The Impact of Urbanization, Culture, and Wealth on Public Art in Toronto*

Alaina Hu

March 20, 2024

Public art has emerged in many urban cities across the world. Existing research and analysis on public art often focuses on the impact of public art on various factors, making public art the explanatory variable. This paper examines public art as a response variable to explanatory variables such as population, minority population, and income as a measure of urbanization, culture, and wealth respectively for the city of Toronto. Through data visualization, we find a moderately positive relationship between urbanization and public art but no relationship between diversity and public art or wealth and public art.

Table of contents

1 Introduction 2

2 Data 3

2.1 Measurement . . . 3
2.2 Public Art . . . 4
2.3 Ward Profiles (25-Ward Model) . . . 4
2.4 Public Art by Ward . . . 5
2.5 Public Art, Population, Minority Population, Income by Ward . . . 8
2.5.1 Examining Population on Public Art . . . 9
2.5.2 Examining Cultural Diversity on Public Art . . . 10
2.5.3 Examining Income and Public Art . . . 11

3 Model 11

3.1 Model set-up . . . 11
3.1.1 Simple Linear Regression . . . 12

*Code and data are available at: https://github.com/alainahu/public_art_revised

3.1.2 Multiple Linear Regression . . . 12
3.1.3 Poisson Regression . . . 12
3.1.4 Negative Binomial Regression . . . 13
3.2 Model justification . . . 13

4 Results **15**

5 Discussion **16**
5.1 Key Findings . . . 16
5.2 Weaknesses and Limitations . . . 16
5.3 Validity . . . 17
5.4 Further Steps . . . 18

Appendix **19**

A Model details **19**

B Checking for issues **19**
B.1 Posterior predictive check . . . 19

References **21**

1 Introduction

Public art installations, or art pieces available to the general public in open spaces can be found in many urban cities across the world. The city of Toronto has over 400 public art installations provided by two organizations, Public Arts & Monument Collection and the Percent for Public Art Program. Public art not only enhances the urban environment, but it also brings together visitors and residents with local culture and history, creating an educational and bonding experience (Barone 2023).

Public art design is heavily integrated with urban planning and development as a whole. It is crucial for overall public art design to consider the overall coordination of the environment, meaning that public art should consider and reflect the environment that it is in (Meng and Ye 2022). Furthermore, public art is also an important symbol of maturity in a city (Liu 2021). Thus, our research is interested in the factors that influence public art design and selection of location. We analyze the trend between urbanization, cultural diversity, and wealth on the locations of public art installations in Toronto.

Existing research on public art often focuses on the effects of public art, making public the explanatory variable. Wright and Herman (2018) examine how public art created by Black artists in the Third Ward of Houston engages with gentrification and ghettoization. Smith

(2014) looks at the effect of public art on societal interactions that reflect complexities of Egyptian society. Our research adds on to the literature by conducting analysis on public art as a response variable. While the societal and public art relationship could be bidirectional, we are interested in exploring the relationship between socioeconomic factors on public space usage with statistical methods. The estimand is the impact of population, minority population, and income on number of public art installations. We find a positive relationship between population and number of public art installations but no relationship between minority population and public art or income and public art. We conclude that urbanization increases the the number of public art installations in an area when holding cultural diversity and wealth constant.

Our research paper begins with the Data section (Section 2) to visualize and further understand the measurement, source, methodology, and variables we are examining. Then, we introduce the Model (Section 3) used to understand the relationships in the data and report the findings in the Results section (Section 4). Finally, we include the Discussion (Section 5) of the findings, summarizing the takeaway and future of this research.

2 Data

Data analysis is performed in `R` (R Core Team 2022), and additional help is provided by libraries such as `dplyr` (Wickham et al. 2023), `ggplot2` (Wickham 2016), `ggrepel` (Slowikowski 2024), `tidyverse` (Wickham et al. 2019), `kableExtra` (Zhu 2021), `knitr` (Xie 2023), and `sf` (Pebesma and Bivand 2023), `opendatatoronto` (Gelfand 2022), `readxl` (Wickham and Bryan 2023), `here` (Müller 2020), `rstanarm` (Goodrich et al. 2024), `arrow` (Richardson et al. 2023), `tidybayes` (Kay 2023), `modelsummary` (Arel-Bundock 2022), `broom` (Bolker and Robinson 2022), and `parameters` (Lüdecke et al. 2020). Data for this research comes from Open Data Toronto (Gelfand 2022), an open source data portal containing various topics of data for the city. For the data involved in this paper, we combine `Public Art` (Toronto 2023a) and `Ward Profiles (25-Ward Model)` (Toronto 2023b)

2.1 Measurement

In order to examine urbanization, cultural diversity, and wealth, we use population, minority population, and average household income as the explanatory variables to represent these factors. Population dynamics - the driving forces behind the growth and movement of populations, play an important role in the urbanization process (Salas 1986). Toronto is an important cultural center, with many ethnic groups organizing traditional festivals, entertainment events, and social activities (Howarth, n.d.). Household income is not the sole indicator of wealth, but it is heavily related to wealth as both income and wealth are key indicators of financial security (Schaeffer 2021). Supported by the literature, we use population, minority population and household income data to measure urbanization, cultural diversity, and wealth. To analyze this trend across Toronto, we look at the effect of these variables across the 25 wards of Toronto.

2.2 Public Art

The Public Art raw dataset includes information such as the source, artist, title, medium, installation year, and ward location on the 409 public art pieces in Toronto. Since we are only interested in the art pieces and their respective ward location, we focus on these two aspects of the data. As shown in Table 1, the data… . shows the part of the public art data we focus on.

Table 1: Sample of cleaned public art data

Art ID	Ward	Ward Name
1	14	Toronto-Danforth
2	13	Toronto Centre
3	11	University-Rosedale
4	11	University-Rosedale
5	14	Toronto-Danforth

2.3 Ward Profiles (25-Ward Model)

In the 25-Ward model, the city of Toronto is separated into 25 neighborhoods or voting districts. To better understand the effect of urbanization, cultural diversity, and wealth on the location of public art pieces, we compare the effect of population, minority population, income on the number of art works across the 25 wards. The `Ward Profiles (25-Ward Model)` data contains demographic, social, and economic information for each ward such as population, households, families, education, ethnocultural composition, spoken languages, income and housing costs. For our purpose of research, we are interested in population, minority population, and average income of each ward.. A sample of the cleaned dataset for the wards is shown below in Table 2.

Table 2: Sample of cleaned Toronto 25 Ward data

Ward	Population	Minority Population	Income
1	115120	90130	95200

2	117200	37210	146600
3	139920	48675	127200
4	104715	30445	127200
5	115675	67120	88700

2.4 Public Art by Ward

As we are interested in the number of public art pieces in each ward, we take the cleaned art data and group the wards together. To better understand our dependent response variable of the number of public art installations by ward, we organize the artworks by ward collect a count. Table 3 shows the number of art pieces by ward for the 25 wards of Toronto. To better visualize the spread of the number of pieces in each ward, we create a bar graph shown in Figure 1.

Table 3: Number of public art installations by ward.

Ward	Ward Name	Number of Public Art Pieces
1	Etobicoke North	1
2	Etobicoke Centre	3
3	Etobicoke-Lakeshore	13
4	Parkdale-High Park	23
5	York South-Weston	7
6	York Centre	9
7	Humber River-Black Creek	0
8	Eglinton-Lawrence	4
9	Davenport	0
10	Spadina-Fort York	157
11	University-Rosedale	46
12	Toronto-St. Paul's	14
13	Toronto Centre	61
14	Toronto-Danforth	12
15	Don Valley West	2
16	Don Valley East	2
17	Don Valley North	19
18	Willowdale	14
19	Beaches-East York	3
20	Scarborough Southwest	3
21	Scarborough Centre	11
22	Scarborough-Agincourt	1
23	Scarborough North	1
24	Scarborough-Guildwood	2
25	Scarborough-Rouge Park	1

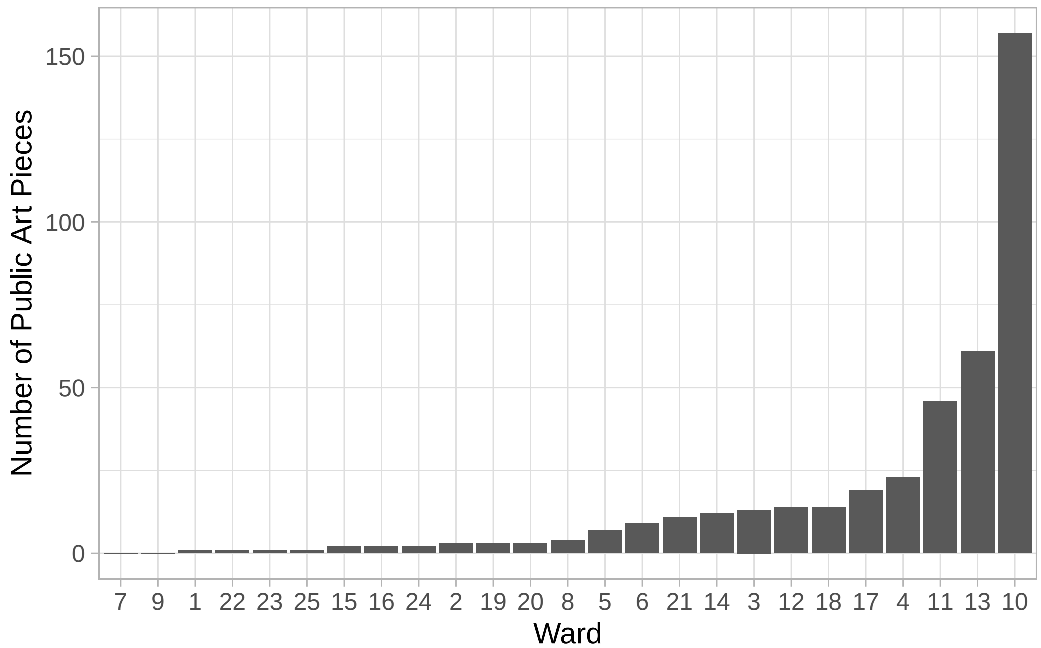


Figure 1: Number of public art installations by ward.

From Figure 1, we can see that Ward 10: Spadina-Fort York, Ward 13: Toronto-Centre, and Ward 11: University-Rosedale have the most number of public art installations while Ward 7: Humber River-Black Creek, Ward 9: Davenport, and Ward 1: Etobicoke North have the least number of public art installations. This aligns with intuition and literature regarding the connection between public art and urbanization. This can be visualized in the map below in Figure 2. The 3 wards with the most public art installations are highlighted. As shown, the wards with the most public art installations are in the main urban center of the city, Downtown Toronto.

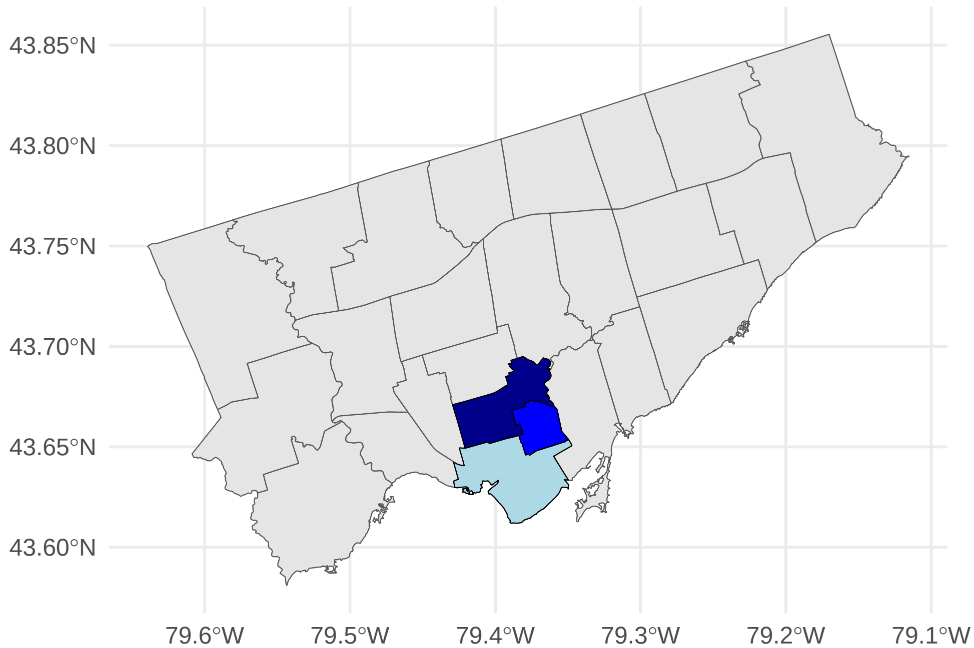


Figure 2: Map of Toronto highlighting the wards with the most art work

Table 4: Summary Statistics

Variable	Mean	Median	Standard. Deviation	Min	Max
Art Works	16.36	4	32.72	0	157
Population	120096.00	107300	33980.64	85700	224800
Minority Population	61492.00	67120	20380.90	30445	90130
Income	110451.60	110095	10593.87	94025	139920

2.5 Public Art, Population, Minority Population, Income by Ward

As the goal of this research is to analyze the impact of urbanization, cultural diversity, and income on the locations of public art pieces, we combine the Public Art Data grouped by ward with the Ward Profile Data to create the analysis data we are interested in. The analysis dataset includes the ward number, ward name, population of the ward, minority population in the ward, and average total income of households in 2020. Below in Table 5 is a sample of the analysis data.

Table 5: Sample of Data of All Variables

Ward	Ward Name	Population	Minority Population	Income	Public Art
1	Etobicoke North	115120	90130	95200	1
2	Etobicoke Centre	117200	37210	146600	3
3	Etobicoke-Lakeshore	139920	48675	127200	13
4	Parkdale-High Park	104715	30445	127200	23
5	York South-Weston	115675	67120	88700	7

2.5.1 Examining Population on Public Art

As one of our variables of interest, we are determined to examine the relationship between human population and the locations of public art at the ward level. We expect population and the number of public art installations to be positively related because the prosperity and development of cities provide space and development opportunities for the development of public art (Yin and Chang 2019). Public art is often a reflection of urban development, and population distribution is a direct proxy of urbanization (Qizhi, Ying, and Kang 2016). To visualize the relationship of interest, we plot population with the number of public art works.

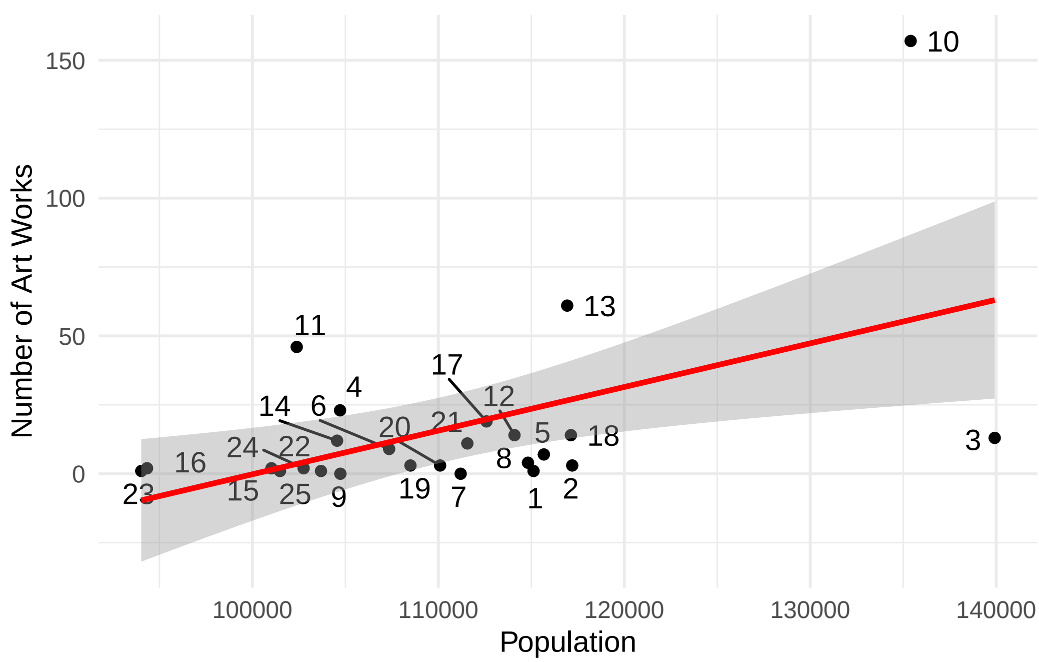


Figure 3: Positive relationship in population and public art by ward number

As expected, we see a moderate positive relationship between population of a ward and the number of public art installations in the ward. Furthermore, Ward 10 is an outlier and has both a high population and high number of public art pieces. Intuitively, this aligns with our beliefs and confirms the trend between high population and urbanization with high concentration of public art pieces.

2.5.2 Examining Cultural Diversity on Public Art

Another explanatory variable we are interested in is the cultural diversity of a region. In this case, we visualize the relationship between minority population and the number of public art works per ward. We expect to see a positive relationship between these variables. Figure 4 displays the visualization.

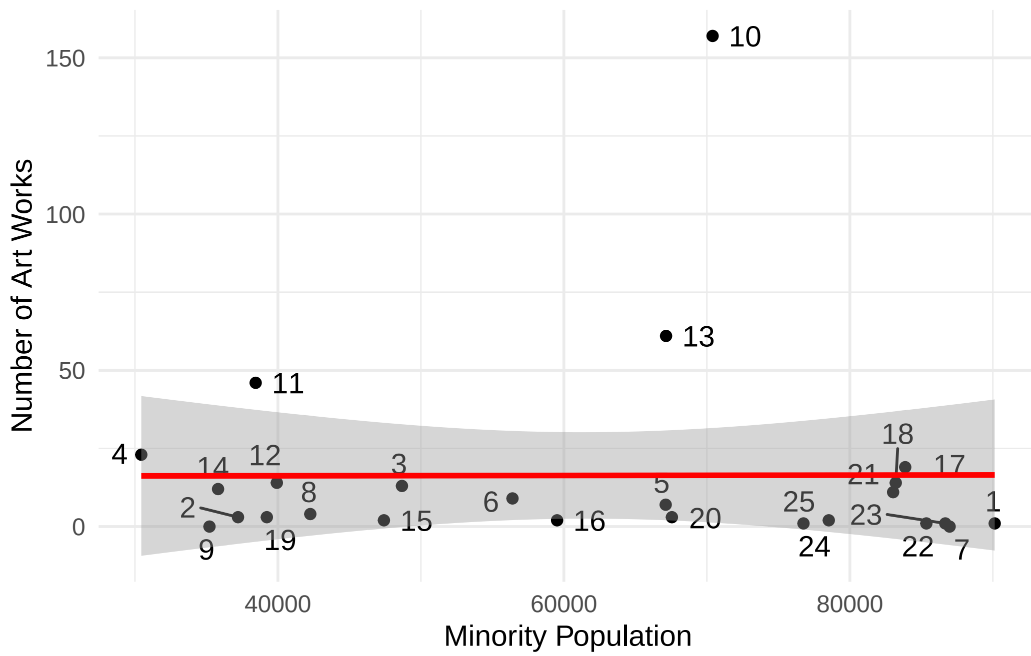


Figure 4: No relationship between minority population and public art works by ward number

Through the plot, we see that there is no relationship between minority population of a ward and the number of public art works in the ward. It is possible that the location of public art pieces is not a reflection of its cultural diversity. Ward 11 is in the Downtown region but has a lower minority population compared to Ward 13 and Ward 10, also wards located in Downtown. This observation leaves room for further research and investigation into the history and background behind Ward 11's population makeup.

2.5.3 Examining Income and Public Art

Lastly, in this research we are hoping to look at the relationship between income by ward and the number of public art pieces. Figure 5 plots the average household income in 2020 by ward with the number of art pieces by ward.

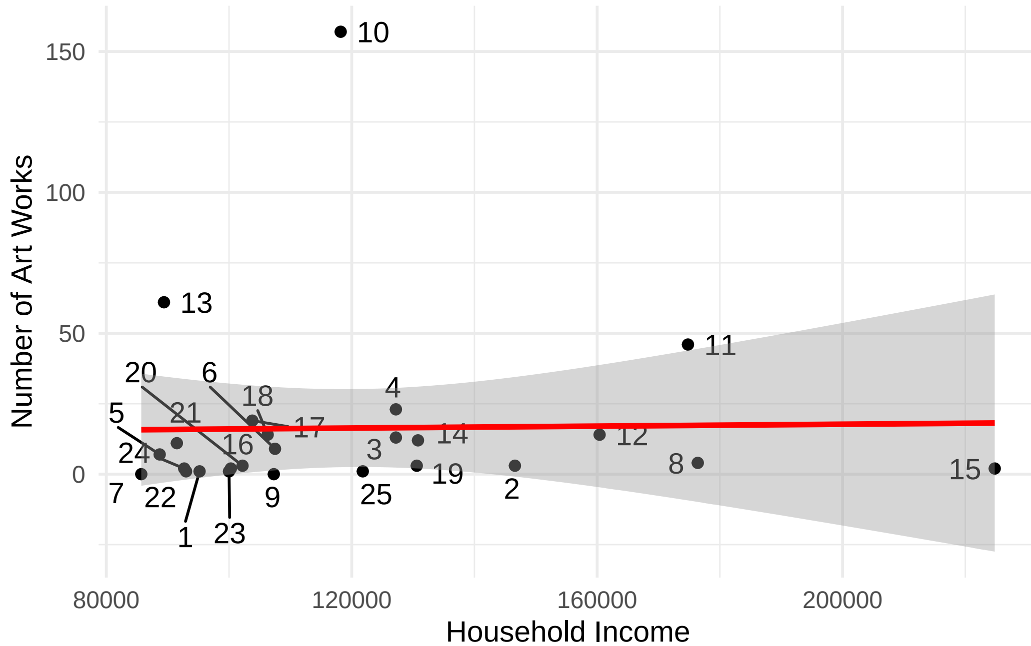


Figure 5: No relationship between income and public art by ward number

Similar to the relationship between minority population and public art, we observe no relationship between average household income level and public art.

3 Model

Here we briefly describe the Bayesian analysis models used to investigate.

3.1 Model set-up

From the data visualization performed in the Data section, we observe a moderately positive relationship between population and number of public art installations. Here, we build a simple linear regression model to further explore the population and public art relationship. Next, we build a multiple linear regression model between all explanatory variables and public art. Although our data visualizations did now show any relationship between minority population with public art and average household income level with public art, we add these two variables into the multiple regression model to act as controls. Through the multiple regression model, we are able to assess the number of public art installations related with population while adjusting for the explanatory variables of minority population and income. Ih the poisson and negative binomial regression models, we include all explanatory variables as well.

3.1.1 Simple Linear Regression

Define y_i as the number of public art pieces in the ward i. Then $population_i$ is the population of ward i

$$
\begin{aligned}
y_i|\mu_i, \sigma &\sim \text{Normal}(\mu_i, \sigma) && (1)\\
\mu_i &= \beta_0 + \beta_1 \times \text{population}_i && (2)\\
\beta_0 &\sim \text{Normal}(0, 82) && (3)\\
\beta_1 &\sim \text{Normal}(0, 7.7) && (4)\\
\sigma &\sim \text{Exponential}(0.031) && (5)
\end{aligned}
$$

3.1.2 Multiple Linear Regression

Define y_i as the number of public art pieces in the ward i. Then $population_i$ is the population of ward i, $minority_i$ is the minority population of ward i, and $income_i$ is the average household income of ward i.

$$
\begin{aligned}
y_i|\mu_i, \sigma &\sim \text{Normal}(\mu_i, \sigma) && (6)\\
\mu_i &= \beta_0 + \beta_1 \times \text{population}_i + \beta_2 \times \text{minority}_i + \beta_3 \times \text{income}_i && (7)\\
\beta_0 &\sim \text{Normal}(0, 82) && (8)\\
\beta_1 &\sim \text{Normal}(0, 7.72) && (9)\\
\beta_2 &\sim \text{Normal}(0, 4.01) && (10)\\
\beta_3 &\sim \text{Normal}(0, 2.41) && (11)\\
\sigma &\sim \text{Exponential}(0.031) && (12)
\end{aligned}
$$

3.1.3 Poisson Regression

Define y_i as the log count of public art pieces in the ward i. Then $population_i$ is the population of ward i, $minority_i$ is the minority population of ward i, and $income_i$ is the average household income of ward i.

$$
\begin{aligned}
y_i|\lambda_i &\sim \text{Poisson}(\lambda_i) && (13)\\
log(\lambda_i) &= \beta_0 + \beta_1 \times \text{population}_i + \beta_2 \times \text{minority}_i + \beta_3 \times \text{income}_i && (14)\\
\beta_0 &\sim \text{Normal}(0, 2.5) && (15)\\
\beta_1 &\sim \text{Normal}(0, 0.236) && (16)\\
\beta_2 &\sim \text{Normal}(0, 0.123) && (17)\\
\beta_3 &\sim \text{Normal}(0, 0.074) && (18)
\end{aligned}
$$

3.1.4 Negative Binomial Regression

Define y_i as the log count of public art pieces in the ward i. Then $population_i$ is the population of ward i, $minority_i$ is the minority population of ward i, and $income_i$ is the average household income of ward i.

$$
\begin{aligned}
y_i|\lambda_i, \alpha &\sim \text{NegativeBinomial}(\lambda_i, \alpha) && (19)\\
\log(\lambda_i) &= \beta_0 + \beta_1 \times \text{population}_i + \beta_2 \times \text{minority}_i + \beta_3 \times \text{income}_i && (20)\\
\beta_0 &\sim \text{Normal}(0, 2.5) && (21)\\
\beta_1 &\sim \text{Normal}(0, 0.236) && (22)\\
\beta_2 &\sim \text{Normal}(0, 0.123) && (23)\\
\beta_3 &\sim \text{Normal}(0, 0.074) && (24)\\
\alpha &\sim \text{Exponential}(1) && (25)
\end{aligned}
$$

We run the model in R (R Core Team 2022) using the `rstanarm` package of Goodrich et al. (2024). Initially, we use the default priors from `rstanarm`, however, we allow `rstanarm` to improve the priors by scaling them based on the data. We allow auto-scaling and run both models with the updated priors specified above.

3.2 Model justification

We expected a positive linear relationship between population and the number of public art installations. Thus, our first model was a simple linear regression model with population as the only explanatory variable. From our exploratory data analysis, we observed that there was a positive correlation between population and public art installations but no relationship from minority population and income. However, we wanted to look at the effect of population on public art while controlling for the other variables. This prompted us to develop a multiple regression model with all the variables of interest. From Appendix Figure 8a, we see that the multiple regression model is not a good fit for the observed data.

To improve our model, we consider the Poisson regression model with all the explanatory variables because we have count data, the number of public art pieces. From Appendix Figure 8b, we see that the Poisson regression is an improved fit from the multiple regression model. However, the key assumption that the mean and variance are equal is violated. From Table 4, we see that mean and variance are not equal.

Since an important assumption for the Poisson regression model does not hold, we build a negative binomial model. We can relax the assumption of mean and variance as equal in negative binomial model. The negative binomial regression model is a close variant of the Poisson model with looser assumptions. However, from Appendix Figure 8c, we see that the model does not capture the full range of observed data, so this model does not capture the data well and would need adjustment.

Table 6: Explanatory models of public art works based on population

	Simple linear	Multiple regression	Poisson	Negative binomial
(Intercept)	−156.58	−172.93	−6.20	−6.74
	[−284.37, −29.66]	[−337.35, −6.51]	[−7.60, −4.80]	[−14.33, 0.62]
population	1.56	1.58	0.07	0.09
	[0.43, 2.70]	[0.40, 2.76]	[0.06, 0.08]	[0.04, 0.15]
minority_population		0.08	0.01	−0.02
		[−0.73, 0.87]	[0.00, 0.02]	[−0.06, 0.01]
income		0.06	0.01	0.00
		[−0.41, 0.53]	[0.00, 0.01]	[−0.02, 0.03]
Num.Obs.	25	25	25	25
R2	0.239	0.280		
R2 Adj.	−0.076	−0.135		
Log.Lik.	−119.188	−119.735	−304.694	−86.414
ELPD	−126.6	−127.6	−379.3	−91.4
ELPD s.e.	11.3	10.5	104.0	8.6
LOOIC	253.2	255.2	758.6	182.7
LOOIC s.e.	22.7	21.0	208.0	17.2
WAIC	250.9	253.3	809.2	181.4
RMSE	27.54	27.54	26.56	45.84

4 Results

Table 6 displays the results for the simple linear model, multiple regression model, Poisson model, and negative binomial model. All four models have negative intercepts and positive coefficients for the population variable.

In the simple linear model, the negative y-intercept of -156.58 means that when population in a ward is zero, the number of public art installations is -156.58. However, this is not a practical interpretation and the y-intercept just helps with level effects. The coefficient of 1.56 for population means that for every increase of 1000 in population, the number of public art pieces increases by 1.56.

In the multiple regression model, the negative y-intercept of -172.93 helps level the number of art installations and has no practical significance because we cannot have a negative number of public art pieces. The coefficient of 1.58 means that when controlling for minority population and income, an increase of 1000 in the population increases the number of public art pieces by 1.58. When comparing the population coefficient of the multiple regression model with the population coefficient in the simple linear regression, we notice that the coefficient is larger in the multiple regression model. After controlling for minority population and income, population has a greater effect on the number of public art pieces.

In the Poisson model, the negative intercept continues to have no practical meaning. The coefficient for the population variables indicates that for each increase in population of 1000 people, the log count of public art pieces is expected to increase by 0.07 after holding minority population and income constant.

In the negative binomial model, the negative intercept has no practical meaning again. The coefficient for the population variables indicates that for each increase in population of 1000 people, the log count of public art pieces is expected to increase by 0.09 after holding minority population and income constant.

The simple linear model, multiple regression model, Poisson model, and negative binomial model show that there is no relationship between minority population and art work or income and art work as the coefficients for these variables all have a confidence interval that includes zero.

The simple linear model, multiple regression model, Poisson model, and negative binomial model all show a positive relationship between population and the number of public art installations even after controlling for minority population and income. The model results differ in the magnitude of the effect of population on the number of public art pieces.

5 Discussion

5.1 Key Findings

From the data visualizations and observed relationships in the Data section, we see that there is a positive relationship between population and number of public art installations, but there is no relationship between minority population with number of public art pieces or income level with number of public art pieces. We further investigate this relationship by building models. We develop a simple linear regression model, multiple linear regression model, Poisson regression model, and negative binomial regression model. The model results for the four models all show a positive relationship between population and the number of public art installations even after controlling for minority population and income. The model results differ in the magnitude of the effect of population on the number of public art pieces. For the variables of minority population and income, we see little or no effect on public art due to their small coefficients that are close to zero.

5.2 Weaknesses and Limitations

From the Appendix posterior prediction checks in Figure 8a, Figure 8b, and Figure 8c, we can see that the multiple regression model and negative binomial model do not fit the data as well as the Poisson regression model. This is an interesting case because our data violates an important assumption of the Poisson regression model: equal mean and variance. In the data, we have 25 observations for the variables since we are comparing data at the ward level across the 25 Toronto wards. The small number of observations combined with the model fit indicates a potential data problem in the research. The relatively small sample size of 25 observations is a critical factor to consider, as it might not only impact the robustness of the statistical models but also reflect on the generalizability of the findings. Small sample sizes can lead to higher variability and may affect the model's ability to accurately capture the underlying distribution of the data.

The superior fit of the Poisson regression model, despite the violation of its core assumption in our data, suggests that the model has inherent flexibility or that the impact of this assumption might not be as critical under certain conditions. This finding prompts a deeper investigation into the nature of the data and the model's assumptions.

Moreover, this scenario underscores a potential issue within the dataset itself. The mismatch between the expected model conditions and the observed data characteristics suggests that there may be underlying factors affecting the data quality or distribution that were not accounted for in the initial analysis. This could range from measurement errors to unaccounted-for variables that could significantly influence the outcomes of the ward-level comparisons. Therefore, this analysis does not merely highlight a statistical anomaly but points to a larger data problem that could have implications for the research's validity and reliability.

In conclusion, the comparative analysis of the regression models provides valuable insights into the complexities of statistical modeling, especially when dealing with real-world data that may not perfectly adhere to theoretical assumptions. The findings call for a cautious approach to interpreting model fits and a critical examination of data quality and assumptions in research. This case serves as a compelling example of the nuances involved in statistical analysis and the importance of adaptability and thoroughness in research methodologies

5.3 Validity

To add on, we discuss the internal validity and external validity. Internal validity is concerned with the degree to which a study can establish a causal relationship between its variables without external influence. As mentioned previously, our small data problem can affect the internal validity of the research in question. Since we have a small data set, there is a reduction in statistical power, the probability of correctly rejecting a false null hypothesis. Additionally, having a small number of observations increases the risk of overfitting. In this case, the model learns the noise in the data instead of the underlying pattern. External validity extends the concern to how well the study's findings can be generalized beyond the specific conditions, populations, and settings examined. The data has limited number of observations and is at the ward level in Toronto, so generalizing our findings to cities outside Toronto may raise external validity concerns. Toronto's unique socio-economic, cultural, and environmental characteristics may influence the study's variables in ways that are not replicable in other cities. Factors such as policies and economic conditions vary significantly from one city to another, potentially affecting the applicability of the findings elsewhere.

5.4 Further Steps

With the purpose of investigating the relationship between urbanization, cultural diversity, and wealth on public art in Toronto, we find a moderate positive relationship between population and number of public art installations by ward supported by data visualization as well as statistical models. The research finds no relationship between minority population or wealth on the number of public art pieces even after analysis with multiple models, suggesting that there may be stronger factors than minority population and income that affect the locations of public art in the city. Although a positive relationship is shown through the Poisson regression model, internal and external validity concerns also leave room for further improvement in the research. Violation in the Poisson model assumptions as well as a small data set calls for a larger data set in future research to address both internal and external validity concerns. With a larger number of observations, model inferences will improve, addressing internal validity. To expand the data set, we can consider gathering data from other cities. Through this, we are also able to address external validity. With the inclusion of a large data set with many cities, we can gather data with socio-economic diversity. When working with more diverse data, we can potentially generalize our findings to urban cities across the world, helping us learn more about the factors that affect public art in urban areas across the world. In conclusion, extending our research to cities across the world could improve the statistical models and validity of the research, enabling us to better understand the effect of urbanization on public art in urban cities.

Appendix

A Model details

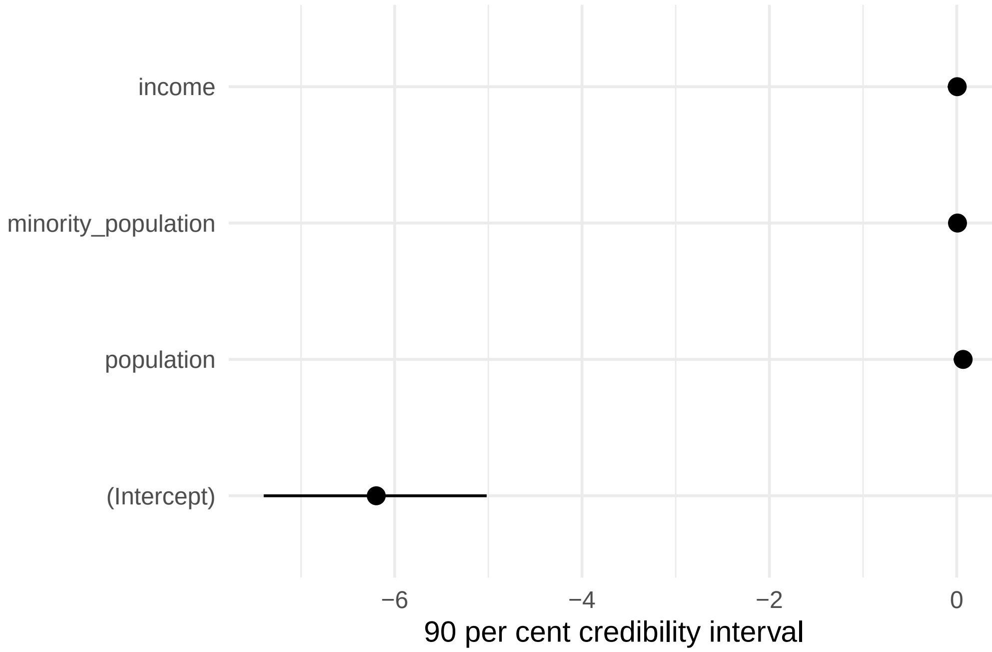


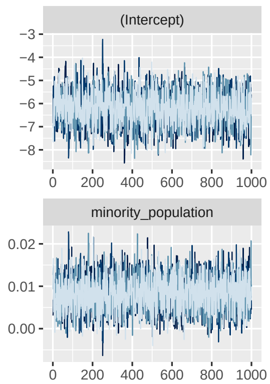
Figure 6: Explanatory models of public art works based on population

B Checking for issues

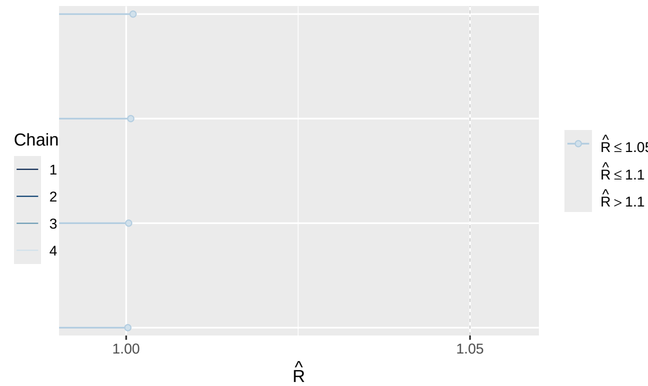
The Markov chain Monte Carlo sampling algorithm checks for signs that the algorithm has issues. We consider a trace plot Figure 7a, and a Rhat plot Figure 7b. In Figure 7a, we see horizontal lines that bounce around and have overlap between the chains. In Figure 7b, we see that everything is close to 1. We do not see anything out of the ordinary in the trace plot or Rhat plot, indicating that the algorithm did not run into any issues.

B.1 Posterior predictive check

In Figure 8a we implement a posterior predictive check on the multiple regression model. This shows the comparison between the actual outcome variable (public art installations) with simulations from the posterior distribution. From the figure, we can see that the observed data has a peaked distribution while the posterior predictive distributions are more dispersed.



(a) Trace plot



(b) Rhat plot

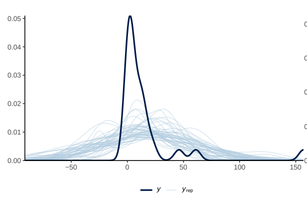
Figure 7: Checking the convergence of the MCMC algorithm

This means that the model is not a good fit and does not replicate the observed distribution well.

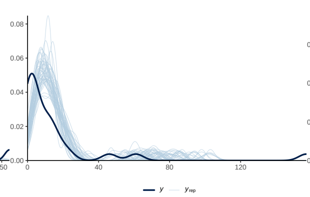
In Figure 8b we implement a posterior predictive check on the poisson regression model. This shows the comparison between the actual outcome variable (public art installations) with simulations from the posterior distribution. Here the observed data and posterior predictions have some overlap. This model is a better fit than the multiple regression model.

In Figure 8c we implement a posterior predictive check on the negative binomial regression model. This shows the comparison between the actual outcome variable (public art installations) with simulations from the posterior distribution. The observed data shows a peak near zero, but the predictive lines do not show any visible peaks despite being concentrated around zero. This does not capture the full range of observed data.

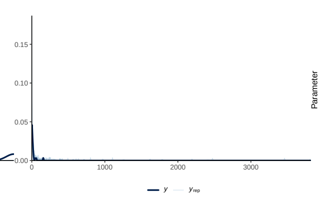
In Figure 8d we compare the posterior with the prior. This shows how much the estimates of the coefficients of our variables population, minority population, and income have changed once data was taken into account.



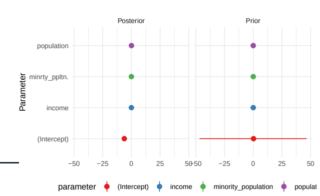
(a) Posterior prediction check



(b) Posterior prediction check



(c) Posterior prediction check



(d) Comparing the posterior with the prior

Figure 8: Examining how the model fits, and is affected by, the data

References

Arel-Bundock, Vincent. 2022. "modelsummary: Data and Model Summaries in R." *Journal of Statistical Software* 103 (1): 1–23. https://doi.org/10.18637/jss.v103.i01.

Barone, Jeanine. 2023. "Toronto's Vibrant Public Art Scene." https://www.forbes.com/sites/jeaninebarone/2023/08/24/torontos-vibrant-public-art-scene/?sh=ee40f0130aa3.

Bolker, Ben, and David Robinson. 2022. *Broom.mixed: Tidying Methods for Mixed Models*. https://CRAN.R-project.org/package=broom.mixed.

Gelfand, Sharla. 2022. *Opendatatoronto: Access the City of Toronto Open Data Portal*. https://sharlagelfand.github.io/opendatatoronto/.

Goodrich, Ben, Jonah Gabry, Imad Ali, and Sam Brilleman. 2024. "Rstanarm: Bayesian Applied Regression Modeling via Stan." https://mc-stan.org/rstanarm/.

Howarth, Thomas. n.d. "Cultural Life." *Encyclopædia Britannica*. Encyclopædia Britannica, inc. https://www.britannica.com/place/Toronto/Cultural-life.

Kay, Matthew. 2023. *tidybayes: Tidy Data and Geoms for Bayesian Models*. https://doi.org/10.5281/zenodo.1308151.

Liu, Chang. 2021. "Research on Urban Public Art Design Based on Digital Information Technology." *Journal of Physics: Conference Series*. https://doi.org/10.1088/1742-6596/1992/2/022081.

Lüdecke, Daniel, Mattan S. Ben-Shachar, Indrajeet Patil, and Dominique Makowski. 2020. "Extracting, Computing and Exploring the Parameters of Statistical Models Using R." *Journal of Open Source Software* 5 (53): 2445. https://doi.org/10.21105/joss.02445.

Meng, Yuan, and Meng Ye. 2022. "Analysis on the Integral Design of Public Art from the Perspective of Urban Culture." *Arts Studies and Criticism*, 125–28. https://doi.org/10.32629/asc.v3i2.838.

Müller, Kirill. 2020. *Here: A Simpler Way to Find Your Files*. https://CRAN.R-project.org/package=here.

Pebesma, Edzer, and Roger Bivand. 2023. *Spatial Data Science: With Applications in r*. Chapman; Hall/CRC. https://doi.org/10.1201/9780429459016.

Qizhi, Mao, Long Ying, and Wu Kang. 2016. "Spatio-Temporal Changes of Population Density and Urbanization Pattern in China (2000 – 2010)." *China City Planning Review V*. https://www.researchgate.net/profile/Kang-Wu-10/publication/311707586_Spatio-Temporal_Changes_of_Population_Density_and_Urbanization_Pattern_in_China_2000_-_2010/links/58d516f592851c44d4470dd3/Spatio-Temporal-Changes-of-Population-Density-and-Urbanization-Pattern-in-China-2000-2010.pdf.

R Core Team. 2022. *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing. https://www.R-project.org/.

Richardson, Neal, Ian Cook, Nic Crane, Dewey Dunnington, Romain François, Jonathan Keane, Dragoș Moldovan-Grünfeld, Jeroen Ooms, Jacob Wujciak-Jens, and Apache Arrow. 2023. *Arrow: Integration to 'Apache' 'Arrow'*. https://CRAN.R-project.org/package=arrow.

Salas, Rafael M. 1986. "Population and Urbanization." *A&Pacific Population Journal*. https://www.unescap.org/sites/default/d8files/knowledge-products/APPJ-Vol-1-No-1_0.pd f.

Schaeffer, Katherine. 2021. "'What's the Difference Between Income and Wealth?' And Other Common Questions about Economic Concepts." *Decoded*. Pew Research Center. https://www.pewresearch.org/decoded/2021/07/23/whats-the-difference-between-income-and-wealth-and-other-common-questions-about-economic-concepts/.

Slowikowski, Kamil. 2024. *Ggrepel: Automatically Position Non-Overlapping Text Labels with 'Ggplot2'*. https://ggrepel.slowkow.com/.

Smith, Christine. 2014. "Art as a Diagnostic: Assessing Social and Political Transformation Through Public Art in Cairo, Egypt." *Social & Cultural Geography*. https://www.tandfonline.com/doi/abs/10.1080/14649365.2014.936894.

Toronto, Open Data. 2023a. "Public Art." https://open.toronto.ca/dataset/public-art/.

———. 2023b. "Ward Profiles (25-Ward Model)." https://open.toronto.ca/dataset/ward-profiles-25-ward-model/.

Wickham, Hadley. 2016. *Ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag New York. https://ggplot2.tidyverse.org.

Wickham, Hadley, Mara Averick, Jennifer Bryan, Winston Chang, Lucy D'Agostino McGowan, Romain François, Garrett Grolemund, et al. 2019. "Welcome to the tidyverse." *Journal of Open Source Software* 4 (43): 1686. https://doi.org/10.21105/joss.01686.

Wickham, Hadley, and Jennifer Bryan. 2023. *Readxl: Read Excel Files*. https://readxl.tidyverse.org.

Wickham, Hadley, Romain François, Lionel Henry, Kirill Müller, and Davis Vaughan. 2023. *Dplyr: A Grammar of Data Manipulation*. https://dplyr.tidyverse.org.

Wright, Willie Jamaal, and Cameron Khalfani Herman. 2018. "No 'Blank Canvas': Public Art and Gentrification in Houston's Third Ward." *Special Issue: Urban Public Art: Geographies of Co-Production*. https://anthrosource.onlinelibrary.wiley.com/doi/10.1111/ciso.12156.

Xie, Yihui. 2023. *Knitr: A General-Purpose Package for Dynamic Report Generation in r*. https://yihui.org/knitr/.

Yin, Lingling, and Jia Chang. 2019. "Innovation of Urban Public Art Design in the New Era." https://webofproceedings.org/proceedings_series/ART2L/CLLA%202019/CLLA079.pdf.

Zhu, Hao. 2021. *kableExtra: Construct Complex Table with 'Kable' and Pipe Syntax*. http://haozhu233.github.io/kableExtra/.