Green Space and Socioeconomic Factors: A Spatial Analysis of NYC's Urban Tree Planting

Yuxi Xiao

22nd December 2024

QMSSGR5070_001_2024_3 – GIS & Spatial Analysis-SOC SCI

This project examines the relationship between the tree density in New York City neighbourhoods and various socio-demographic factors, including population density, income levels, and ethnic composition (Asian, White, and Black populations). The goal is to identify patterns of tree distribution relative to these variables and explore inequalities in access to urban greenery. Insights from this research aim to inform urban planning and environmental justice initiatives.

Research questions:

- How does tree density vary across NYC neighborhoods with different population densities?
- Is there a relationship between income levels and tree density in an area?
- Do higher-income or less densely populated areas have more trees, and if so, to what extent?
- Which ethnic group has the highest amount living in high tree-density areas?

Purposes:

Urbanization often deprives people of their connection to nature. Past and recent studies have consistently shared a consensus that this lack of access to green spaces can deteriorate people's stress levels, mental health, and even physical health (Vlahov and Galea, 2002; Jiang, et al., 2021; Ventriglio, et al., 2020). For example, exposure to greenery has been linked to reduced symptoms of depression, improved cardiovascular health, and better community well-being (Gianfredi, et al., 2021). However, as cities grow denser, urban residents often find themselves disconnected from such benefits. Therefore, understanding the distribution of greenery in urban areas is critical to addressing health disparities and promoting well-being across different populations. Apart from improving health, street trees have a series of benefits for other urban organisms, such as food and shelter (Comes et al., 2023; Shackleton, 2016). They also help maintain healthier habitats for many species, such as birds, squirrels, and mistletoes (de Menezes et al., 2023; Wood and Esaian, 2020).

Despite the importance of urban greenery to residents' health and living conditions, ample studies that investigated different parts of the world all suggest the phenomenon that socioeconomic inequalities have led to an uneven distribution of urban green space, which can exacerbate social vulnerability and contribute to environmental degradation in cities as a whole (Lin, et al., 2021; Pistón, et al., 2022). New York City (NYC), as one of the most unequal cities in America—with the top 20% of earners making 53 times more than the bottom 20%—is worth investigating the correlation between this significant income inequality and the distribution of tree resources. The city is also very culturally diverse, home to a population that speaks over 200 languages, making it an ideal case study to examine how socioeconomic and demographic factors, including income, population density, and ethnicities, influence access to urban greenery. By exploring these patterns, this study aims to highlight potential inequities in urban tree distribution and contribute to the existing literature by providing actionable insights for addressing them.

Overall, urban greenery is an important part of urban planning to restore some of this balance, reduce social disparities, and secure a sustainable future for cities. It is crucial to formulate policies to improve equal tree planting while knowing which areas or social groups are most in need of change and support so that the policies are sufficient.

Hypotheses:

Hypothesis 1:

Tree density (tree per tract area) is negatively correlated with population density, such that more densely populated neighborhoods have lower tree density. This prediction comes from the substituting idea that more residents in an area due to the sacrifice of tree space, which individuals are less proximate to trees.

Hypothesis 2:

Neighborhoods with higher income levels have significantly higher tree density compared to lower-income neighborhoods.

Hypothesis 3:

Neighborhoods with a higher proportion of White residents have greater tree density compared to neighborhoods with a higher proportion of Black or Asian residents.

Data and Methodology

I will be conducting tree data from the 2015 NYC OpenData Street Tree Census dataset. This data provides a detailed point shapefile containing the recorded trees across NYC in 2015, which includes various tree factors such as their exact geographic locations, the diameter of the tree, and the status of the tree. This data was collected by volunteers and staff under the organization of NYC Parks & Recreation and partner organizations. It took over a year to complete, from May 2015 to October 2016, and the results show that there are 666,134 trees planted along NYC's streets. It is the third decadal street tree census and the largest citizen science initiative in NYC Parks' history, which strengthens the importance of this dataset and reinforces the reason for choosing it. This data source is appropriate for my research because it is comprehensive and spatially detailed. By downloading the data directly from NYC OpenData, I ensure it is sourced from a more reliable government platform. Hence, when combined with other demographic and socioeconomic data, it will enable accurate and robust interpretations that address the research questions and hypotheses.

The “status” column indicates whether the tree documented is alive, stump, or dead. I will exclude rows containing either dead or stump, leaving the data with only alive trees. This step, though it reduces the data size, makes sure that the trees I use are more specifically defined as those still alive and can contribute to the actual greenery. This is important because only living trees are able to produce the scenery, fresher air, and represent better environmental management, while dead trees and stumps may lead to the opposite results and contradicted conclusions. Therefore, the first map in part 4 clearly demonstrates areas with living trees and areas with dead trees or stumps, with purple representing the former and green representing other tree statuses.

The New York City census tracts will be derived from the TIGER shapefile provided by the US Census Bureau. This shapefile includes neighborhood boundaries and every area in the city. Census tracts are an ideal unit of analysis because they align with both demographic and tree data.

The neighborhood-level demographic and socioeconomic data are also collected from the US Census Bureau, including total population, median household income levels, and ethnic populations (Asian, White, Black). These are the independent variables and they would be used to examine the tree density (dependent variable). For these variables, I decided to use the data in 2020 and it is appropriate since it represents the most recent and complete demographic information available at the neighborhood level. Moreover, I cleaned the data for these demographic and socioeconomic factors by changing each column’s NULL cells into zero, with others remaining constant. This ensures that we will not leave some of the areas blank. However, the tree numbers and planted locations might have changed during the 5-year gap between 2015 and 2020, potentially introducing bias into the analysis.

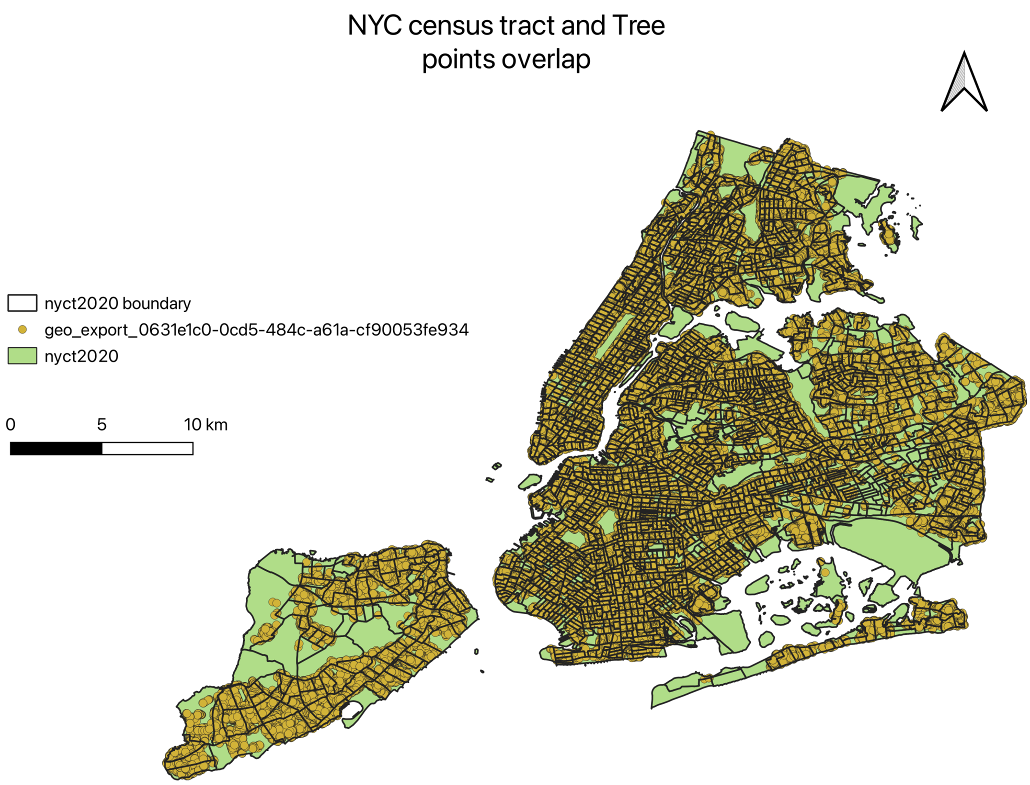
After obtaining all three fundamental datasets (the first one is a point shapefile, the second one is a census shapefile, and the last one is a table CSV file), I first performed “join attributes by location” to merge the tree points shapefile with the NYC census shapefile, thereby generating a new spatial map that is similar to the city shapefile but with integrated data. Then, I was able to distinguish trees by their status (alive, stump, and dead) by creating a new variable that marked “alive” rows as purple (as 1) and marked “stump” and “dead” rows as green (as 0). Hence, I could visualize the spatial distribution of tree status across NYC neighborhoods, which provided an initial understanding of the concentration of “useful” trees versus those needing removal or replacement. Second, I joined the newly formed spatial layer with the demographic data in its layer properties matched by their GEOID.

To find the tree density, I used “count points in polygon” to count the number of trees in each tract, then divided these counts by tract areas and multiplied by 100. Similarly, to find the ethnic composition, I divided each ethnic group by the total population in each tract and times by 100, which gave me the percentage of each ethnicity in each neighborhood.

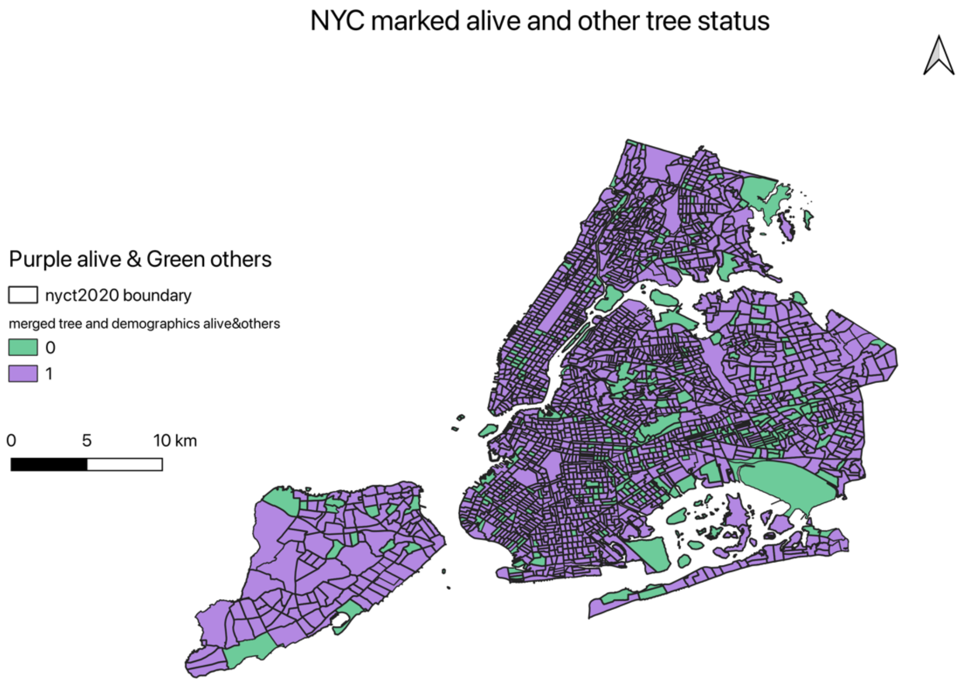
Also, I used GeoDa to create several maps and tables necessary for analysis, including spatial autocorrelation and spatial regressions.

Results and Analysis

This first map shows an overlay of the NYC spatial map and the point locations of trees. The map indicates that the points are very densely located in most NYC tracts, while a few areas in the southeast and on Staten Island are absent of trees. Nonetheless, the dense tracts and trees make it difficult to see the actual patterns between the trees and population and other demographic factors. Therefore, I formed Map 2 that merges points into the spatial map. The green highlighted tracts indicate an intersection with either stumps or dead trees, while purple tracts have living trees. Although green tracts account for a minority of lands, it is important to consider that even if later areas are found with high tree density, the green mark suggests that they do not necessarily signify well-maintained environments or a high-quality living environment for residents. Instead, high tree density in these areas may reflect the need for maintenance or renewal, as the trees are either dead or have been reduced to stumps, indicating environmental neglect rather than sustainable greenery.

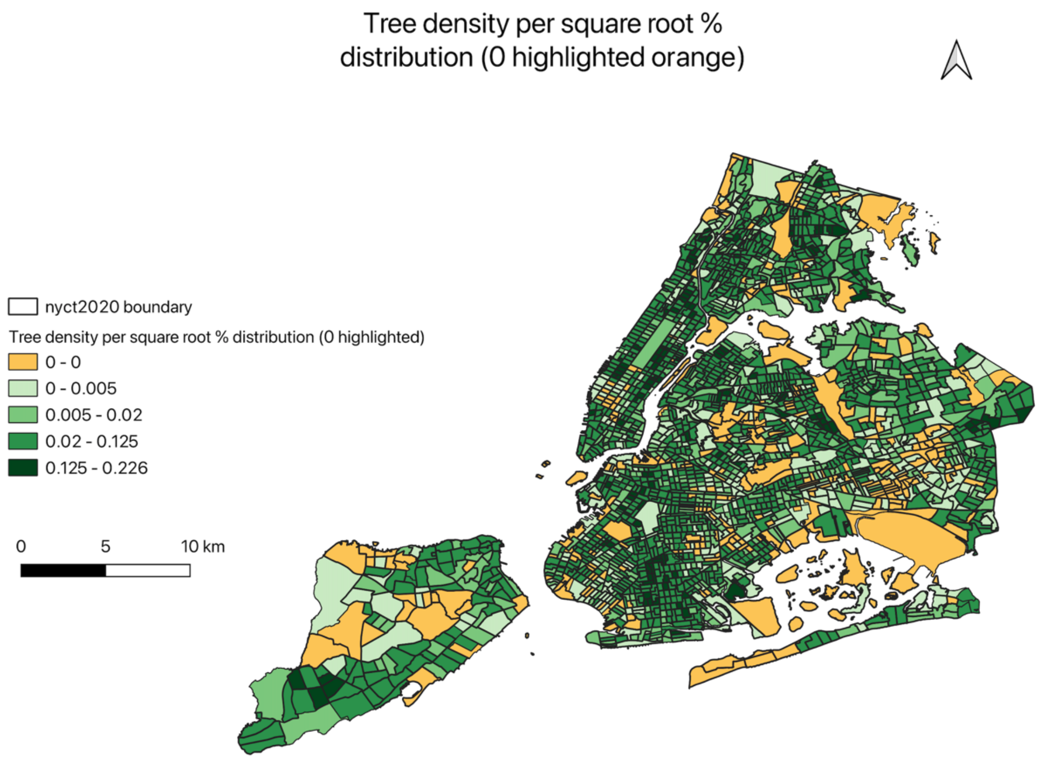


Map 1



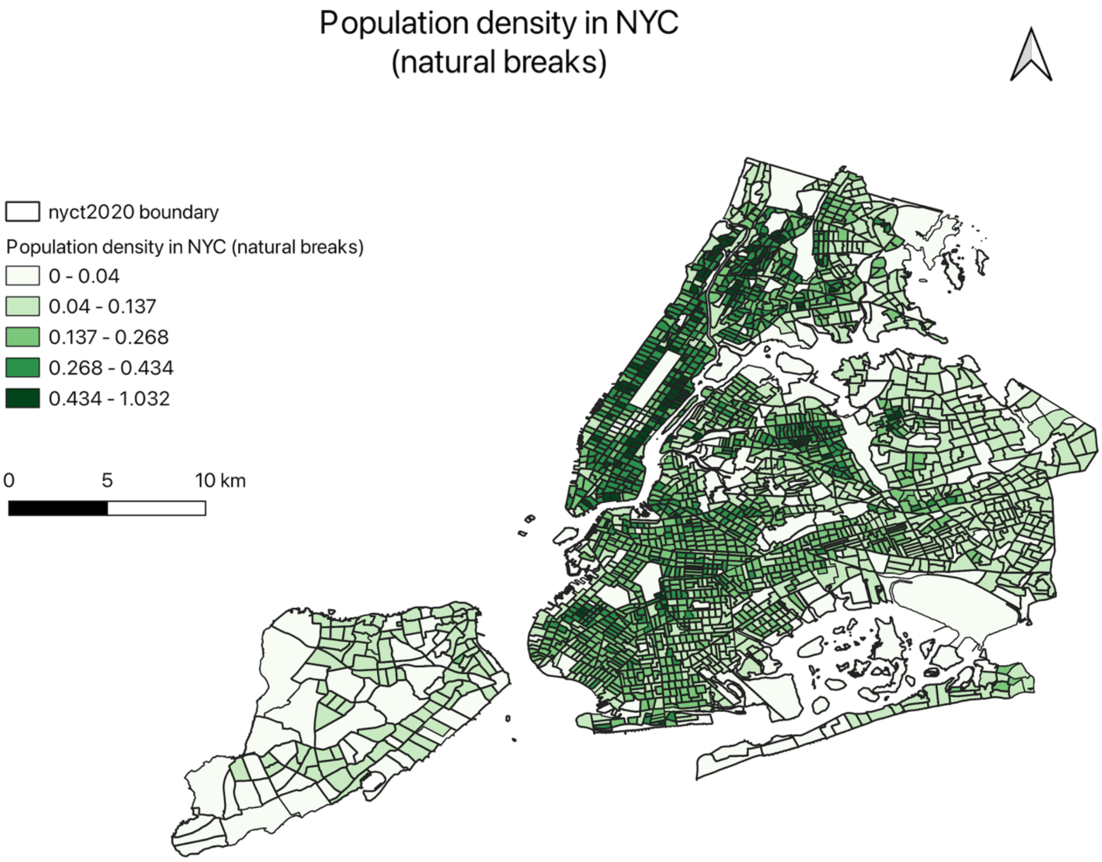
Map 2

Map 3 visualizes the tree density distribution, again using the natural breaks method. The areas with very small tree density that can be rounded up to 0% are highlighted in orange. This orange highlight is spread across NYC, particularly visible in parts of Staten Island, eastern Queens, and some tracts in the Bronx, which indicate regions with no significantly recorded tree. The only pattern found here is that fewer tracts in Manhattan are marked orange (zero percentage). The darker green areas represent higher tree density and vice versa. Higher tree density is concentrated in parts of each borough and is also spread across the city without a strongly distinguishable pattern that shows which borough has a generally higher tree density concentration It does emphasize disparities in tree distribution as certain tracts completely lack tree coverage while others maintain significant density.



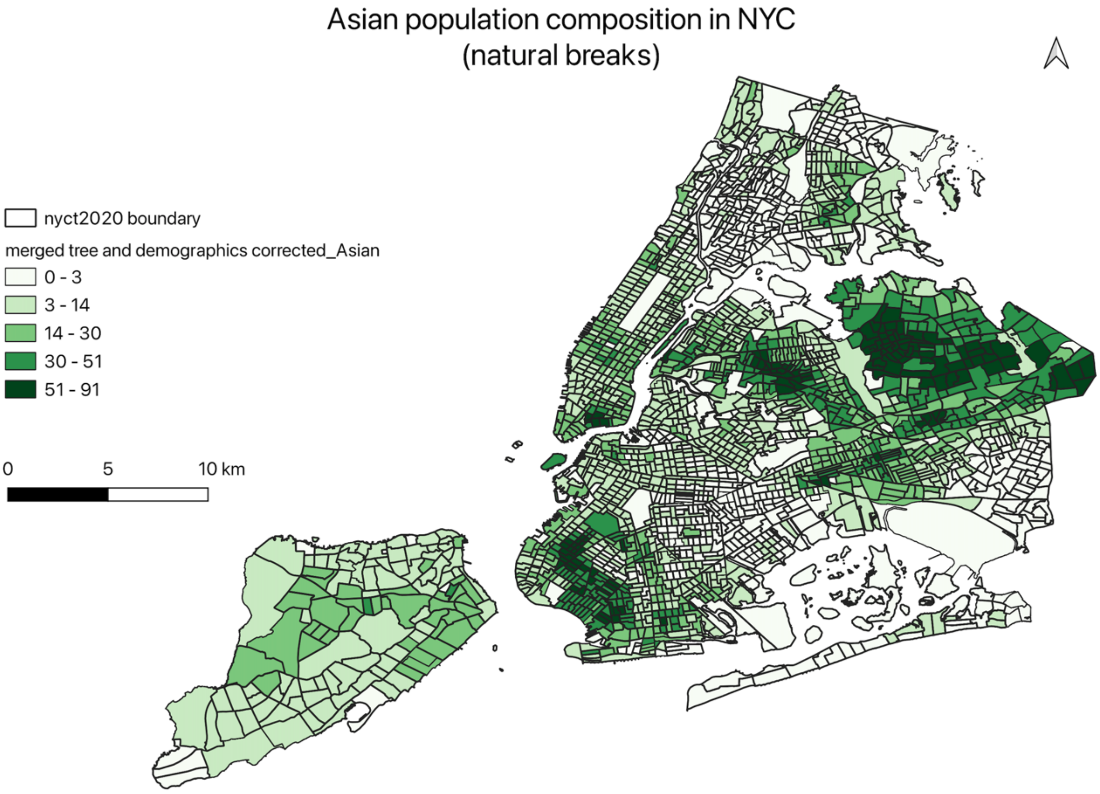
Map 3

The four choropleth maps below all represent the distribution of different populations. Map 4 shows the distribution of the population density in NYC. The total population density data shows a very opposite relationship to the tree concentration, especially in Manhattan and Brooklyn, which have a very dense population. Although this negative correlation is also shown in the east part of Richmond, the west part of the city where Staten Island is located has a quite similar pattern for tree density and population density choropleth maps.



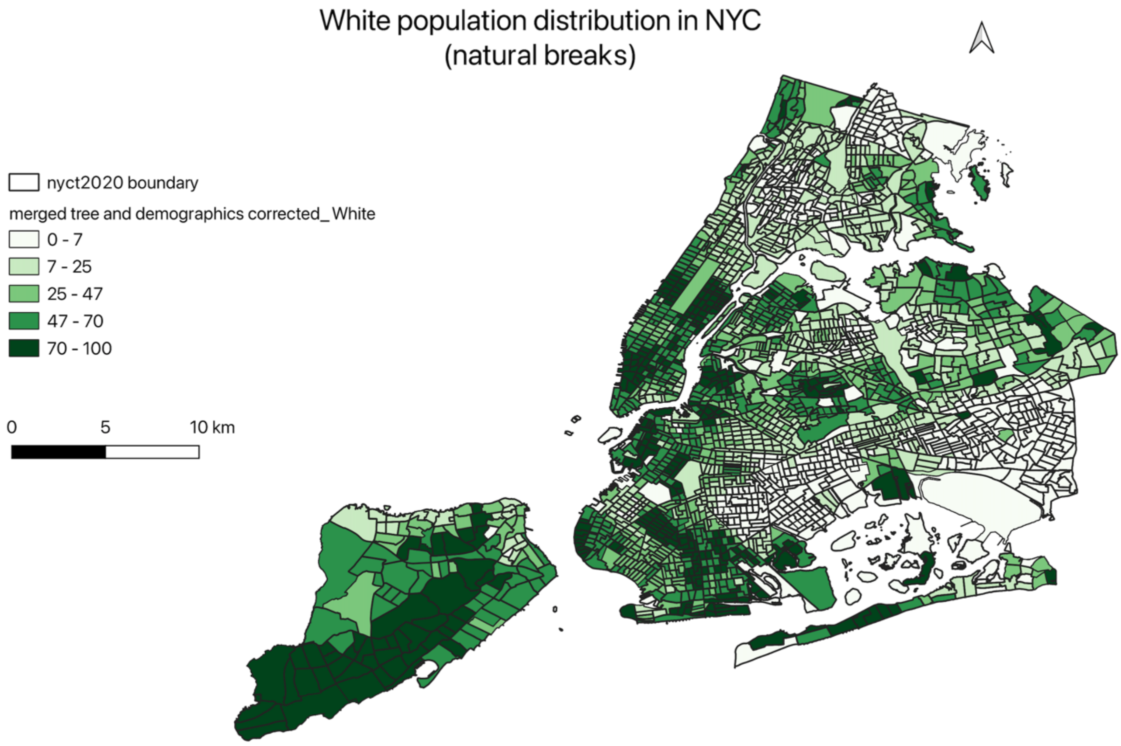
Map 4

Map 5 illustrates the distribution of the Asian population percentage across NYC using a natural breaks classification. The darker green areas represent higher proportions of the Asian population, while lighter green areas represent the opposite. Therefore, we can see that neighborhoods such as Flushing in Queens and parts of Brooklyn (e.g., Sunset Park and Bensonhurst) have dark green clusters, and from hindsight, we know that these areas are known for having significant Asian communities. Some clusters of high concentration also appear in midtown and downtown Manhattan, which has Chinatown and Ktown. It is noticeably lighter in most of Staten Island, the Bronx, and some areas of Queens and Brooklyn.



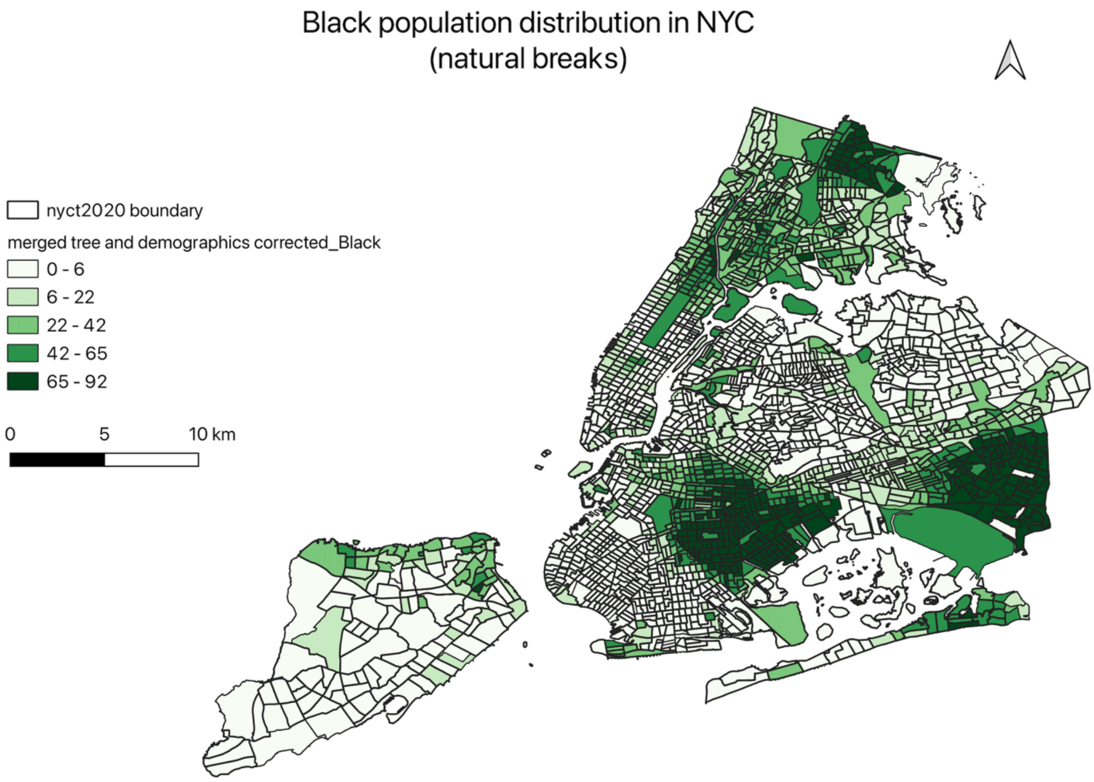
Map 5

Maps 6 and 7 all used the natural breaks method to show the distribution of their specific ethnic populations. Map 6 shows the distribution of the White population composition. It suggests that most White people live on the city’s west side in boroughs such as Staten Island and midtown to downtown Manhattan. Whereas many areas in the Bronx, Brooklyn, and Queens have only 0 to 7 percent of the White population.



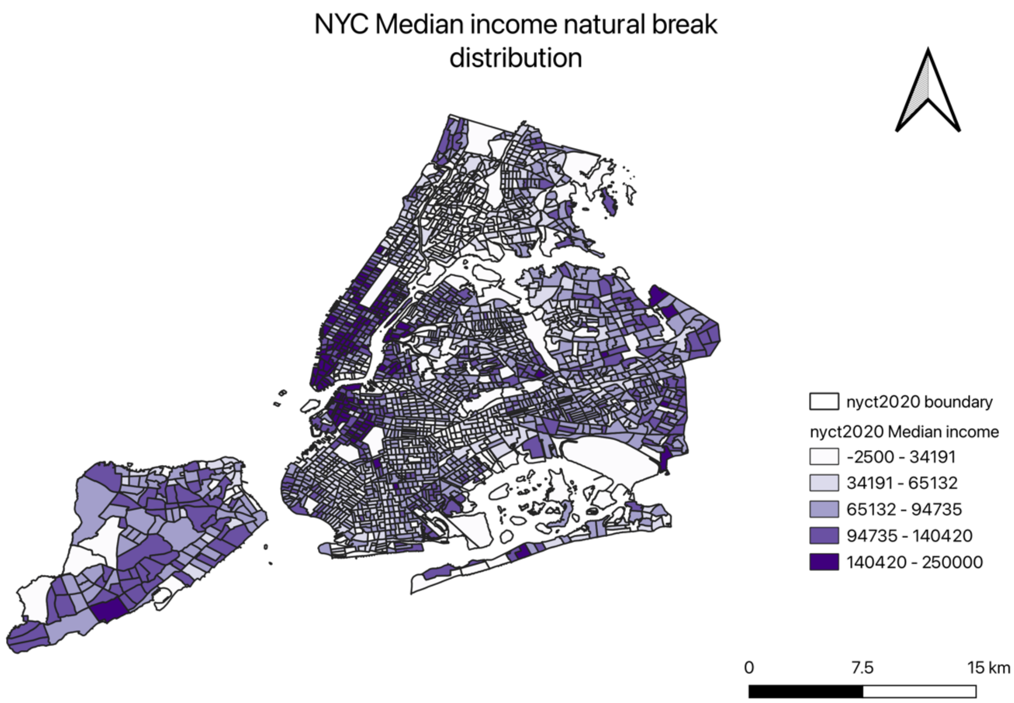
Map 6

Map 7 illustrates the Black population composition in NYC. It has a high concentration in mainly three parts of the city—Central Brooklyn, Southeast Queens, and South Bronx, where most neighborhoods have a 65% to 92% Black composition. Moreover, the Bronx has an overall higher concentration of this ethnicity compared to other ethnic groups.



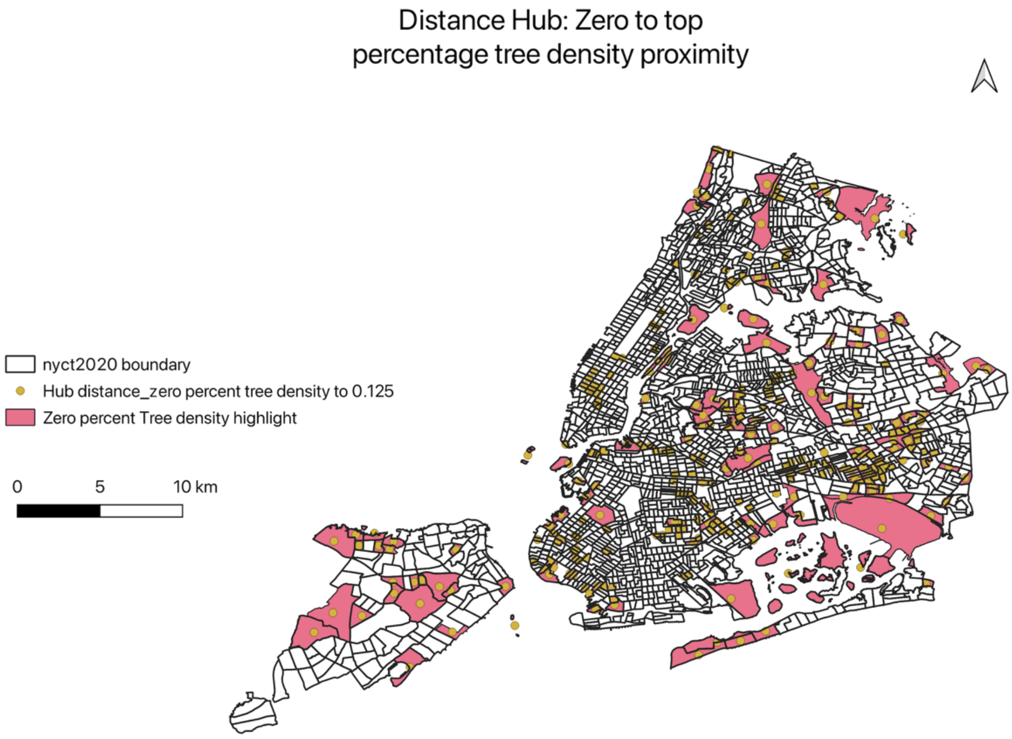
Map 7

Map 8 shows the households' median income distribution in NYC using the natural breaks method. Midtown and downtown Manhattan have the most tracts in the highest median income class. Some parts of Staten Island, Brooklyn, and Queens also have higher income concentrations. In contrast, low median-income tracts are found dominantly in the Bronx, central Brooklyn, and parts of northern Manhattan. Many of these tracts may even have negative household income, which means that residents in these areas are likely to live on social assistance programs or rely heavily on public resources. This map clearly highlights the existence of stark wealth inequality, which might be a factor when examining access to urban greenery and environmental justice.



Map 8

After showing the choropleth map of tree density, I created another three layers representing the zero percent areas, above half tree density areas (0.113), and above 0.125 areas. Then, I formed a hub distance points layer (Map 9) that indicates the proximity between those zero percent areas to the highest (above 0.125 percent) tree density areas. Most of the pink areas (zero percent) intersect with the hub distance points, which suggests while these areas lack trees themselves relative to their land size, they may still be near high-density tree areas, making green spaces somewhat accessible and hence overcome some of the negativity of having limited tree density in these neighborhoods.



Map 9

Spatial Autocorrelation and Spatial Regressions (Lag and Errors) with LISA Map:
Using GeoDa, I constructed the OLS regression model, spatial lag model, and spatial error model. I chose to use the tree density as the dependent variable. The independent variables are the total population density, the three ethnic population compositions (Asian, Black, and White), and household median income.

The Ordinary Least Squares (OLS) model has an R-squared value of around 0.081, indicating that these independent variables account for 8.1% of the changes in tree density. Demograp_2 represents the total population. However, all three ethnic populations produce a negative coefficient that suggests their negative correlation with tree density. This might mean that these ethnic groups tend to live in dense neighborhoods with relatively less tree coverage. Hence, the positive relationship between tree density and Demograp_2 may be caused by other ethnic groups. Nonetheless, the White population clearly has a much lower negative relationship with tree density, while the Black population has the worst correlation. The Median income level's coefficient is positive, showing a positive relationship with the tree density. That is, the higher the median income of a census tract, the higher its tree coverage (i.e., better tree planting and preservation), and vice versa. This highlights the likelihood of income inequality affecting access to urban greenery. These findings suggest that urban planning and environmental policies should consider more socioeconomic disparities to ensure equitable distribution of tree coverage across all communities.

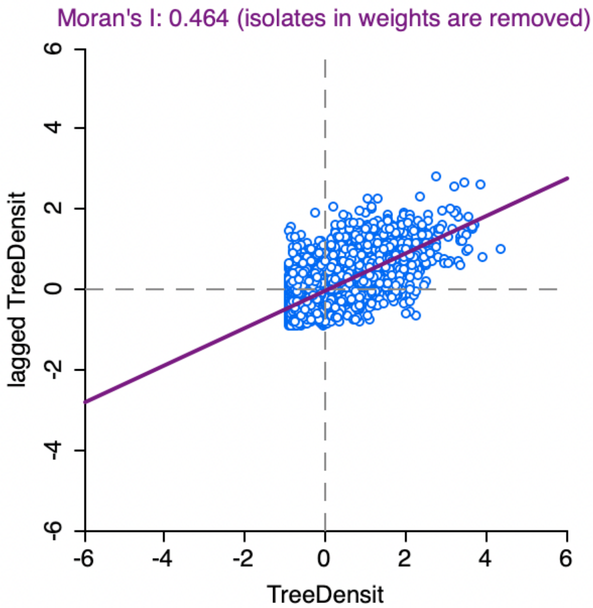
Compared to the OLS model, the R-squared value gets massively improved in the SLM and the SEM (0.467 and 0.477). The positive and highly significant lag coefficient (Rho) at 0.7248 indicates strong spatial autocorrelation in the dependent variable in which areas with high or low tree density are influenced by the tree density of their neighbors. The weighted tree density variable has a coefficient of 0.7248, again showing that tree density in neighboring areas strongly influences tree density in the focal area. The Demograp_2 (total population)'s coefficient is positive and significant. It suggests that a higher population correlates with higher tree density, which supports the result of the OLS model. Demograp_3, 4, and 5 show negative coefficients that also correspond to the OLS model and are all statistically significant with high absolute z values above 1.96. The median income variable has a positive relationship with tree density, which also confirms the OLS result.

SEM obtained similar outcomes but had a less statistically significant result for the coefficient of Demograp_3 (White population) with a z value of only -1.792. This might indicate its weaker relationship to tree density than in the SLM.

```
REGRESSION
----------
SUMMARY OF OUTPUT: ORDINARY LEAST SQUARES ESTIMATION
Data set            :  TREE DENSITY %
Dependent Variable  :  TreeDensit  Number of Observations: 2325
Mean dependent var  :   0.0391686  Number of Variables   :    6
S.D. dependent var  :   0.0429134  Degrees of Freedom    : 2319

R-squared           :    0.080929  F-statistic           :     40.8403
Adjusted R-squared  :    0.078948  Prob(F-statistic)     : 2.21282e-40
Sum squared residual:     3.93512  Log likelihood        :      4119.5
Sigma-square        :   0.0016969  Akaike info criterion :       -8227
S.E. of regression  :   0.0411935  Schwarz criterion     :    -8192.49
Sigma-square ML     :  0.00169253
S.E of regression ML:   0.0411403

-----------------------------------------------------------------------------
       Variable      Coefficient      Std.Error    t-Statistic   Probability
-----------------------------------------------------------------------------
          CONSTANT     0.0175349     0.00244723        7.16521      0.00000
        Demograp_2   6.96036e-06    8.74055e-07         7.9633      0.00000
        Demograp_3  -2.71078e-06    1.19775e-06       -2.26323      0.02371
        Demograp_4  -9.25744e-06    1.39286e-06       -6.64636      0.00000
        Demograp_5  -7.60829e-06    1.47999e-06       -5.14078      0.00000
         Medincome    1.5601e-07    2.58281e-08        6.04031      0.00000
-----------------------------------------------------------------------------
```

```
REGRESSION
----------
SUMMARY OF OUTPUT: SPATIAL LAG MODEL - MAXIMUM LIKELIHOOD ESTIMATION
Data set            : TREE DENSITY %
Spatial Weight      : TREE DENSITY % weight
Dependent Variable  :  TreeDensit  Number of Observations: 2325
Mean dependent var  :   0.0391686  Number of Variables   :    7
S.D. dependent var  :   0.0429134  Degrees of Freedom    : 2318
Lag coeff.   (Rho)  :    0.724833

R-squared           :    0.466834  Log likelihood        :     4614.36
Sq. Correlation     : -            Akaike info criterion :    -9214.71
Sigma-square        : 0.000981858  Schwarz criterion     :    -9174.45
S.E of regression   :   0.0313346

-----------------------------------------------------------------------------
       Variable       Coefficient     Std.Error       z-value    Probability
-----------------------------------------------------------------------------
      W_TreeDensit       0.724833     0.0173199       41.8496      0.00000
          CONSTANT    -0.00399312    0.00190804      -2.09279      0.03637
        Demograp_2     4.5221e-06   6.74063e-07       6.70873      0.00000
        Demograp_3   -2.60065e-06    9.1186e-07      -2.85203      0.00434
        Demograp_4    -4.9114e-06   1.07744e-06      -4.55839      0.00001
        Demograp_5    -4.0482e-06   1.13458e-06      -3.56802      0.00036
         Medincome    1.15807e-07    1.9796e-08       5.85003      0.00000
-----------------------------------------------------------------------------
```

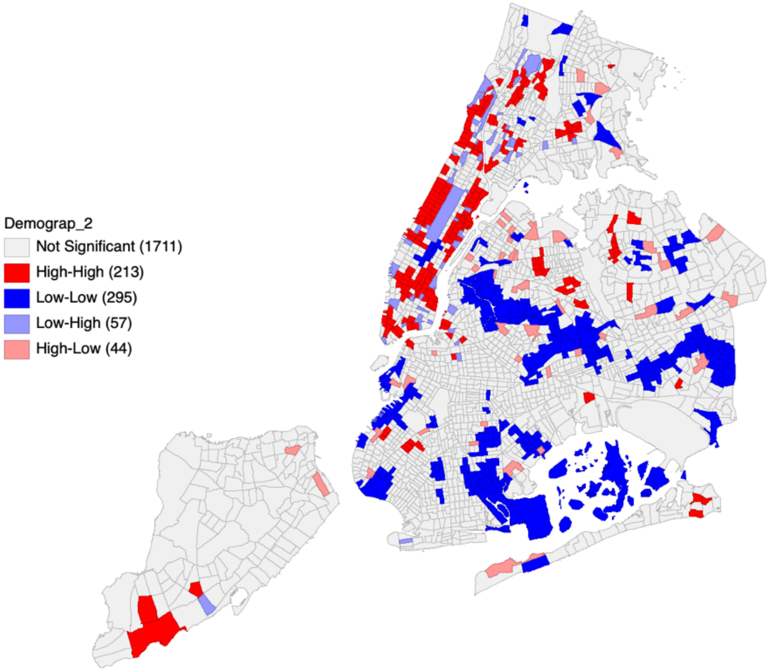
```
REGRESSION
----------
SUMMARY OF OUTPUT: SPATIAL ERROR MODEL - MAXIMUM LIKELIHOOD ESTIMATION
Data set            : TREE DENSITY %
Spatial Weight      : TREE DENSITY % weight
Dependent Variable  :  TreeDensit  Number of Observations: 2325
Mean dependent var  :    0.039169  Number of Variables   :    6
S.D. dependent var  :    0.042913  Degrees of Freedom    : 2319
Lag coeff. (Lambda) :    0.741603

R-squared           :    0.477057  R-squared (BUSE)      : -
Sq. Correlation     : -            Log likelihood        : 4628.278361
Sigma-square        : 0.000963031  Akaike info criterion :    -9244.56
S.E of regression   :   0.0310327  Schwarz criterion     :    -9210.05

-----------------------------------------------------------------------------
       Variable       Coefficient      Std.Error       z-value    Probability
-----------------------------------------------------------------------------
          CONSTANT     0.0215405     0.00314604        6.84686      0.00000
        Demograp_2   6.30983e-06    1.07204e-06         5.8858      0.00000
        Demograp_3  -2.63245e-06    1.46885e-06       -1.79218      0.07310
        Demograp_4  -1.02455e-05    1.84298e-06       -5.55919      0.00000
        Demograp_5    -6.171e-06    1.91799e-06       -3.21743      0.00129
         Medincome   1.53225e-07    2.36474e-08        6.47957      0.00000
            LAMBDA      0.741603      0.0173598        42.7195      0.00000
-----------------------------------------------------------------------------
```

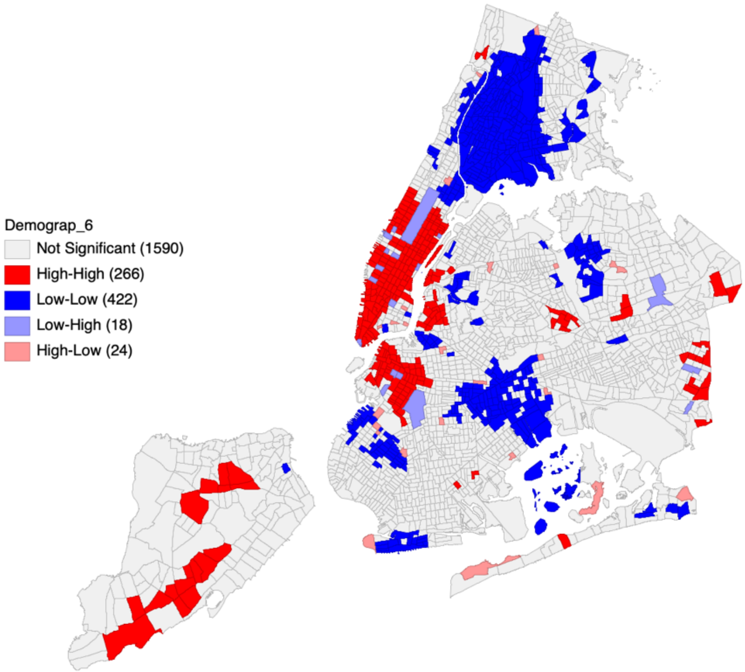


This scatterplot with Moran's I value of 0.464 suggests that similar values of tree density tend to cluster together spatially. That is, areas with a high number of tree densities are near other areas with similarly high percentages, and vice versa. This value is in the range of moderate to strong positive autocorrelation, meaning that the clustering pattern is quite pronounced. The distribution of points is not random, further indicating the existence of these significant spatial patterns.

Below are all the LISA cluster maps that categorize tracts into High-High, Low-Low, High-Low, and Low-High clusters based on the spatial weights. The first map that looks at the total population in NYC has High-High clusters (red) concentrated in areas such as Manhattan and Staten Island. They indicate neighbourhoods with both high tree density weights and high total population. From geographical knowledge, we know that the red tracts are mainly centred around Central Park in Manhattan, though they are also noticeably present in both downtown and uptown of Manhattan compared to other boroughs. These results are likely reflecting urban areas with greater investment in tree planting and maintenance, potentially wealthier neighbourhoods. The blue neighbourhoods are the Low-Low clusters, which are predominantly found in areas of the Bronx, Brooklyn, and parts of Queens. These neighbourhoods show both low tree plant weights and low population density. This might indicate potential underinvestment in greenery in less populated regions.

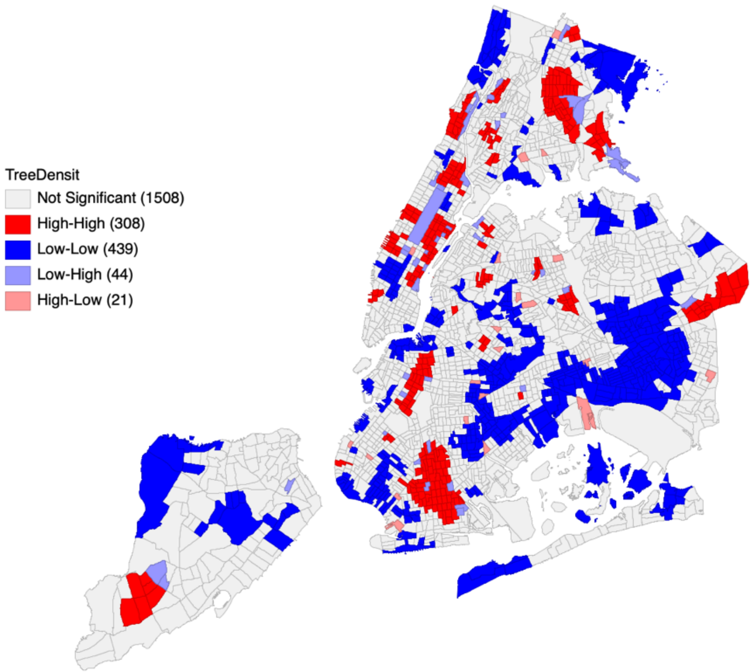
The second map with demograp_6 represents the relationship between tree density and household median income level. Similarly, the high-high regions lie in mid- and downtown Manhattan, indicating they contain higher median income and a high tree density. In contrast, neighborhoods in uptown Manhattan are showing a low-low cluster. This means that these “blue” neighborhoods generally have lower-income households and a low tree density. These results might reveal the likelihood of urban environmental inequality that is worth investigating. Overall, in this map, Manhattan is the borough with the most significant clustering, while other boroughs contain mainly insignificant regions.





The dependent variable (tree density)'s local LISA cluster map below shows areas of spatial autocorrelation in tree density across the boroughs of New York City. The High-High clusters are concentrated in specific regions across the city that are likely to represent major parks or green zones. For example, in Staten Island, regions with concentrated high tree density that are surrounded by similar green areas are identified with red highlights. Indeed, we can see on the map that these areas are surrounded by various large parks such as Bloomingdale Park and Clay Pit Ponds State Park. The

Low-Low clusters are represented in blue, which are found primarily in denser urban cores, mainly in the Bronx, Brooklyn, and Queens. The Spatial outliers are these High-Low and Low-High areas represented in pink and light blue, respectively. The High-Low areas are quite limited in NYC, which are usually the smaller and isolated parks in dense urban areas. Similarly, we found only a few small Low-High areas that have low tree density themselves but are surrounded by neighbors with higher tree density. However, given the definition, it seems unusual that Central Park in Manhattan has been marked as light blue as it contains 590 tree counts ranked at the top 37. Therefore, this might be due to its vast open space that may have much lower tree coverage compared to dense residential neighborhood areas with street trees, leading to a Low-High designation. Lastly, more than half of the tracts on this map are shown as not significant, indicating that tree density in these areas does not exhibit statistically substantial spatial clustering. They may have random or mixed distributions of tree density.



This tree density per square root percentage distribution map has zero percentage areas highlighted in orange. This map separates tracts into different classes according to their tree density. The darker the green is, the higher the percentage of tree density in the area, and vice versa. The lowest percentage can be rounded up to 0 in three significant figures and the highest tract's canopy coverage is 0.226%.

The dark green areas (high tree density classes) correspond to High-High clusters (red) in the LISA map. The orange areas (0% tree density) in the density map also align closely with some of the Low-Low clusters (blue), particularly in Manhattan, Staten Island, and parts of Brooklyn.

Discussion and Interpretation

Comparing Map 3 and Map 4, it is clear that many large area tracts have population density that seems to align with the distribution of tree density. This positive relationship is further supported by the OLS regression table that indicated a positive coefficient value. This finding objects to Hypothesis 1 as we initially believed they would have a reverse relationship that aligns with the general studies suggesting that population density involving heavy urban development and land use (e.g., roads and buildings) is in competition with forest land and displaces urban forests (Lin, et al., 2021).

Maps 5 to 7 highlight the residential segregation patterns among different racial groups in New York City. It is evident that people of the same race tend to live in neighboring areas, creating distinct geographic clusters based on race. Although maps have some overlapping areas, they reveal clear spatial patterns of racial distribution, where certain racial groups are more concentrated in specific neighborhoods. This finding can partly reflect the uneven distribution of tree density. The correlation between tree density and these demographic and social factors indicates the existence of social stratification in NYC. Social stratification implies that subpopulations lack the mobility to move to neighborhoods with greater greenery…(Lin, et al., 2021), and this idea is quite visible through these spatial maps.

The negative coefficients found in regression tables for all three racial groups suggest they are overall negatively correlated with tree density, as analysed earlier. These findings are surprising to me as I expected them to have a similar sign as the coefficient of the total population density. However, they indicate that a higher composition of either of these races is associated with a decrease in tree density, which satisfies the argument that more human existence reduces land for trees. Among these racial groups, the White population has the smallest negative coefficient, Asians the second, and Blacks the highest, which supports Hypothesis 3.

The household median income distribution map suggests great income disparities in this city. It has the clearest and strongest distinguishment between neighborhoods compared to maps of other variables, which further indicates the inequalities and the high concentration of wealth held in certain regions. Wealthier areas, such as parts of Manhattan, Staten Island, and certain neighborhoods in Brooklyn and Queens, are distinctly separated from lower-income areas like the Bronx and central Brooklyn. This stark contrast is anticipated, as previous studies have found. However, here we are comparing the income map with the tree density map, which we can conclude that this economic divide in NYC has significantly impacted resource allocation, at least the access to urban greenery. Fortunately, the income effects might only partially explain the tree distribution (Hypothesis 2) because many small area tracts in boroughs like Brooklyn and the Bronx are matched with quite high tree density. Although these high-density findings might result from smaller land areas, they still include the possibility of having fairly equal environmental management in these neighborhoods, which indicates less exploitation of the poor in urban tree planting. However, it is also worth noting that more lower-income tracts are highlighted in orange compared to neighborhoods in downtown Manhattan and southern Staten Island that are in higher-income classes. This finding may still emphasize the need for policies that address economic inequities and improve tree management, especially in areas where low-income overlaps with the orange mark or with low access to green spaces (light green).

The idea that higher distributional injustice more often happens in tracts with a greater proportion of Black residents is shared by various studies (Watkins, et al., 2016; Grant, et al., 2022). In fact, it is evident that the poor and minority communities are almost always facing fewer resources in street trees, which might indicate a sign of consistent socioeconomic marginalization. This disparity in tree distribution reflects deeper systemic issues, such as historical redlining, unequal investment in urban infrastructure, and persistent racial and economic segregation. Hence, the allocation of street trees is not just an environmental concern but a symbol of broader inequalities in more essential public resource distribution, urban planning, and prioritization of community well-being. Addressing this imbalance requires tackling the structural factors that perpetuate these injustices. We might need more specific reforms, such as policies that protect people from discrimination and fund public services in minority neighborhoods. Without resolving these root causes in the political or justice system, the inequities in urban greenery will continue to mirror and reinforce the socioeconomic divides in cities like NYC.

Limitations:

For this project, I used the 2015 NYC tree data and calculated the tree density after counting the tree numbers in each tract. This tree density as my dependent variable can be replaced by tree canopy cover (TCC), which refers to the proportion of an area covered by the vegetative portion of trees. TCC might provide more insights into the ecological and environmental benefits trees can offer, such as shading, air purification, and habitat for wildlife. Many studies also included different independent variables such as looking at education-, age-, and household characteristic-based inequalities (Lin, et al., 2021; Pistón, et al., 2022).

As mentioned earlier, the datasets I used do not come from the same year. These year differences might lead to data collection biases, thereby affecting the accuracy of the results. For example, tree planting and removal are likely to have occurred between 2015 (tree census data) and 2020 (demographic data), which might not fully reflect the current state of tree density or population distribution. Although resident movements are often less mobile, the street tree density is very likely to change over five years. Additionally, the use of aggregated data at the tract level could obscure finer-grained variations within neighborhoods, limiting the study's ability to capture micro-level disparities. Finally, this project does not account for the potential influence of historical policies, which might have long-term effects on tree distribution and demographic patterns. Hence, these effects should be considered as in the unobserved error term. These limitations suggest that while the findings provide valuable insights into tree density inequalities, further research with more synchronized and detailed data is necessary to improve the generalizability and accuracy of the results.

Conclusion

In conclusion, the crucial role of urban greenery and the problem of environmental resource inequality are highlighted through my findings and supported by past literature. Despite the revelation of the issue, uneven resource allocation persists, underscoring the urgent need for more effective actions to promote urban environmental equity. The main findings of this study include that racial groups correspond to a different decrease in tree density (White the least and Black the most). Additionally, while higher total population density is unexpectedly related to an overall increase in tree density, this may reflect differences in urban planning priorities rather than equitable resource distribution. Also, I can conclude that higher-income tracts show a positive association with tree density, though the income effect is relatively smaller than the racial effect.

These results partially support the hypotheses and emphasize the interplay of socio-demographic factors in determining tree distribution. The findings reveal systemic disparities that warrant policy interventions aimed at reducing social and environmental inequalities. Future research can incorporate additional factors such as tree health and status or explore the dynamic impact of tree planting and removal. Based on my findings, policymakers should prioritize green initiatives in underserved areas, particularly in low-income and Black communities, hence bridging the gap in environmental resources and enhance the equitable access to the benefits of urban greenery.

References:

2015 Street Tree Census - Tree Data | NYC Open Data. (n.d.). NYC Open Data. https://data.cityofnewyork.us/Environment/2015-Street-Tree-Census-Tree-Data/pi5s-9p35

De Menezes, M. O. T., Moro, M. F., De Souza, M. O., De Almeida, J. M., De Freitas Braga, D. V., De Alencar, P. L. R., & De Araújo, F. S. (2023). Patterns of mistletoe infestation in neotropical urban areas. *Urban Ecosystems*, *26*(4), 1081–1097. https://doi.org/10.1007/s11252-023-01363-7

Gianfredi, V., Buffoli, M., Rebecchi, A., Croci, R., Oradini-Alacreu, A., Stirparo, G., Marino, A., Odone, A., Capolongo, S., & Signorelli, C. (2021). Association between Urban Greenspace and Health: A Systematic Review of Literature. *International Journal of Environmental Research and Public Health*, *18*(10), 5137. https://doi.org/10.3390/ijerph18105137

Gomes, I. N., Bosenbecker, C., Silva, V. H., Cardoso, J. C., Pena, J. C., & Maruyama, P. K. (2023). Spatiotemporal availability of pollinator attractive trees in a tropical streetscape: unequal distribution for pollinators and people. *Urban Forestry & Urban Greening*, *83*, 127900. https://doi.org/10.1016/j.ufug.2023.127900

Grant, A., Millward, A. A., Edge, S., Roman, L. A., & Teelucksingh, C. (2022). Where is environmental justice? A review of US urban forest management plans. *Urban Forestry & Urban Greening*, *77*, 127737. https://doi.org/10.1016/j.ufug.2022.127737

Jiang, T., Deng, Z., Zhi, Y., Cheng, H., & Gao, Q. (2021). The effect of urbanization on population health: Evidence from China. *Frontiers in Public Health*, *9*. https://doi.org/10.3389/fpubh.2021.706982

Lin, J., Wang, Q., & Li, X. (2020). Socioeconomic and spatial inequalities of street tree abundance, species diversity, and size structure in New York City. *Landscape and Urban Planning*, *206*, 103992. https://doi.org/10.1016/j.landurbplan.2020.103992

Pistón, N., Filho, D. S. S., & Dias, A. T. (2022). Social inequality deeply affects people's perception of ecosystem services and disservices provided by street trees. *Ecosystem Services*, *58*, 101480. https://doi.org/10.1016/j.ecoser.2022.101480

Shackleton, C. (2016). Do Indigenous Street Trees Promote More Biodiversity than Alien Ones? Evidence Using Mistletoes and Birds in South Africa. *Forests*, *7*(7), 134. https://doi.org/10.3390/f7070134

Ventriglio, A., Torales, J., Castaldelli-Maia, J. M., De Berardis, D., & Bhugra, D. (2020). Urbanization and emerging mental health issues. *CNS Spectrums*, *26*(1), 43–50. https://doi.org/10.1017/s1092852920001236

Vlahov, D. (2002). Urbanization, urbanicity, and health. *Journal of Urban Health*, *79*(90001), 1S – 12. https://doi.org/10.1093/jurban/79.suppl_1.s1

Watkins, S. L., Mincey, S. K., Vogt, J., & Sweeney, S. P. (2016). Is Planting Equitable? An examination of the spatial distribution of nonprofit urban Tree-Planting programs by canopy cover, income, race, and ethnicity. *Environment and Behavior*, *49*(4), 452–482. https://doi.org/10.1177/0013916516636423

Wood, E. M., & Esaian, S. (2020). The importance of street trees to urban avifauna. *Ecological Applications*, *30*(7). https://doi.org/10.1002/eap.2149