# Speak for Tree - Environment

Environment Group

2021-12-19

## Introduction

In the environment part, we try to relate the tree canopy to the pollution(PM2.5) and the surface heat(temperature).

### Data Source

The data sets we used are from the Landsat satellite, which captures images in 30x30 meters resolution. We took surface temperature and PM 2.5 into account. When we first got the data, they are both in raster formats, only contain pixels, so we used the average pixel value for each census tract. We also tried to find two parts of different areas' data and combined them together in order to make our data cover the whole Boston area. The PM 2.5 data estimates particulate concentrations for 2018 using several remote sensing sources and models of particulate matter transport. The temperature data were collected by the Landsat satellite on August 22, 2021. On this day, cloud cover was estimated to be 7.5 %, and according to the National Weather Service, the average temperature at Logan Airport was 72.5 F. ### Method

We used the model from the Demographic group and combined it with tree canopy to analyze temperature and PM2.5.

1.Generate the plots which relate the canopy to Surface Temperature & PM2.5 using QGIS;

2.Do Kriging for spatial mixed effects using R.

# PM2.5

### Using QGIS

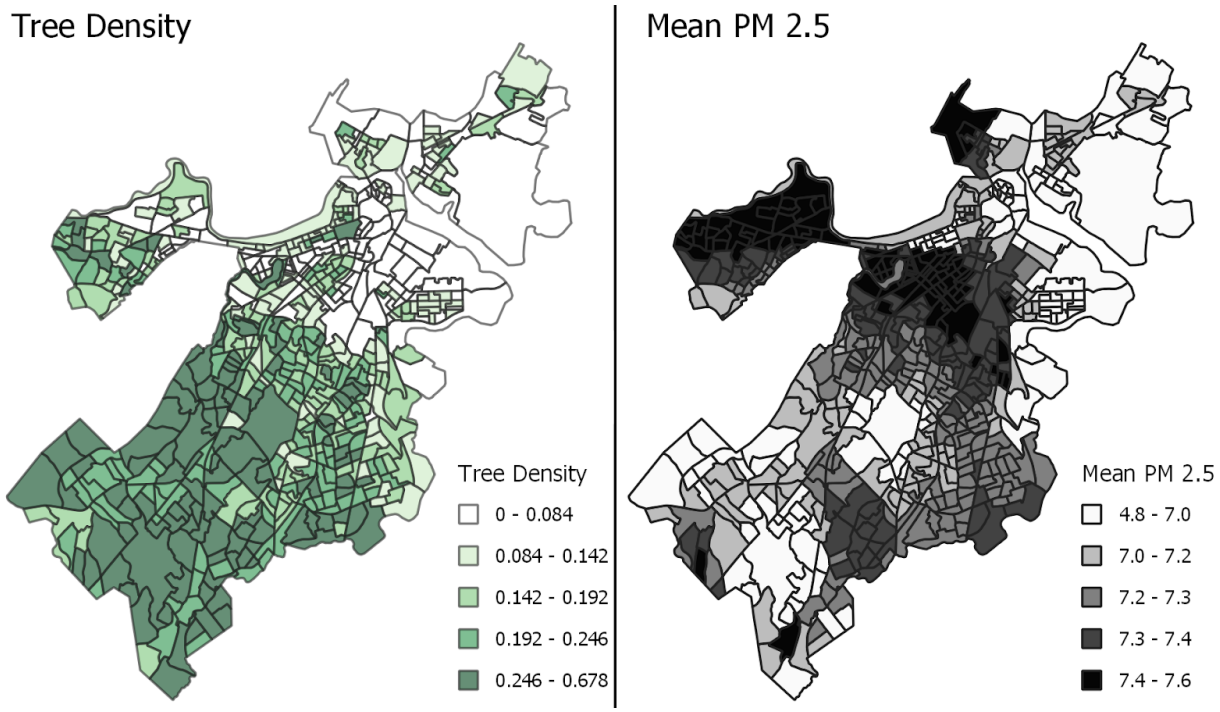



Figure 1: The PM2.5 Situation in Boston

In the north west of the map (Allston/Brighton), we can see there are few trees over there because we can see the color of the tree density is green but in the PM2.5 part, it is dark one which is unusual, we guess this happens because there are many people living there.

In the middle part (Back Bay and South End), there are not many trees so the air pollution is serious, which makes sense.

In the south west of the map (West Roxbury and Roslindale), the trees are dense and there is less air pollution.

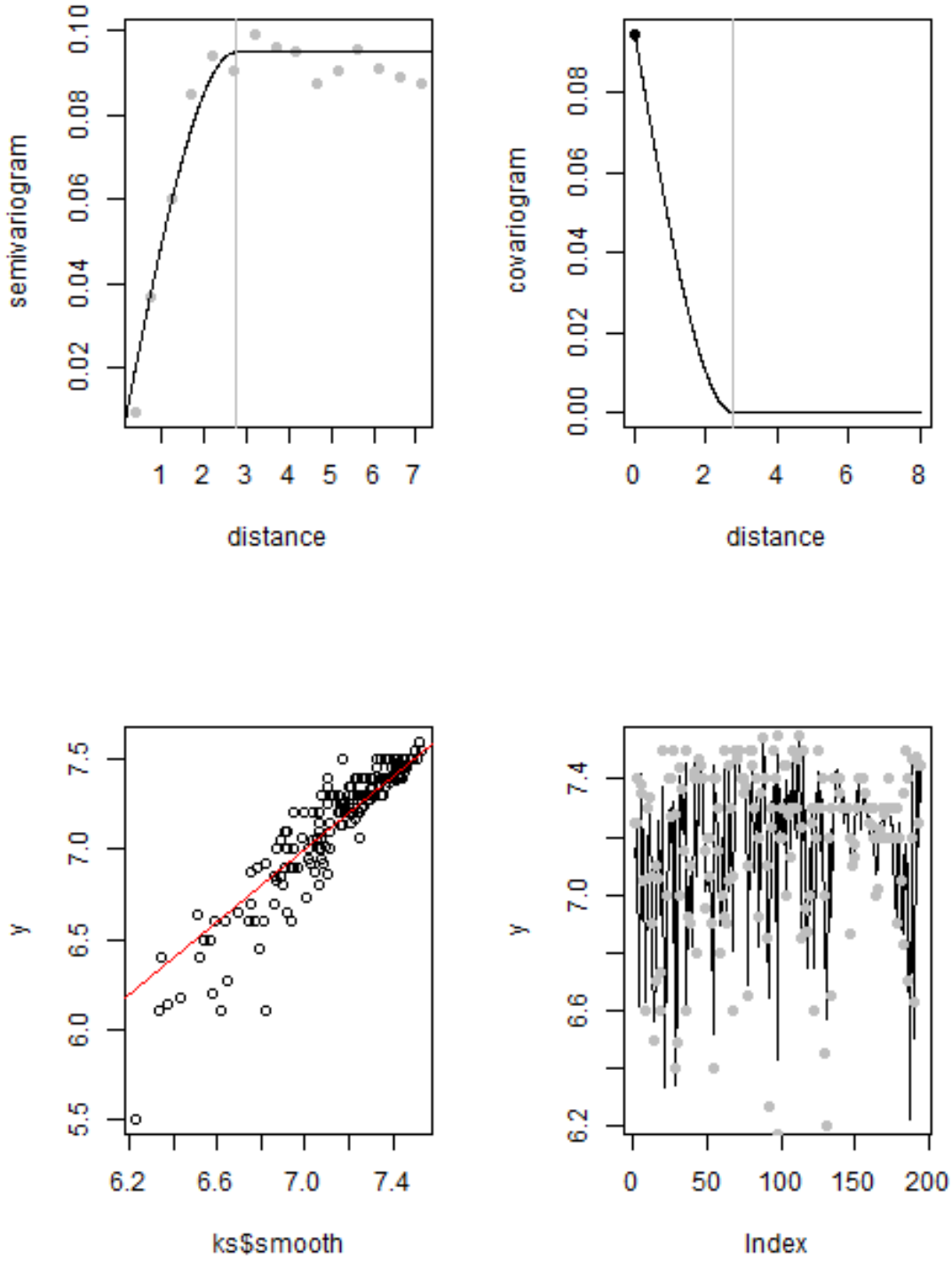
If we look closely at what the images show us, we can see two strange things.

(1) Around the middle, there is Franklin Park. Although there are a lot of trees in the park, we think the air should be flowing, so there should also be a degree of air pollution because the surroundings have a high level of PM2.5, but the PM 2.5 concentration is low here.

(2) In the north east of the map, there is the airport. Generally speaking, the air problem at the airport should be serious, but maybe because the airport is in a coastal area, the air quality is not very serious.

### PM2.5 Model

Based on the spherical variogram assumptions and krigring smooth function, predictors obtained from demographic group such as smoothed income, income residual, logarithmic form of population density, smoothed tree canopy and tree canopy residual are included in the testing variogram assumptions and obtain the smoothed PM 2.5 data. Including the residual predictors aims to avoid spatial confounding. In other words, similar variogram relationships between predictors will not be included when there is actually no variogram relationship between them.




Result shows that PM2.5 fits the spherical model well, with the stable value of variogram, which means the value of PM2.5 of one census block has an impact on the PM2.5 of census blocks nearby.

| PM2.5 | (Intercept) | Mdn_Inc_s | Mdn_Inc_r | log(PpDnsty) | tr_smth | tre_rsd |
|---|---|---|---|---|---|---|
| Coefficient | 28.060604 | -1.8453608 | 0.020124982 | 0.003232229 | 0.02885575 | -0.03057803 |
| Sd. | 1.906169 | 0.1683352 | 0.007705991 | 0.004304010 | 0.02328049 | 0.00616595 |

Regression model result shows that the coefficients of income are negatively significant which means richer places have lower levels of PM2.5. This might be due to some specific methods dealing with the environment although tree density is not high compared to areas with less citizens living. The coefficient of logarithmic population density is comparatively not that significant which means the relationship between the PM2.5 and the population density is not certain.

## Temperature

### Using QGIS

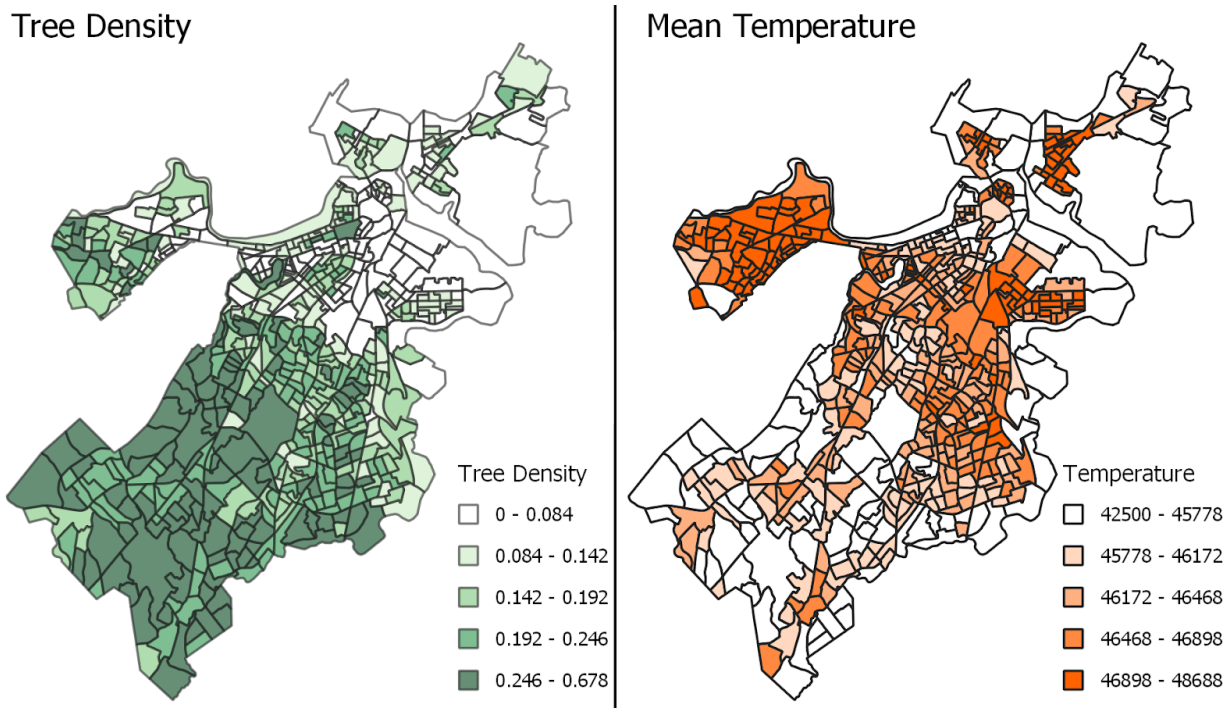
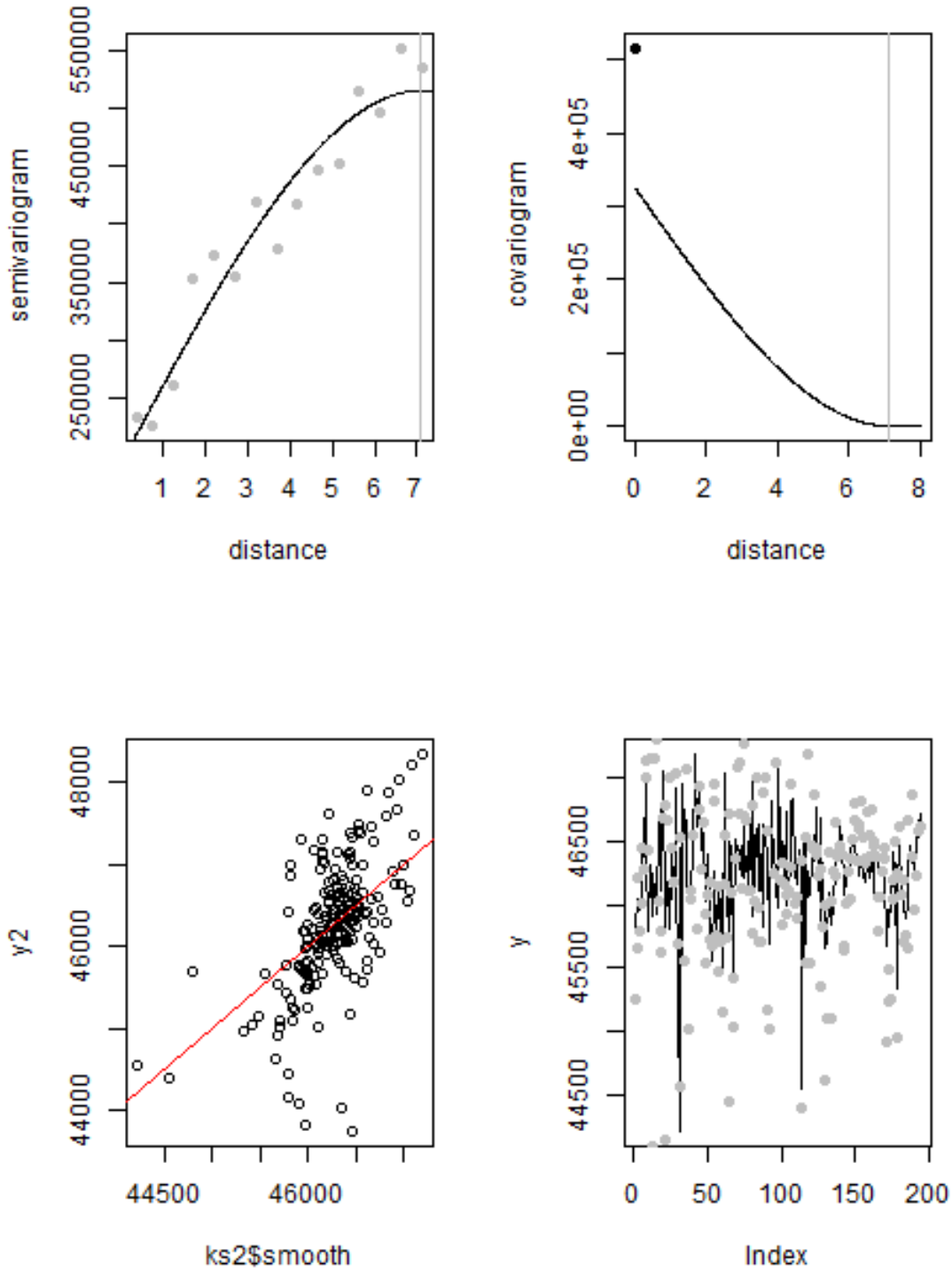



Figure 2: The Temperature Situation in Boston

Next we did the same thing with temperature data. And here the conclusion from the graph is much easier. Where trees are denser, temperatures are cooler.

**Temperature Model**




Similarly, smoothed mean temperature is obtained by the same method as mentioned before in the PM2.5 model part. However, what is different is that temperature does not fit the spherical model well because there

is no stable value of variogram from the variogram plot. This is partly because the temperature information is not enough when this project just focuses within the Boston area. If larger areas are considered, there might be a variogram relationship shown in the plots because in that case farther distance is considered and more temperature information is included. Also, the model does not fit well can be explained by the 30 meters by 30 meters resolution used in calculating the temperature and tree canopy. If we have a finer resolution, the variogram can also be improved.

| Temperature | (Intercept) | Mdn_Inc_s | Mdn_Inc_r | log(PpDnsty) | tr_smth | tre_rsd |
|---|---|---|---|---|---|---|
| Coefficient | 56610.397 | -1022.3532 | 141.27618 | 300.14727 | -2.212849 | -433.00393 |
| Sd. | 5207.884 | 470.4496 | 71.46459 | 35.27686 | 95.797697 | 65.69616 |

Regression model result shows that all the predictors coefficients are significant and we do include the main predictors related to the temperature. The smoothed income is negatively related to the temperature, which means it is hotter in richer places. The coefficient of logarithmic population density is positively correlated with the temperature and it makes sense that it is hotter in places living with more people.

## Before-and-After Smooth

Here we compared the original plots with the ones after smoothing both for temperature and PM2.5. We discovered that the smoothed image shows the gradient of color more clearly, which means for both temperature and PM2.5, the size of the data in each small area is affected to some extent by the size of the surrounding values.

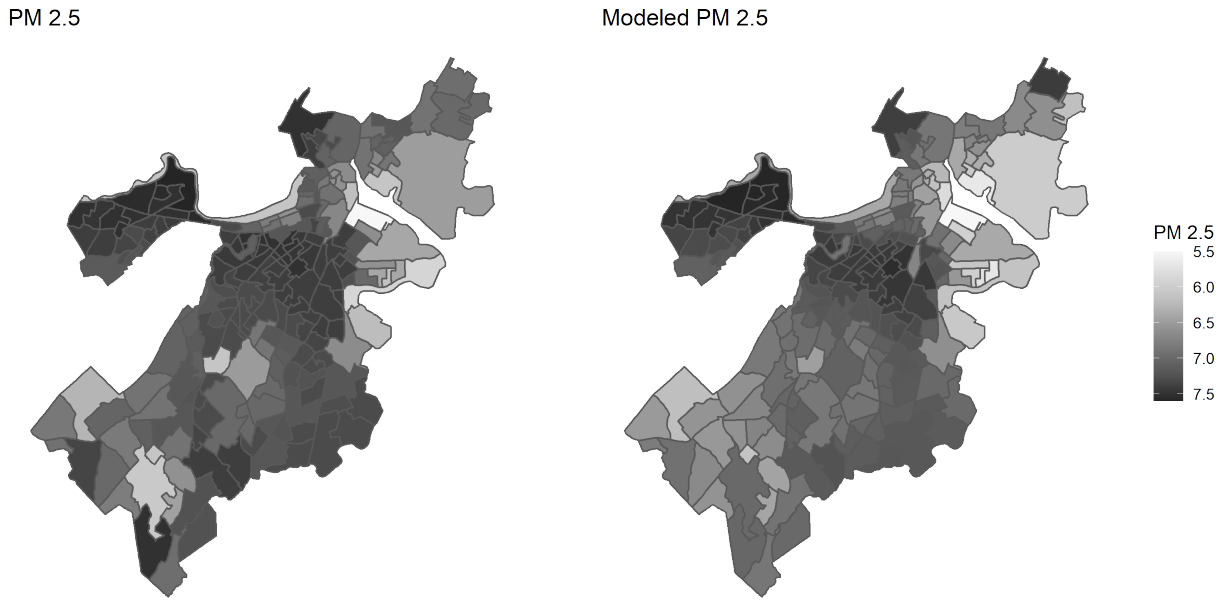



Figure 3: The PM2.5 Situation Before-and-After Smooth

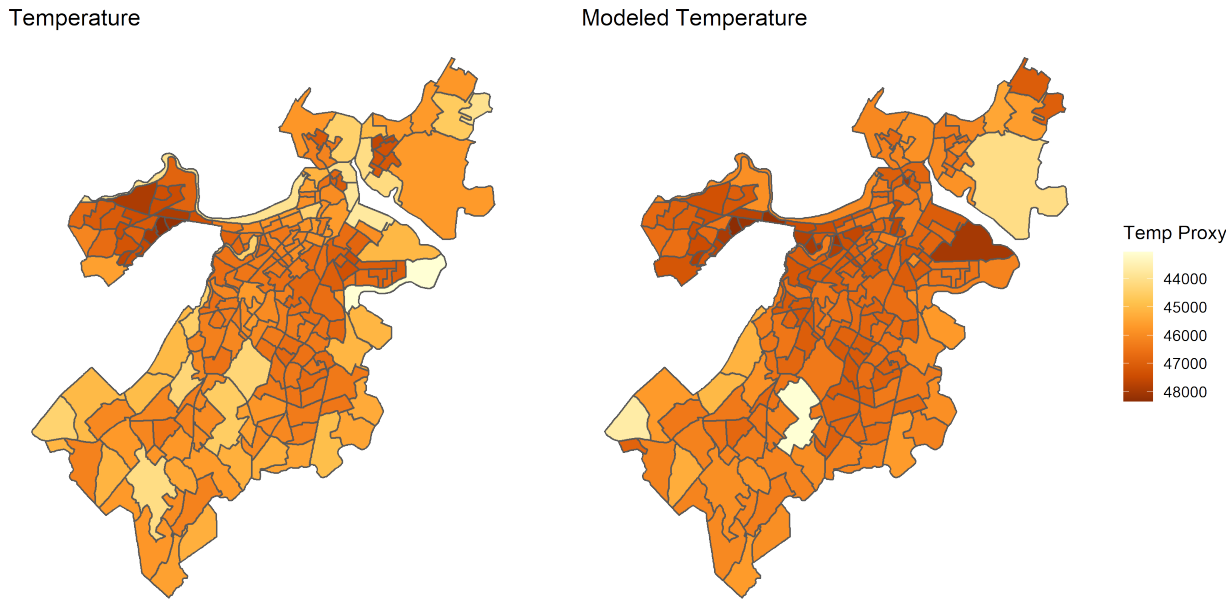



Figure 4: The Temperature Situation Before-and-After Smooth

## Summary

Both PM2.5 and temperature perform well in modeling tree canopy and predictors from demographic groups from population density to income. Smoothed PM2.5 and temperature information are obtained and variances are smaller after kriging, which can be noticed when comparing the maps.

## Reference and Source

https://data.boston.gov/dataset/census-2020-block-groups

https://www.usgs.gov/landsat-missions/landsat-provisional-surface-temperature?qt-science_support_page_related_con=0#qt-science_support_page_related_con

https://stackoverflow.com/questions/49032217/inner-joining-two-sf-objects-by-non-sf-column

https://www.youtube.com/watch?v=J-IB4_QL7Oc

http://www.goldensoftware.com/variogramTutorial.pdf

https://ggplot2-book.org/scale-colour.html