

Key Drivers of Community Garden Growth: Analyzing Trends in PollinateTO Funding Across Toronto's Neighborhoods*

A Bayesian Approach to Understanding the Impact of Time, Type, and Location
on Garden Size and Number.

Aliza Abbas Mithwani

November 27, 2024

First sentence. Second sentence. Third sentence. Fourth sentence.

Table of contents

1	Introduction	2
1.1	Overview	2
1.2	Estimand	2
1.3	Results	3
1.4	Significance	3
1.5	Telegraphing	3
2	Data	3
2.1	Software Packages	3
2.2	Data Source	4
2.3	Data Features	4
2.4	Data Measurement*** (check key)	4
2.5	Data Consideration	4
2.6	Methodology	5
2.7	Data Visualization	5
3	Model	8
3.1	Model set-up	9
3.1.1	Model justification	9

*Code and data are available at: [<https://github.com/alizamithwani/PollinateTO.git>]

4 Results	9
5 Discussion	9
5.1 First discussion point	9
5.2 Second discussion point	10
5.3 Third discussion point	10
5.4 Weaknesses and next steps	10
Appendix	11
A Additional data details	11
B Model details	11
B.1 Posterior predictive check	11
B.2 Diagnostics	11
References	12

1 Introduction

1.1 Overview

Urban biodiversity plays a crucial role in sustaining ecosystem health, yet cities face challenges in designing equitable and effective strategies to protect pollinators. Toronto’s PollinateTO initiative seeks to address this by funding community-led projects that create and enhance publicly visible pollinator gardens. Since its inception in 2019, the program has supported over 190 projects, resulting in approximately 500 gardens and 25,500 m² of pollinator habitat. While these efforts align with the city’s Pollinator Protection and Biodiversity Strategies, the allocation patterns of these grants remain an open question. Specifically, understanding whether certain types of gardens (e.g., rain gardens, boulevard gardens) or locations (e.g., Neighbourhood Improvement Areas) are more likely to receive funding can provide valuable insights into optimizing resources and maximizing program impact.

1.2 Estimand

To address this question, this paper investigates the relationship between garden characteristics and PollinateTO funding decisions. Our analysis focuses on how variables such as garden type, location, and community engagement strategies influence both the frequency and magnitude of funding allocations. The estimand centers on whether specific garden features are associated with higher funding likelihoods and amounts, filling a critical gap in understanding how urban biodiversity programs distribute resources.

1.3 Results

Our results reveal that certain garden types, such as _____, are more likely to secure funding, particularly in underserved areas identified as _____. We also find that projects emphasizing _____ receive higher allocations, highlighting the program’s alignment with broader engagement and ecological priorities. These findings offer actionable insights into how urban biodiversity initiatives can balance ecological goals with social equity considerations.

1.4 Significance

The implications of this study extend beyond Toronto, providing a framework for other cities seeking to implement or refine pollinator protection strategies. By shedding light on funding allocation patterns, this research informs policymakers and urban planners on how to enhance the effectiveness and equity of biodiversity programs. These insights contribute to the growing body of literature on sustainable urban development and community-driven ecological stewardship.

1.5 Telegraphing

The remainder of this paper is structured as follows. Section 2 describes the data used in the analysis, including its sources, key features, and considerations regarding measurement and visualization. Section 3 outlines the methodology, detailing the data visualization techniques employed and the justification for using a Bayesian logistic regression model to analyze the relationship between garden characteristics and funding outcomes. Section 4 presents the results, including model validation and key findings. Section 5 offers a discussion of the results, focusing on their implications through key points, followed by an examination of the study’s limitations and suggested future steps for further exploration. Section 6 concludes the paper. The Appendix provides supplementary materials, including raw data previews, data feature definitions, data visualizations, and technical model details such as posterior predictive checks, MCMC convergence diagnostics, and additional figures.

2 Data

2.1 Software Packages

We use the statistical programming language R (R Core Team 2023), rstanarm (Goodrich et al. 2022), tidyverse (Wickham et al. 2019) Our data (Environment & Climate 2024).... Following Alexander (2023), we consider...

2.2 Data Source

This study uses data from the PollinateTO Primary Project Garden Locations dataset, hosted on Toronto’s Open Data Portal. The dataset provides information on community-led garden projects funded by PollinateTO, including geographic location, garden type, funding year, and other features relevant to understanding the distribution and characteristics of these initiatives. It is publicly accessible and serves as a vital resource for analyzing PollinateTO’s impact in Toronto neighborhoods.

2.3 Data Features

The dataset includes several key variables:

YEAR_FUNDED: The year the project received funding from PollinateTO. **GARDEN_TYPE:** Categories of gardens (e.g., food gardens, pollinator gardens, rain gardens). **NIA_OR_EN:** Classification of garden location as part of a Neighbourhood Improvement Area (NIA), Emerging Neighborhood (EN), or other areas. **WARD_NAME:** The city ward where the garden is located. **IS_INDIGENOUS_GARDEN:** A binary variable indicating whether the garden is Indigenous-led. **PROJECT_NAME:** Names assigned to specific garden projects. **LATITUDE** and **LONGITUDE:** Geographic coordinates of each garden. These features allow for diverse analyses, from exploring spatial trends to assessing equity and accessibility in funding distribution.

2.4 Data Measurement*** (check key)

The dataset translates real-world community garden projects into structured entries by standardizing key attributes. For example:

YEAR_FUNDED records when projects were initiated, reflecting the temporal dimension of PollinateTO’s funding priorities. **GARDEN_TYPE** is based on project applications, categorized to highlight garden purposes and ecological roles. **NIA_OR_EN** and **WARD_NAME** capture the geographic and socioeconomic contexts of gardens, derived from city planning frameworks. **IS_INDIGENOUS_GARDEN** identifies projects promoting Indigenous cultural or ecological knowledge, based on project self-identification during funding application. The dataset provides structured snapshots of projects, enabling analysis of broader trends, equity impacts, and ecological contributions of urban greening initiatives.

2.5 Data Consideration

Several considerations affect the dataset’s completeness and reliability:

Self-Reported Data: Much of the data, including garden types and Indigenous-led status, relies on self-reporting by project applicants. Selection Bias: Funding decisions prioritize certain areas (e.g., NIAs), potentially overrepresenting them compared to other regions. Temporal Incompleteness: The dataset spans 2019 to present, limiting analyses to recent funding trends. Granularity: Certain variables, such as project outcomes or specific plant species, are not included, restricting the scope of ecological impact analysis. Geographic Precision: While coordinates are provided, some garden boundaries and precise sizes may not be captured, introducing potential spatial analysis challenges.

2.6 Methodology

Several considerations affect the dataset’s completeness and reliability:

Self-Reported Data: Much of the data, including garden types and Indigenous-led status, relies on self-reporting by project applicants. Selection Bias: Funding decisions prioritize certain areas (e.g., NIAs), potentially overrepresenting them compared to other regions. Temporal Incompleteness: The dataset spans 2019 to present, limiting analyses to recent funding trends. Granularity: Certain variables, such as project outcomes or specific plant species, are not included, restricting the scope of ecological impact analysis. Geographic Precision: While coordinates are provided, some garden boundaries and precise sizes may not be captured, introducing potential spatial analysis challenges.

2.7 Data Visualization

To understand the data provided better, we visualize it using tables and graphs. To begin with, let’s see how many gardens each type of neighborhood (NIA, EN or None) has and separate them by whether or not they are indigenous-led in Figure 1 and Figure 2.

Neighborhood Type (NIA or EN)	Total Gardens	Indigenous-Led Gardens	Non-Indigenous Gardens	Proportion Indigenous (%)
EN	6	1	5	16.666667
NIA	32	1	31	3.125000
None	111	3	108	2.702703

Figure 1: Summary of Gardens by Neighborhood Type

Next, we visualize the trend of garden sizes and numbers across time to examine the impact of the initiative over time in Figure 3.

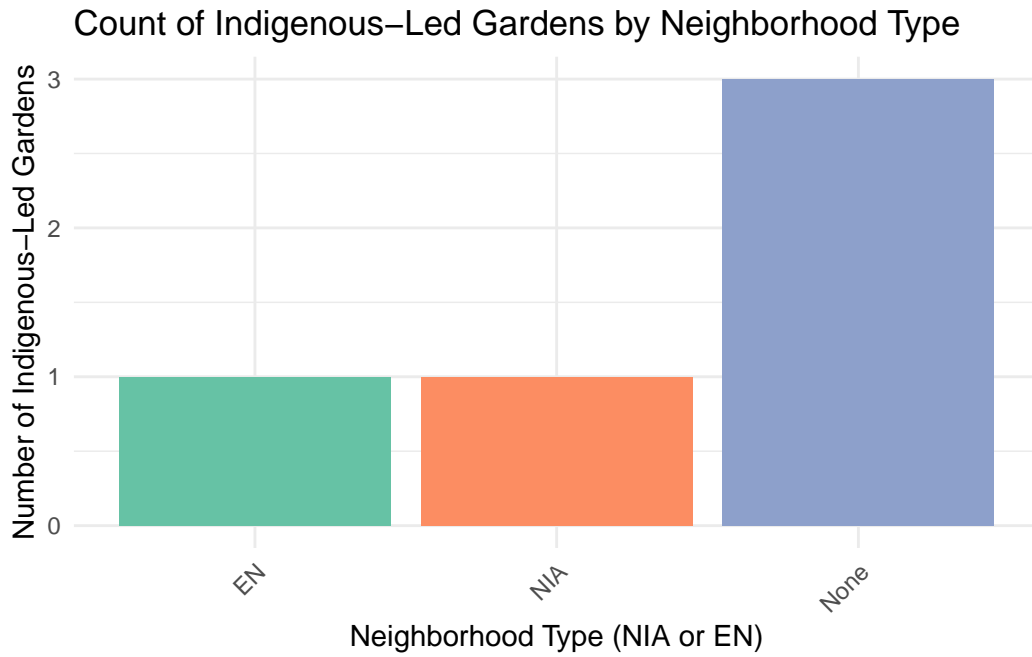


Figure 2: Distribution of Garden Types Funded Over Time

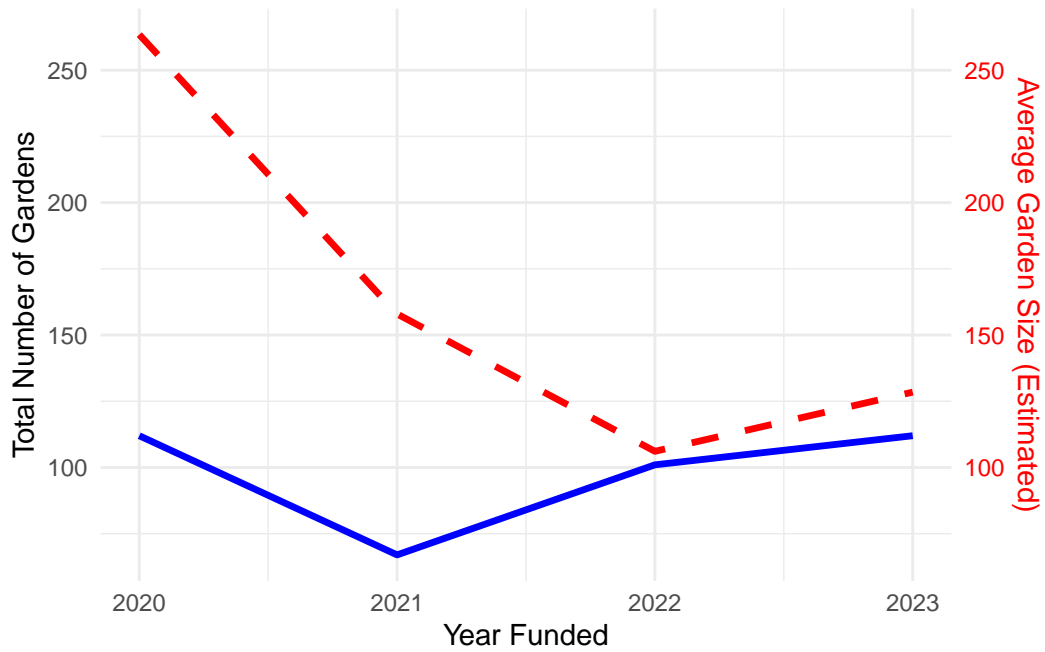


Figure 3: Trend of Garden Size and Number Over Time

Now, let's analyze the proportion of each garden type visually across time in Figure 4 to see if certain garden types increased in number and proportion overtime, reflecting shifting funding priorities and initiatives.

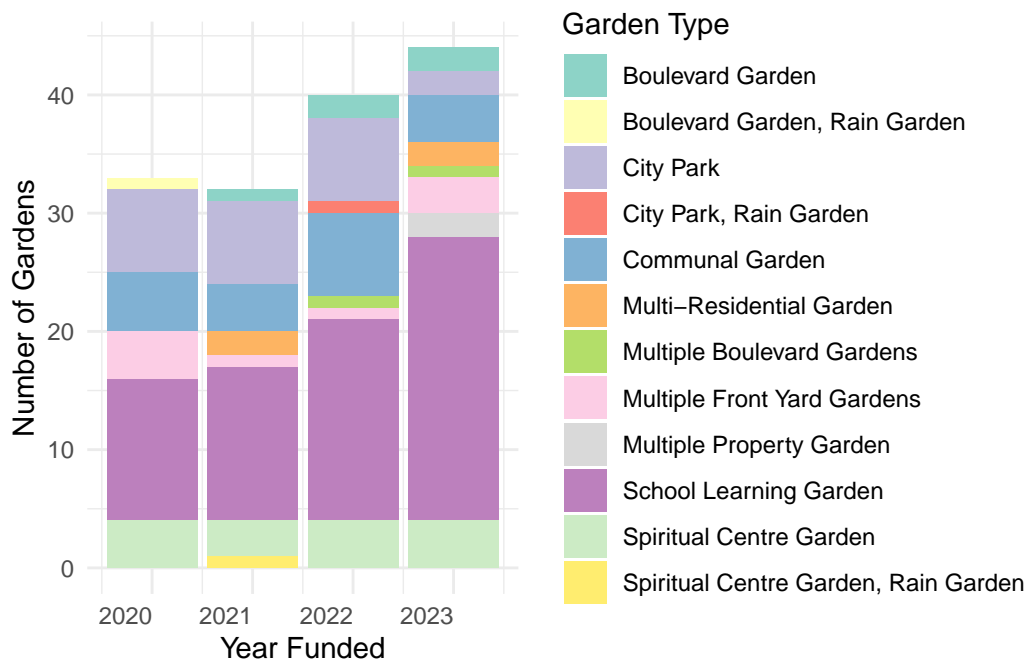


Figure 4: Distribution of Garden Types Funded Over Time

To prevent outliers from making graphs unreadable, we use log-transformed garden sizes to analyze differences in garden sizes based on neighborhood type in Figure 5. A boxplot is used to do this where the 3 neighborhood categories are EN (Emerging Neighborhoods), NIA (Neighborhood Improvement Areas) and None (neither EN nor NIA).

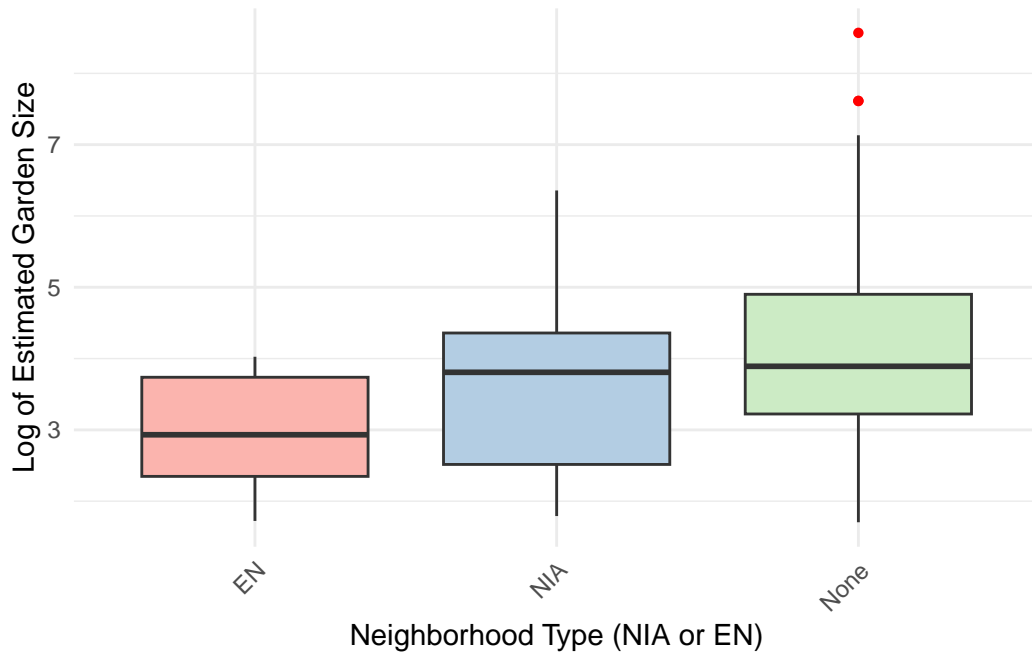


Figure 5: Log-Transformed Garden Sizes by Neighborhood Type

3 Model

In the analysis, I used a Bayesian Logistic Regression Model to examine the relationship between garden size and other factors like the type of garden, the year the project was funded, the neighborhood type (emerging neighborhood, EN vs neighborhood improvement area, NIA), and if it is an indigenous-led project.

The goal of our modelling strategy is twofold. Firstly,...

Here we briefly describe the Bayesian analysis model used to investigate... Background details and diagnostics are included in Appendix B.

Table 2: Explanatory models of flight time based on wing width and wing length

3.1 Model set-up

Define y_i as the number of seconds that the plane remained aloft. Then β_i is the wing width and γ_i is the wing length, both measured in millimeters.

$$y_i | \mu_i, \sigma \sim \text{Normal}(\mu_i, \sigma) \quad (1)$$

$$\mu_i = \alpha + \beta_i + \gamma_i \quad (2)$$

$$\alpha \sim \text{Normal}(0, 2.5) \quad (3)$$

$$\beta \sim \text{Normal}(0, 2.5) \quad (4)$$

$$\gamma \sim \text{Normal}(0, 2.5) \quad (5)$$

$$\sigma \sim \text{Exponential}(1) \quad (6)$$

We run the model in R (R Core Team 2023) using the `rstanarm` package of Goodrich et al. (2022). We use the default priors from `rstanarm`.

3.1.1 Model justification

We expect a positive relationship between the size of the wings and time spent aloft. In particular...

We can use maths by including latex between dollar signs, for instance θ .

4 Results

Our results are summarized in Table 2.

5 Discussion

5.1 First discussion point

If my paper were 10 pages, then should be at least 2.5 pages. The discussion is a chance to show off what you know and what you learnt from all this.

5.2 Second discussion point

Please don't use these as sub-heading labels - change them to be what your point actually is.

5.3 Third discussion point

5.4 Weaknesses and next steps

Weaknesses and next steps should also be included.

Appendix

A Additional data details

B Model details

B.1 Posterior predictive check

In `?@fig-ppcheckandposteriorvsprior-1` we implement a posterior predictive check. This shows...

In `?@fig-ppcheckandposteriorvsprior-2` we compare the posterior with the prior. This shows...

Examining how the model fits, and is affected
by, the data

B.2 Diagnostics

`?@fig-stanareyouokay-1` is a trace plot. It shows... This suggests...

`?@fig-stanareyouokay-2` is a Rhat plot. It shows... This suggests...

Checking the convergence of the MCMC algo-
rithm

References

- Alexander, Rohan. 2023. *Telling Stories with Data*. Chapman; Hall/CRC. <https://tellingstorieswithdata.com/>.
- Environment & Climate. 2024. *PollinateTO Primary Project Garden Locations*. <https://open.toronto.ca/dataset/pollinateto-primary-project-garden-locations/>.
- Goodrich, Ben, Jonah Gabry, Imad Ali, and Sam Brilleman. 2022. “rstanarm: Bayesian applied regression modeling via Stan.” <https://mc-stan.org/rstanarm/>.
- R Core Team. 2023. *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing. <https://www.R-project.org/>.
- Wickham, Hadley, Mara Averick, Jennifer Bryan, Winston Chang, Lucy D’Agostino McGowan, Romain François, Garrett Golemund, et al. 2019. “Welcome to the tidyverse.” *Journal of Open Source Software* 4 (43): 1686. <https://doi.org/10.21105/joss.01686>.