Analyzing Tree Species Distribution and Diameters Across Urban Toronto Wards*

Yihang Xu

September 26, 2024

This paper investigates the distribution and density of tree species across various wards in an urban environment. Through the analysis of tree diameters and the prevalence of specific tree species, we uncover that certain species dominate the landscape, while others are relatively rare. Additionally, variations in tree density across wards highlight significant disparities in green infrastructure distribution.The findings have implications for urban planning, suggesting a need for more diverse species planting and equitable distribution of trees across neighborhoods.

Table of contents

*A GitHub Repository containing all data, R code, and other files used in this investigation is located here: https://github.com/YihangXu04/UrbanTreeData.git

1 Introduction

Urban forests are essential for the sustainability and livability of cities, offering numerous ecological, economic, and social benefits. In Toronto, where the urban heat island effect is a growing concern, urban trees help mitigate rising temperatures, improve air quality, and enhance biodiversity (Wolf et al. 2020). A healthy, diverse, and well-distributed urban forest can significantly enhance public health, lower energy costs, and increase property values (Czaja, Kołton, and Muras 2020). However, managing…

In this study, we analyze the distribution of tree species and diameters (DBH) across Toronto's urban wards. The primary objectives are to understand species dominance, disparities in tree density, and the size distribution of trees. This data-driven approach will inform future urban forestry strategies and policies (Pataki et al. 2021).

This paper is organized as follows: the **Data** section describes the dataset, including its sources and data cleaning. The **Results** section presents an analysis of tree species diversity, size distribution, and tree density across Toronto's wards. Finally, the **Discussion** section explores the implications of these findings for urban forestry management.

2 Data

2.1 Dataset Overview

The dataset used in this study is a collection of tree data from Toronto's urban forest, providing detailed records about the trees' physical characteristics and their locations. The data includes key variables such as tree diameters, species names, ward locations, and tree positions. It provides insight into the distribution of trees across different geographic areas of Toronto, with each entry representing an individual tree observation.

2.2 Broader Context

The dataset provides a comprehensive view of the current state of Toronto's urban trees. It is likely collected as part of municipal efforts to monitor and maintain urban forestry, which contributes to environmental sustainability and the well-being of urban residents. Urban forests play a critical role in reducing air pollution, mitigating the heat island effect, and enhancing urban biodiversity. This dataset is an invaluable resource for understanding tree species distribution, identifying areas in need of reforestation, and supporting future urban planning initiatives.

The dataset is unique to Toronto's context. Although other urban forestry datasets are available from cities like Vancouver and New York, they differ in geography, climate, and urban planning priorities. Therefore, while these datasets could offer some comparative insights, they were not chosen for this analysis as the focus is specifically on Toronto's urban ecosystem.

2.3 Key variables

X_id, OBJECTID, STRUCTID: These are unique identifiers for each tree, helping to differentiate individual observations. ADDRESS, STREETNAME, CROSSSTREET1, CROSSSTREET2: These variables provide geographic identifiers for each tree's location, giving insight into where the tree is positioned in the city. DBH_TRUNK (Diameter at Breast Height): This is a key variable used to assess the age and size of a tree. Tree diameters range from small, newly planted trees to larger, mature specimens. WARD: This variable identifies the political ward in which the tree is located, allowing for geographic analysis across Toronto's districts. BOTANICAL_NAME and COMMON_NAME: These variables identify the tree species by both scientific and common names, giving us insight into the species diversity across the city. geometry: Contains the latitude and longitude of each tree's position, allowing for mapping and spatial analysis.

Using the R programming language (R Core Team 2023),and `tidyverse` (**citetidyverse?**) packages.

2.4 Ethical Considerations

The dataset presents minimal ethical concerns. It contains no personally identifiable information (PII) and was collected by municipal authorities as part of efforts to manage Toronto's urban forest (Czaja, Kołton, and Muras 2020).

#Measurements

This dataset captures important attributes of Toronto's urban trees. Each entry in the dataset corresponds to a tree observed in the city, with several key measurements recorded. This section explains how real-world phenomena—such as tree size, species, and location—are translated into the dataset.

2.5 Tree Diameter

Diameter at Breast Height (DBH) is a standard measurement used to assess a tree's size and age. DBH is measured at 1.4 meters above ground. In this dataset, the DBH is recorded in centimeters under the DBH_TRUNK variable. The diameter is measured with a diameter tape or caliper during field surveys. This measurement serves as a proxy for a tree's maturity and overall health.

Importance: DBH is used to estimate the tree's contribution to the urban forest in terms of air quality, shade, and carbon storage. Larger diameters generally indicate older trees, which offer more environmental benefits. Accuracy: Measurements may vary due to obstacles like irregular trunk shapes, but the field process provides a reliable estimate of size.

2.6 Tree Species

Tree species are recorded by both scientific (botanical) and common names. Species identification is carried out by trained arborists based on observable characteristics such as leaves, bark, and growth form.

BOTANICAL_NAME: The scientific name of the tree, recorded for accuracy and consistency. COMMON_NAME: The common or colloquial name, more familiar to the general public. Both names are recorded to ensure the dataset can be used by a wide range of users, from scientists to the public.

2.7 Location Data

Location data is critical for understanding the distribution of trees. Several variables capture a tree's precise location:

STREETNAME, CROSSSTREET1, CROSSSTREET2: The streets nearest to the tree's location. WARD: The municipal ward where the tree is located, allowing analysis of tree distribution across political boundaries. geometry: Latitude and longitude coordinates, captured via GPS during field surveys, provide the exact location of each tree. Accuracy: The GPS coordinates are generally accurate, though occasional signal interference may cause minor discrepancies.

2.8 Unique Identifiers (X_id, OBJECTID, STRUCTID)

Each tree has a unique identifier to ensure that it can be tracked and managed in the dataset:

X_id, OBJECTID, STRUCTID: Automatically generated identifiers used to uniquely distinguish each tree, ensuring that the same tree is not recorded multiple times.

3 Results

3.1 Tree Species Distribution

The analysis revealed that Toronto's urban forest is dominated by a few species, particularly *Acer platanoides* (Norway Maple) and *Gleditsia triacanthos* (Honey Locust). The dominance of these species raises concerns about biodiversity (Pataki et al. 2021).

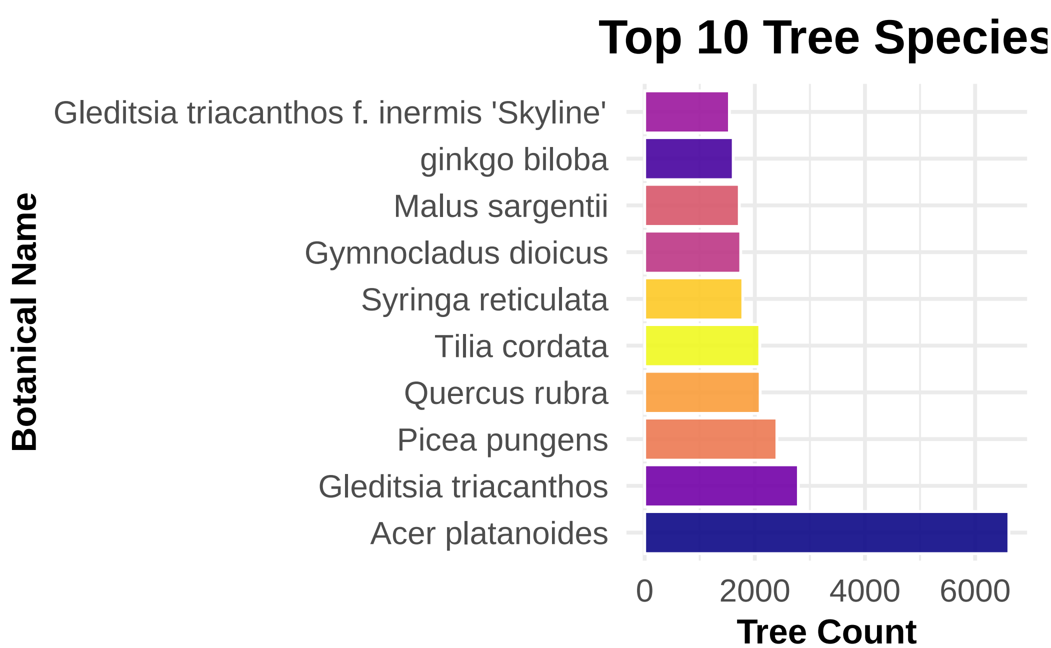

Figure 1: Distribution of Tree Species

3.2 Tree Diameter (DBH) Distribution

Most trees in Toronto's urban forest have a DBH between 10 cm and 50 cm, indicating that many of the trees are relatively young (Wolf et al. 2020).

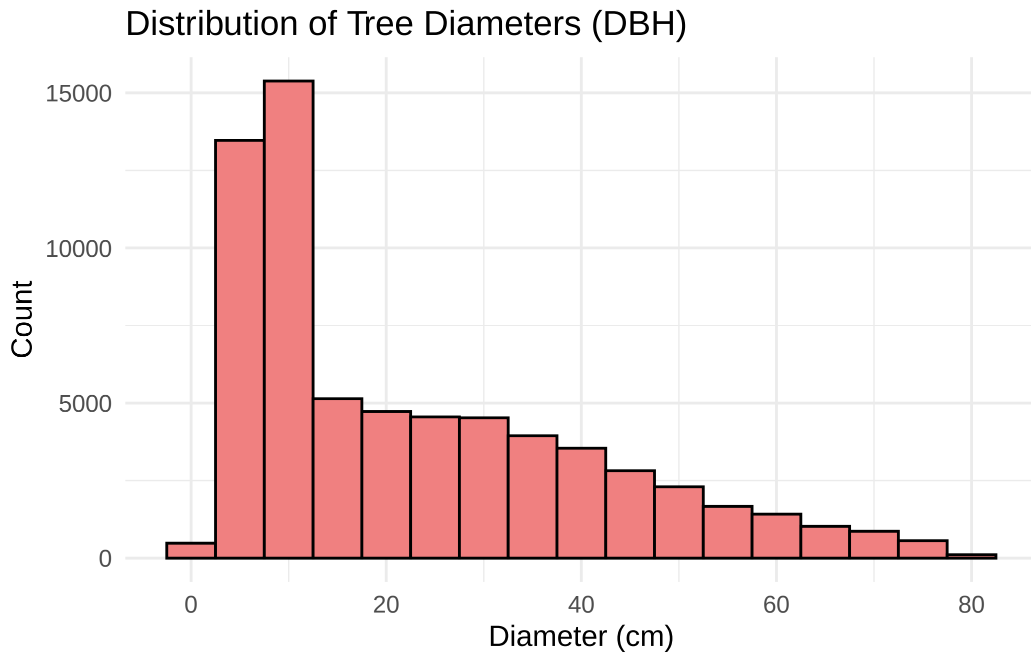

Figure 2: Tree Diameter Distribution

3.3 Tree Density by Ward

Tree density varies significantly across Toronto's wards, with some wards having far more trees than others. This disparity raises concerns about equitable access to green spaces across different neighborhoods (Czaja, Kołton, and Muras 2020).

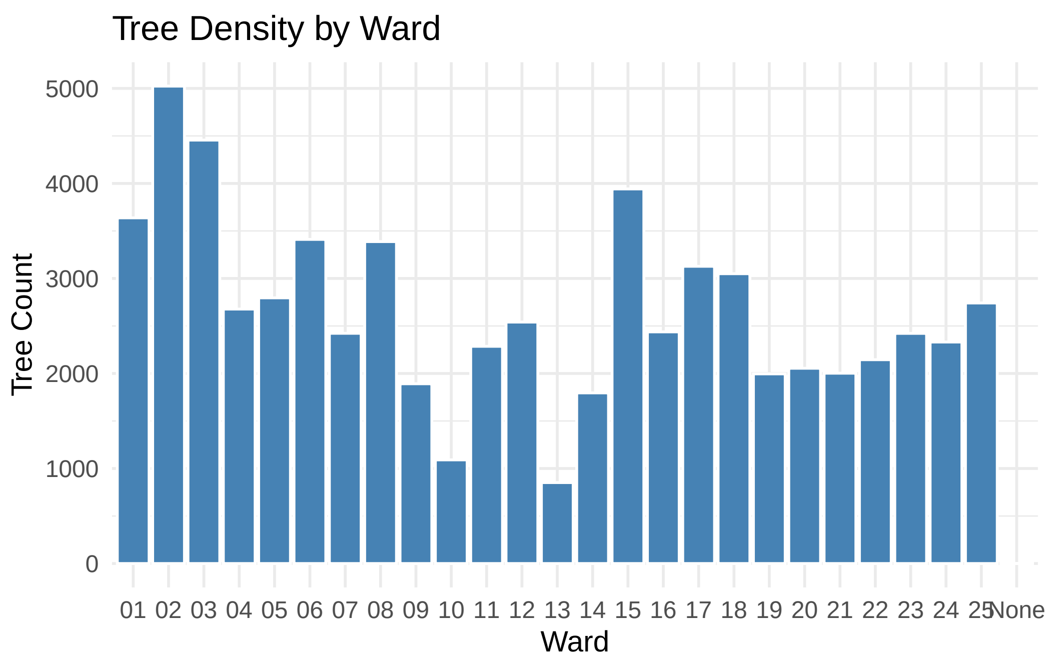

Figure 3: Tree Density by Ward

4 Discussion

The results of this analysis provide several important insights for the management of Toronto's urban forest. Firstly, the dominance of a few tree species, such as *Acer platanoides* (Norway Maple) and *Gleditsia triacanthos* (Honey Locust), highlights a potential risk of reduced biodiversity. When an urban forest lacks diversity, it becomes more vulnerable to pests, diseases, and environmental changes. For instance, if a pest or disease affects one of the dominant species, such as the Colorado Blue Spruce, the urban forest could experience significant tree loss. This vulnerability is well-documented in urban forestry literature, where monocultures or low-diversity tree populations have been shown to be more susceptible to environmental stressors (Pataki et al. 2021).

To address these concerns, urban forestry management in Toronto should prioritize increasing tree species diversity in future planting programs. This could involve introducing more native species that are better suited to local conditions and less likely to become susceptible to pests or diseases. Furthermore, diversifying tree species can create a more resilient ecosystem that offers better habitat opportunities for wildlife and enhances the ecological functions of the urban forest.

The analysis of tree diameters (DBH) also reveals important trends. Most trees in Toronto's urban forest are relatively small, with diameters between 10 and 50 cm. This indicates that many of the trees in the city have been planted recently, likely as part of efforts to increase the city's tree canopy. However, the lack of larger, older trees suggests that many trees may not be surviving to maturity. This could be due to several factors, including poor soil quality, limited space for root growth, exposure to pollution, or physical damage from urban infrastructure (Wolf et al. 2020). Larger, older trees provide greater environmental benefits than smaller trees, as they sequester more carbon, produce more oxygen, and provide more shade. Therefore, it is crucial to ensure that newly planted trees are cared for properly so they can survive to maturity and contribute to the urban environment in the long term.

Policymakers should consider implementing maintenance programs that provide regular care for newly planted trees, such as watering, pruning, and protection from construction and traffic-related damage. In addition, providing adequate space for root growth and improving soil conditions can help ensure that trees have the resources they need to thrive in an urban environment.

Another critical finding is the significant disparity in tree density across Toronto's wards. Some wards have a much higher tree density than others, which raises concerns about environmental justice. Neighborhoods with fewer trees are likely to experience higher temperatures, poorer air quality, and fewer social and psychological benefits compared to areas with more trees (Czaja, Kołton, and Muras 2020). This inequity may disproportionately affect lower-income communities, where access to green spaces is already limited.

A Appendix

A.1 Dataset and Graph Sketches

Sketches depicting both the desired dataset and the graphs generated in this analysis are available in the GitHub Repository.

A.2 Data Cleaning

The data cleaning process involved filtering out some of the columns from the raw dataset and renaming some of the data entries for clarity and simplicity.

A.3 Attribution Statement

"Contains information licensed under the Open Government Licence – Toronto" (**tphlicense?**).

References

Czaja, Monika, Anna Kołton, and Piotr Muras. 2020. "The Complex Issue of Urban Trees—Stress Factor Accumulation and Ecological Service Possibilities." *Forests* 11 (9): 932.

Pataki, Diane E, Marina Alberti, Mary L Cadenasso, Alexander J Felson, Mark J McDonnell, Stephanie Pincetl, Richard V Pouyat, Heikki Setälä, and Thomas H Whitlow. 2021. "The Benefits and Limits of Urban Tree Planting for Environmental and Human Health." *Frontiers in Ecology and Evolution* 9: 603757.

R Core Team. 2023. *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing. https://www.R-project.org/.

Wolf, Kathleen L, Sharon T Lam, Jennifer K McKeen, Gregory RA Richardson, Matilda van Den Bosch, and Adrina C Bardekjian. 2020. "Urban Trees and Human Health: A Scoping Review." *International Journal of Environmental Research and Public Health* 17 (12): 4371.