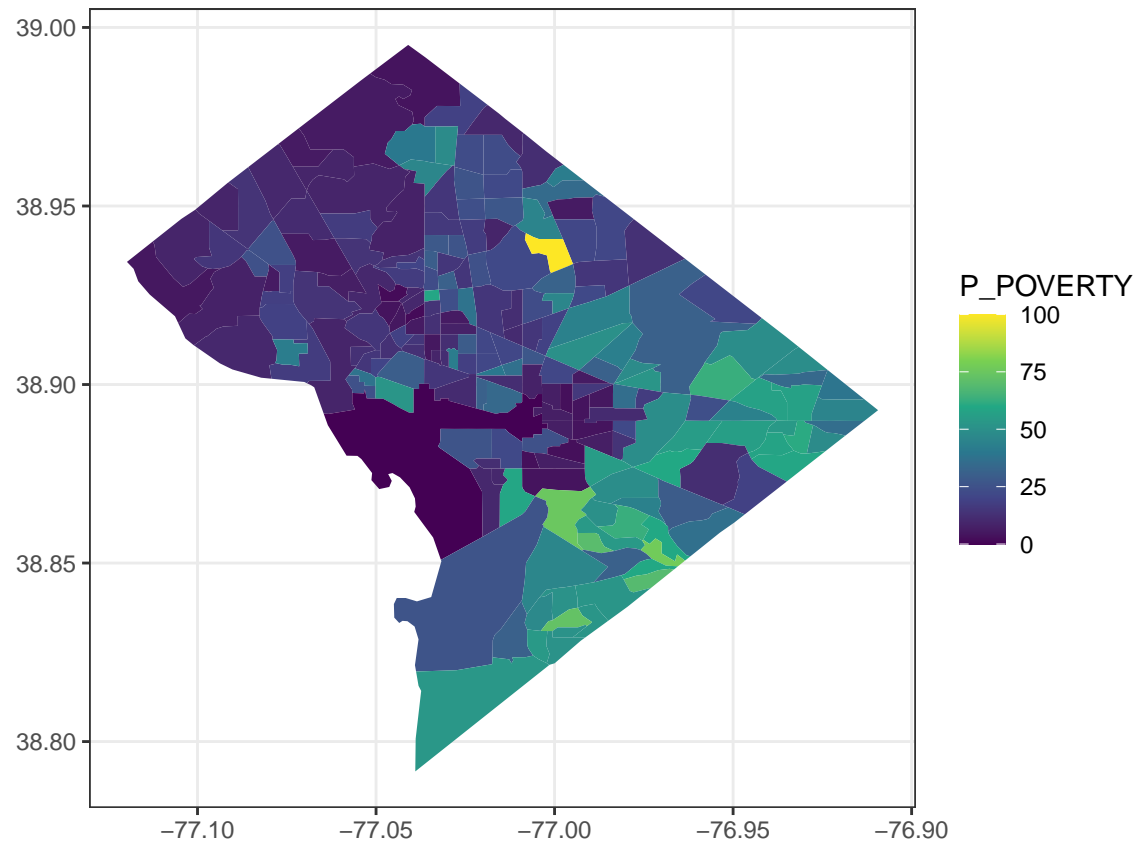
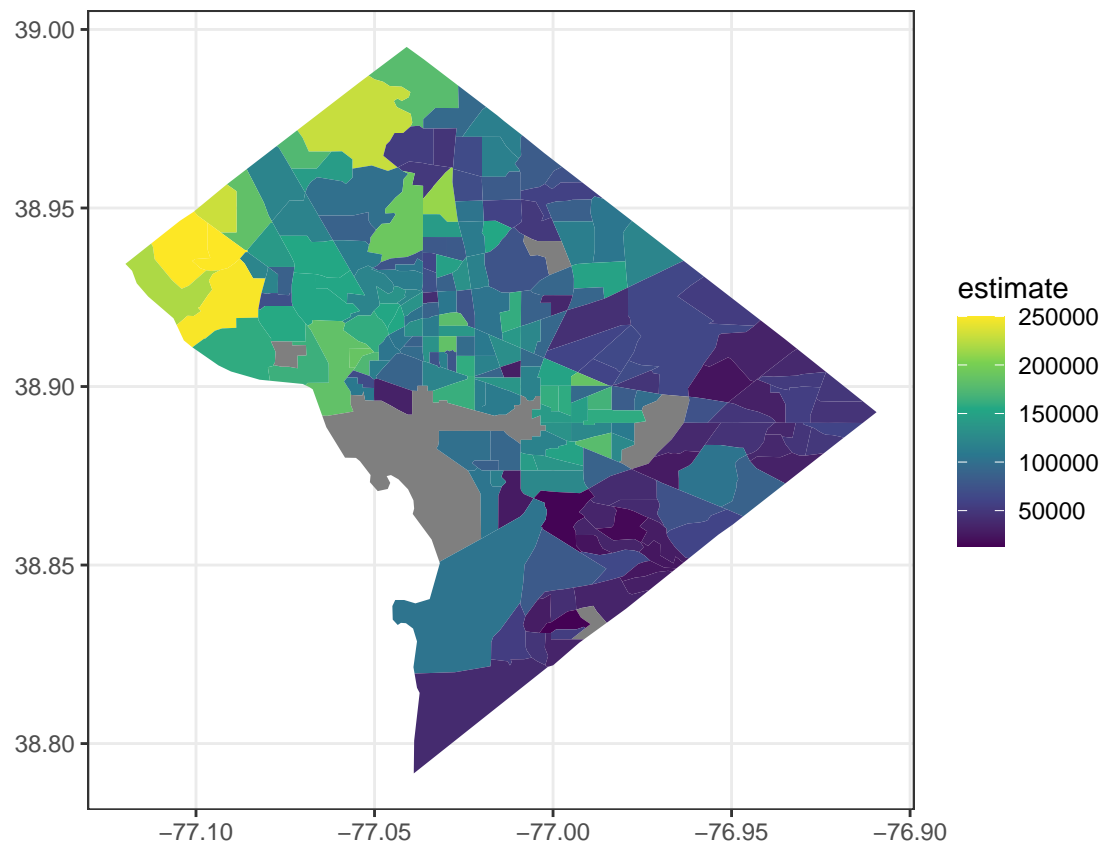
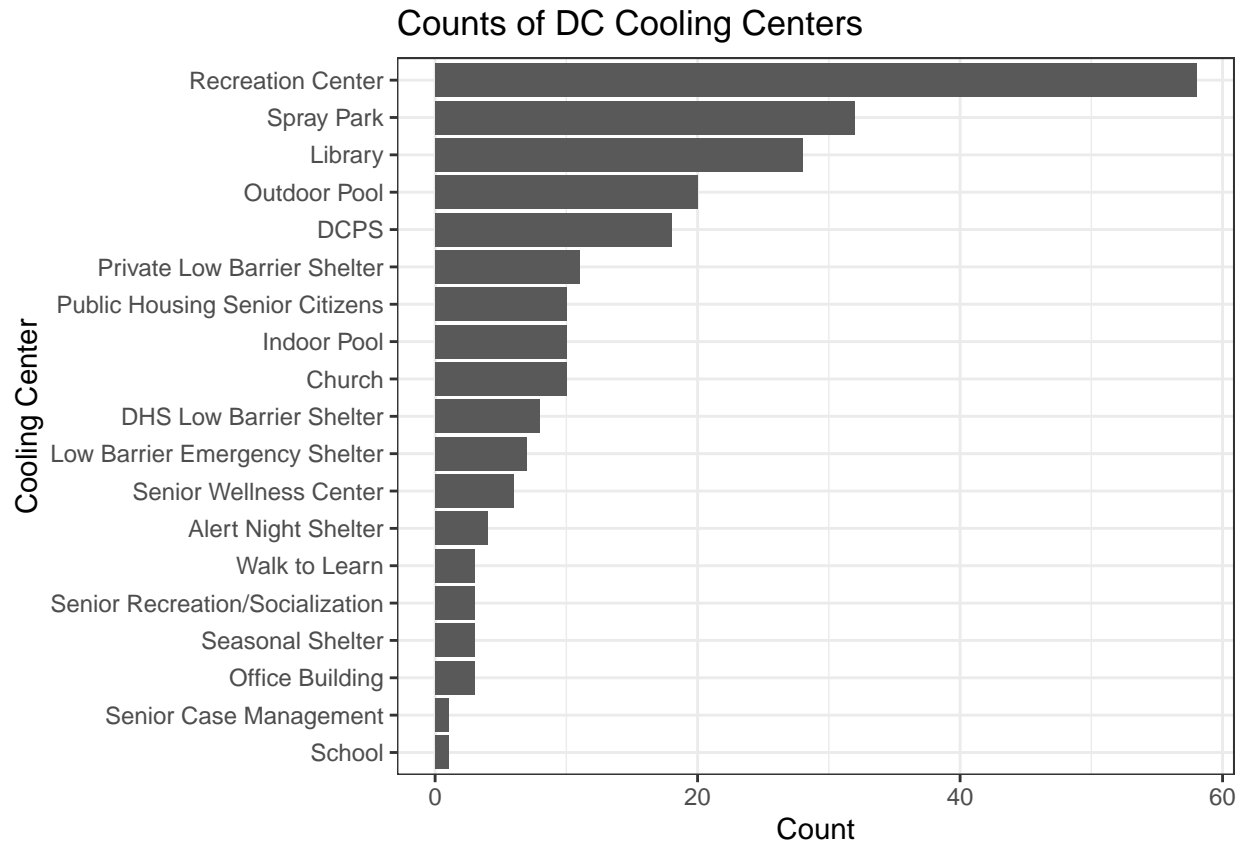


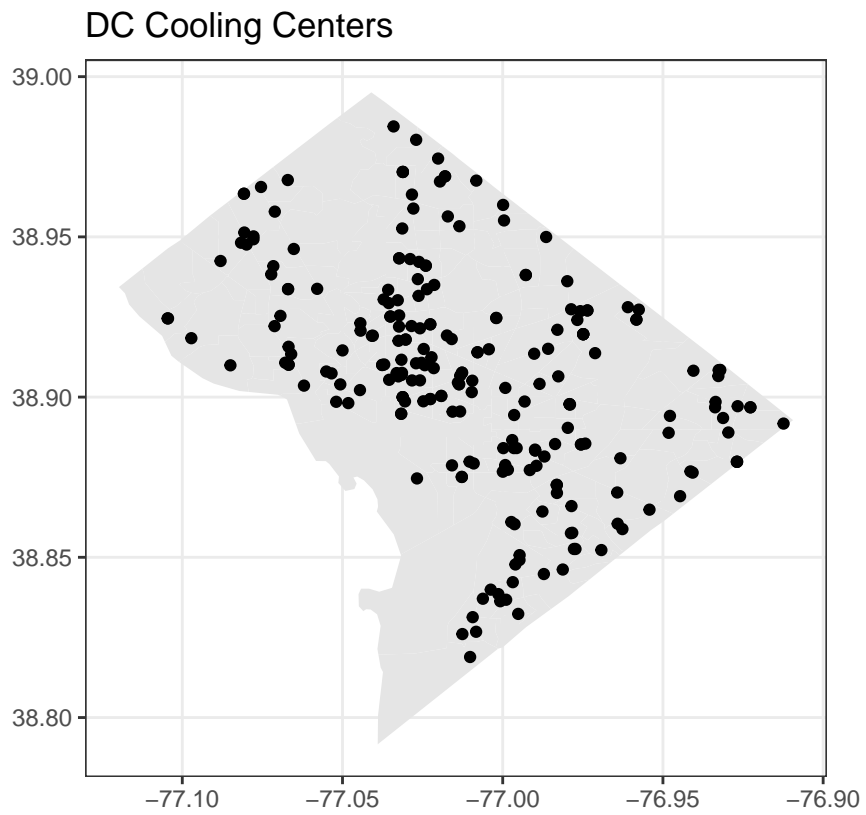
Fig. 8











## Appendix

Table 3: Correlation table of HSI variables

	P_POCP	CHILD	P_ELDERLY	POVERTY	DISABILITY	LIMENG	ASTHMA	CHD	OBSITY	HSI
P_POC	1.000	0.353	-0.032	0.761	0.681	0.165	0.857	0.748	0.928	0.929
P_CHILD	0.353	1.000	-0.062	0.322	0.123	-0.019	0.378	0.287	0.403	0.470
P_ELDERLY	-0.032	-0.062	1.000	-0.131	0.244	-0.037	-	0.396	-0.048	0.147
P_POVERTY	0.761	0.322	-0.131	1.000	0.632	0.100	0.891	0.609	0.823	0.842
P_DISABILITY	0.681	0.123	0.244	0.632	1.000	0.000	0.651	0.650	0.710	0.771
P_LIMENG	0.165	-0.019	-0.037	0.100	0.000	1.000	-	0.011	0.030	0.152
ASTHMA	0.857	0.378	-0.109	0.891	0.651	-0.020	1.000	0.686	0.935	0.903
CHD	0.748	0.287	0.396	0.609	0.650	0.011	0.686	1.000	0.781	0.844
OBSITY	0.928	0.403	-0.048	0.823	0.710	0.030	0.935	0.781	1.000	0.948
HSI	0.929	0.470	0.147	0.842	0.771	0.152	0.903	0.844	0.948	1.000

```

tree <- heat_full %>%
  filter(HEI >0) %>%
  ggplot(aes(P_POVERTY, P_TREECOVER)) +
  geom_point()

no_tree <- heat_full %>%
  filter(HEI >0) %>%
  ggplot(aes(P_POVERTY, P_NOTREE)) +
  geom_point()

temp <- tree <- heat_full %>%
  filter(HEI >0) %>%
  ggplot(aes(P_POVERTY, AIRTEMP_MEAN)) +
  geom_point()

imp_surf <- tree <- heat_full %>%
  filter(HEI >0) %>%
  ggplot(aes(P_POVERTY, P_IMPSURF)) +
  geom_point()

gridExtra::grid.arrange(tree, no_tree, temp, imp_surf, ncol = 2, right = "Fig. 8")

```

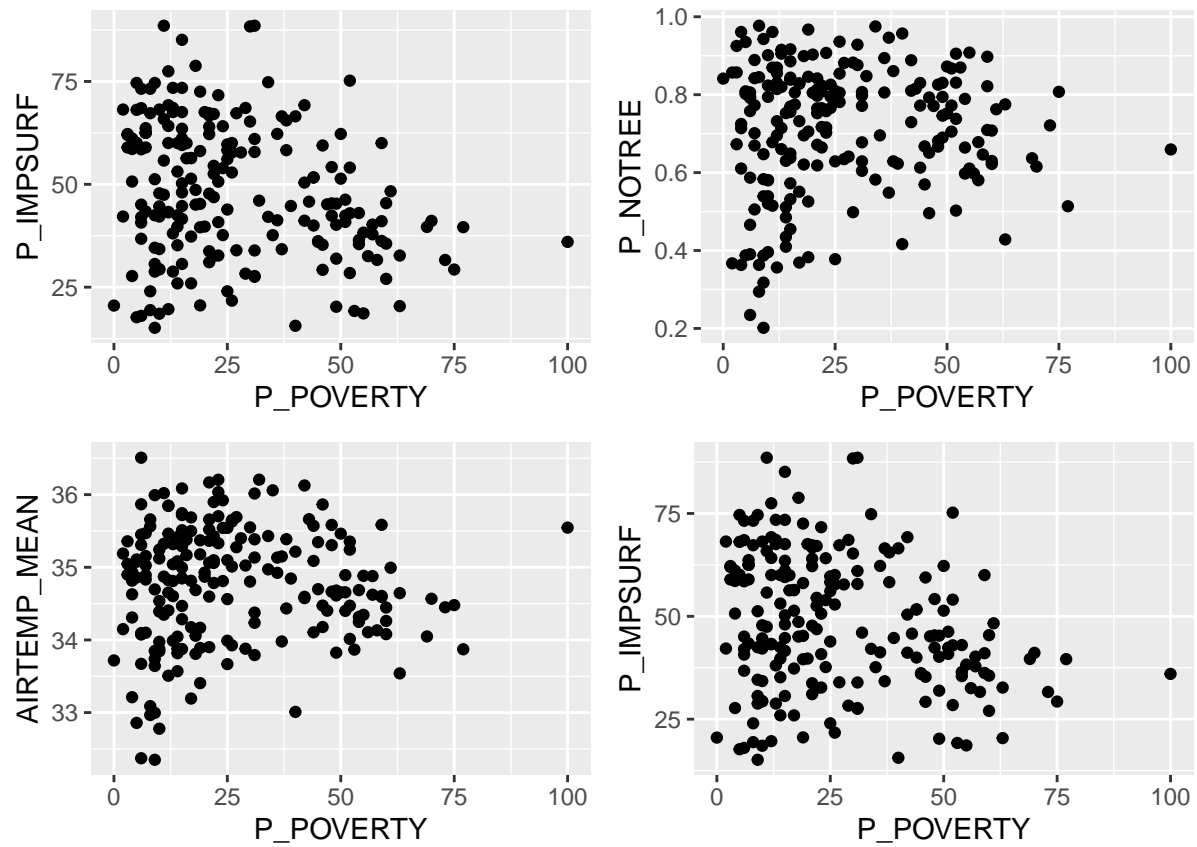


Fig. 8

## References

- Benz, Susanne Amelie, and Jennifer Anne Burney. 2021. “Widespread Race and Class Disparities in Surface Urban Heat Extremes Across the United States.” *Earth’s Future* 9 (7): e2021EF002016.
- Chakraborty, T, A Hsu, D Manya, and G Sheriff. 2020. “A Spatially Explicit Surface Urban Heat Island Database for the United States: Characterization, Uncertainties, and Possible Applications.” *ISPRS Journal of Photogrammetry and Remote Sensing* 168: 74–88.
- Hsu, Angel, Glenn Sheriff, Tirthankar Chakraborty, and Diego Manya. 2021. “Disproportionate Exposure to Urban Heat Island Intensity Across Major US Cities.” *Nature Communications* 12 (1): 1–11.