

Women and Highly Educated Toronto Residents Most Likely to Be Concerned About Local Impacts of Climate Change*

A study on climate change perceptions with the City of Toronto's Open Data

Lorena Almaraz De La Garza

April 26th, 2021

Abstract

This report presents a secondary analysis of the City of Toronto's data on residents' perceptions about climate change. The results indicate that respondents' gender, education, and income, affect their self-reported level of concern, and level of information, about climate change. The analysis consists of logistic regression performed with the statistical programming language R. The results contribute to our understanding of the issue at a local level, enabling effective action to address the ongoing threat of climate change.

Keywords: climate change, climate perceptions, climate concerns, climate action, age and climate, gender and climate, toronto residents, open data toronto

Contents

1	Introduction	2
2	Data	3
3	Methodology	7
4	Results	8
5	Discussion	8
5.1	Findings	9
5.2	Ethics & Implications	11
5.3	Limitations & Future Work	12
6	Appendix 1	15
7	Appendix 2	17
	References	19

*Code and data are available in this GitHub repository: [lalmaraz/toronto_climate_perceptions](https://github.com/lalmaraz/toronto_climate_perceptions).

1 Introduction

Although scientific research on the impacts of climate change has been conducted for decades, the last few years have seen a surge in public participation and activism. In particular, youth climate leaders, with unwavering motivation and forthright campaigns, have guided and inspired thousands around the world (Unicef 2020). In their demands, they often ask that adults enact intergenerational equity, which states that every generation, present and future, “holds the Earth in common” (Oxford Public International Law 2013). They urge older generations to wake up from their profit-driven ignorance and comfort (Hirsi, Coleman, and Villaseñor 2019). But how does age really affect self-reported concern about climate change? To answer this question, the following pages present a secondary analysis of survey data collected by Environics, commissioned by the City of Toronto, on residents’ perceptions of climate change. Using the statistical programming language R, this study performs logistic regression to determine whether a correlation exists between two key variables: age and self-reported level of concern about climate change on a local and global scale. Informed by the youth climate movement, the hypothesis is that younger generations are more inclined to self-describe as concerned about climate change on both scales.

Drawing upon the activism that advocates for an intersectional approach to climate action, a multivariate regression analysis is also performed to include gender, education, income, and region. Intersectionality, as defined by Kimberle Crenshaw and described by Nina Lykke, is the idea that discrimination and systemic inequalities overlap – or intersect – across factors like gender, race, and class, and do not exist in isolation (Lykke 2010). Intersectional environmentalism, therefore, promotes climate action that recognizes how climate change disproportionately affects the already disadvantaged (Madson 2020). In this case, the hypothesis is that participants at the margins will have a greater likelihood of reporting a greater level of concern about climate change.

The results provide valuable insight. The models indicate that when it comes to whether a Toronto resident self-describes as concerned or not concerned about climate change as it affects Toronto and the Greater Toronto Area, gender, education, and income have a statistically significant effect. For reports on concern about climate change as it affects the rest of the world, only gender and education are important. Regarding whether Torontonians claim to be informed about climate change or not, the most important factor is education, unsurprisingly, but also gender. As argued by Jackson and De Coninck, effective climate action, particularly in adaptation and resilience, depends on active participation at a local level (Jackson and Coninck 2019). With this in mind, this paper advocates for the use of open data practices by local governments, including municipalities, to promote data-informed decision-making. For civic engagement and grassroots organizing, access to reliable data is vital. It allows organizers to direct outreach efforts and allocate precious resources efficiently. Due to the scale and complexity of the challenges posed by climate change, ensuring the public has access to quality data can prove immensely beneficial. As poignantly illustrated by Kieran Healy, the crowdsourced data and data analyses that continue to uncover the effects of COVID-19 demonstrate the power of collaboration and open data in addressing challenges of unforeseen scales (Healy 2020). Similarly, climate change - and effective climate action - have a lot to gain from public ingenuity fueled by open data.

The paper first presents an overview of the original dataset and the variables relevant to this study. It also provides critical commentary on the data collection methods and original survey design. In addition, it compares the surveyed sample to the provincial and national populations. Next, the Methodology section discusses the strengths and weaknesses of the implemented model, as well as the model that was considered as an alternative. It includes details on how the data were re-coded to suit the model and how non-response was addressed. Then, the Results section explains the baseline category for all models and an interpretation of the estimated coefficients from each analysis. Finally, the Discussion section is comprised of four subsections: Findings, Ethics & Implications, and Limitations & Future Work situating the study within a broader context.

2 Data

This report, including all necessary data cleaning, analysis and visualization, was produced with the R statistical programming language (R Core Team 2020) in the R Markdown file format. It uses features from several packages. The dataset is accessed through the *opendatatoronto* package (Gelfand 2020). The *tidyverse* package is used for data cleaning and manipulation (Wickham et al. 2019), as well as *janitor* (Firke 2021). Tables and graphs are generated with *kableExtra* (Zhu 2020), *extrafont* (Winston Chang 2014), *modelsummary* (Arel-Bundock 2021), *patchwork* (Pedersen 2020), and *ggplot2* (Wickham 2016). A map is created with *ggmap* (Kahle and Wickham 2013), *osmdata* (Padgham et al. 2017) and *sf* (Pebesma 2018). External data is accessed, imported, and exported with *RCurl* (Temple Lang 2021), *readxl* (Wickham and Bryan 2019), *httr* (Wickham 2020), *xml2* (Wickham, Hester, and Ooms 2020), and *rvest* (Wickham 2021). The models are assessed with *performance* (Lüdtke et al. 2021). To facilitate a reproducible workflow, *here* (Müller 2020) is used to reference file locations, and *bookdown* (Xie 2020) and *knitr* (Xie 2021) to format the report.

The dataset is publicly available through the Open Data Portal hosted by the City of Toronto, and as mentioned, it is accessed through the *opendatatoronto* package (Gelfand 2020). The data was last refreshed on August 15, 2019, and is listed in the catalogue as “Climate Perception Study,” published by the Environment & Energy Division. The data, as the catalogue states, was collected by Environics Research, a private Canadian polling and market research firm based in Toronto. The firm was commissioned by the City of Toronto to conduct an online survey with a representative sample of Toronto residents (City of Toronto 2021b). According to the source, the survey was designed to “capture perceptions about climate change” to “help the City establish goals, priorities and metrics that will measure success in [climate action]” (City of Toronto 2021b). The catalogue includes a report with results from data analysis prepared by Environics under the section “More information.” This report, titled “City-Wide Climate Perceptions Study” (Environics Research 2018a) is referenced throughout this paper to compare results.

The data is provided in two versions in the Open Data Portal, a raw version (V1) and a secondary version coded for quantitative analysis (V2). For the purposes of this paper, and to code the data to this paper’s specific aims, only V1 data is used.

The survey was conducted in 2018, between October 11 and 18, with a total of 404 adult residents of Toronto answering the online questionnaire (City of Toronto 2021b). The study followed a purposive sampling method, having specific predetermined groups of interest, residents aged 18 years and over, and implementing screening questions at the beginning of the questionnaire. As the source claims, sample quotas were set by region, age, and gender, in keeping with federal 2016 Census data (City of Toronto 2021b), presumably following a proportional approach. The study description on the Open Data portal lists that the survey was conducted on a sample of convenience, indicating that the survey respondents were not randomly selected and invited to participate based on systematic methods, but rather based on proximity, availability, or ease. The specific strategies used to recruit participants are not explained. Additionally, the source does not disclose why sampling was limited to adult residents.

The survey collected a wide range of data, including demographic and general information like postal code, gender, age, language of preference, income level, education level, type of residence, and homeownership. In addition, the survey also asked respondents attitudinal and behavioural questions including their level of concern about climate change as it affects Toronto and the Greater Toronto Area, level of concern of climate change as it affects the rest of the world, level of information about climate change, perceived impacts of climate change in the community, awareness of climate action programs nationally and locally, and believes, attitudes, and behaviours towards climate change and climate action. The dataset spans 152 columns which include, in addition to the participants’ direct responses, systemic thematic coding for open-ended questions and all possible responses for question loops. A preview of a few columns and rows from the original dataset as provided by the source can be found in Table 1.

From the dataset, this paper focuses on the following: (1) postal code, a nominal variable; (2) age, measured at the ordinal level; (3) gender, nominal; (4) level of concern of the local impact of climate change, (5) level of concern of the global impact of climate change, (6) level of information about climate change, (7)

Table 1: A preview of the Climate Perceptions Study dataset

QS1	QS2	HIDAGE1	Q1r1	Q1r2	Q2	Q3
m4s	Man	38	Very concerned	Very concerned	Very informed	
m1w	Woman	46	Extremely concerned	Extremely concerned	Extremely informed	yes
M9R	Man	76	Extremely concerned	Very concerned	Not very informed	Changes in weather
M9w	Woman	41	Extremely concerned	Extremely concerned	Very informed	Not sure
M5A	Man	67	Very concerned	Very concerned	Very informed	Weather volatility, flooding from Lake Ontario high water,
M5N	Woman	70	Very concerned	Extremely concerned	Very informed	Air quality is poor Dirt and dust are higher Temperature is extreme
m6m	Woman	69	Very concerned	Very concerned	Very informed	dying front lawn tree's on main streets

education level, and (8) income level, all ordinal variables. Table 2 shows a summary of key demographic features subdivided into a binary, “yes” or “no,” for each of the three dependent variables: concerned about climate change as it affects Toronto and the Greater Toronto Area (local), concerned about climate change as it affects the rest of the world (global), and informed about climate change. A list of the questions used to collect this information, as included in the original questionnaire (Environics Research 2018b), can be found in Appendix 1.

The original report separated responses based on the participants’ district of residence. Similarly, using the first three digits of respondents’ postal code and corresponding approximate latitude and longitude, a map with the approximate locations of respondents can be found in Figure 1.

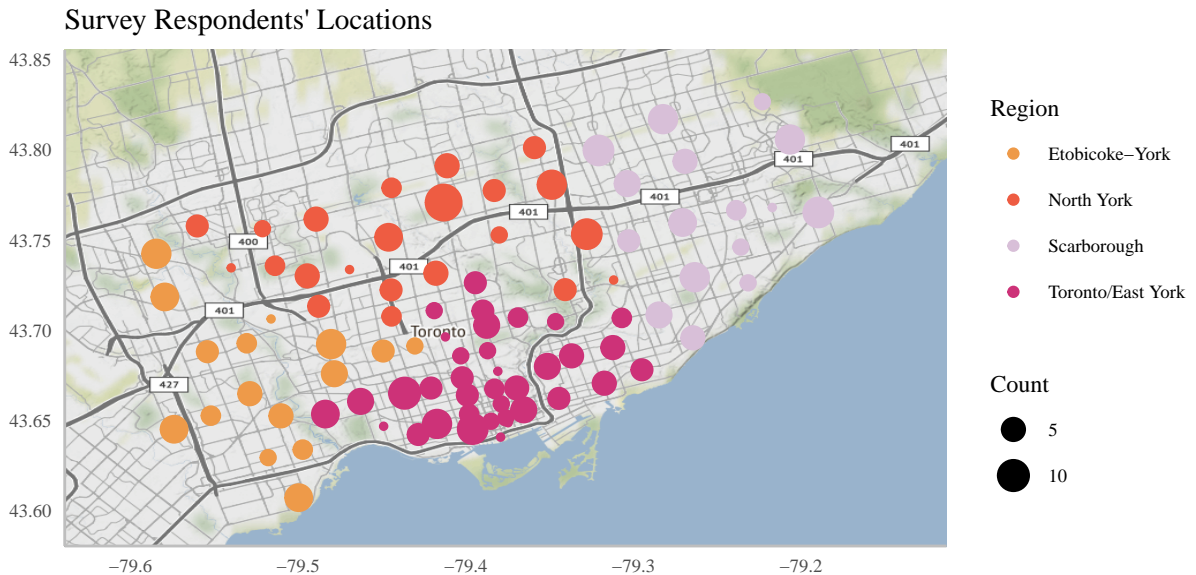


Figure 1: A map showing the approximate location of survey respondents

It is important to acknowledge the strengths and weaknesses of the original study. The source does not state how many people were approached to answer the survey. Therefore, it is impossible to accurately estimate the impact of unit non-response bias in the original study. However, it is safe to assume that the original study took deliberate measures to ensure the mitigation of item non-response. All 404 observations list values for all columns in the dataset, with the exception of question loops that required specific conditions to be asked. Assuming the dataset is presented in its original form or as close to it as possible, this response rate is commendable, particularly for an online survey that asks about behaviours and attitudes. As Marlar explains, “[f]or most attitudinal questions, respondents have never considered how they feel about a topic until they are asked” (Marlar 2018), suggesting that methodological designs that allocate time for reflection might lead to higher response rates. Non-response bias can be addressed with preventative steps, including providing incentives for participation (Turk, Heneghan, and Nunan 2019), yet the steps taken in the study are unclear. As stated in the original questionnaire instructions, there were “only a few key questions where

Table 2: Summary of variables of interest by outcome

		(N = 401*)		(N = 401*)		(N = 404)	
		Concerned: Local		Concerned: Global		Informed	
		No	Yes	No	Yes	No	Yes
Age	Mean	45.87	47.36	49.02	46.89	49.28	46.36
Gender	Female	26	167	18	175	65	129
	Male	57	148	40	165	47	160
	Transgender	0	1	0	1	0	1
	Prefer not to say	1	1	2	0	0	2
Education	High school or less	11	24	8	26	17	18
	Some college, vocational, trade school	3	17	3	17	10	10
	Completed college, vocational, trade school	12	36	8	40	20	28
	Some university	9	32	6	35	13	29
	Completed undergraduate degree	28	100	18	110	32	97
	Post graduate/professional school	16	103	12	107	18	101
	Prefer not to answer	5	5	5	6	2	9
Income	Under \$40,000	6	39	4	41	18	28
	\$40,001 to \$60,000	7	40	4	43	12	35
	\$60,001 to \$80,000	15	48	13	50	24	40
	\$80,001 to \$100,000	6	49	3	52	16	39
	\$100,001 to \$150,000	11	51	8	54	13	49
	More than \$150,000	15	41	10	46	9	47
	Prefer not to answer	24	49	18	55	20	54

* Three responses ["Don't know"] have been removed

[a participant's] answer [was] required" (Environics Research 2018b), indicating that not all responses were mandatory and instead, participants felt compelled to volunteer an answer.

The survey did include options for participants to withhold information if preferred. For the purposes of this paper, "Prefer not to say," "Prefer not to answer" and "Don't know" were considered invalid responses. A total of 78 responses were removed, a significant number that accounts for almost 20% of the sample. The majority of these, 74 in total, were removed due to participants' refusal to report their income level. Because the models in this paper consider income as a predictor variable, it was necessary to remove the responses, sacrificing valuable information from other variables. While the trade-off is not ideal, it provides the model with stability when including the income variable. Ideally, the original study would have accounted for this type of item non-response in the original survey and study design. Non-response in income-related questions is a common issue, as explained by Moore et al. (Moore, Stinson, and Welniak-Jr. 1997). They note this is due to several potential contributors, including cognitive demand. As they explain, participants may refuse to provide their income because they might need definitional precision or clarification on whether it should include or exclude taxes, government subsidies or other financial support, income generated from assets, etcetera (Moore, Stinson, and Welniak-Jr. 1997). Additionally, respondents might not recall an amount or the amounts that comprise their total income and may find the mental process of retrieval too taxing for a quick survey. Finally, Moore et al. point at the issue of sensitivity - the taboo of speaking about money publicly (Moore, Stinson, and Welniak-Jr. 1997). Although not mentioned by Moore et al., growing concerns around data privacy may lead to respondents hesitating on this particular question. Moore et al. recommend several measures in study design to account for this error, including providing a clear definition, specifying a time period, and most importantly, reiterating that confidentiality will be respected throughout the entirety of the study, as well as explaining succinctly but transparently, why exactly this information is necessary (Moore, Stinson, and Welniak-Jr. 1997). While the relevant question in the original questionnaire implements all these considerations to some extent, perhaps it would have been beneficial to emphasize for participants the importance of providing an answer. The original study does not indicate how this type of non-response was dealt with when analyzing the data.

Additionally, the source indicates that sample quotas were established following Census 2016 information from Statistics Canada. However, a comparison between the survey sample and the population of Toronto and Ontario shows some discrepancies. Most salient are the over-representation of participants who identified as male in the 35 to 49 age category, the under-representation of female-identifying participants in the same age group, and the under-representation of participants aged 65 and over of both genders. Because of this, insights gathered from the survey should be considered only for pilot programs or as a starting-point for future, more robust research, with a larger sample that appropriately fits the city’s demographics. It is worth noting as well that, as compared to Ontario, the population of Toronto decreases as age increases, indicating that measures or programs designed for Torontonians might not necessarily be generalizable to the province. A graph comparing the age group representation by gender in the survey sample, the city of Toronto, and the province of Ontario is found in Figure 2.

Convenience sampling, while time- and cost-efficient, is not an appropriate method for the motivations of the original study. As the source states, the study results will “help the City establish goals, priorities and metrics” (City of Toronto 2021b). As indicated in Figure 2, the study sample did not accurately represent the city population. To ensure the study and results have external validity, it is imperative to follow a random selection approach, ensuring all individuals in the population of interest have similar odds of being included in the study. In this case, cluster random sampling would have accounted for the nuance unique to Toronto neighbourhoods and surrounding areas, notably segmented across important demographic features like age, race, cultural background, and wealth.

The observed effects might be exaggerated or attenuated due to selection bias. Individuals that responded to the call to participate and completed the survey might be systematically different from the broader Toronto population. Participants with little interest in climate action, or skeptical or uninformed about climate change, for example, might have decided not to participate or complete the survey. Proper randomization and participation incentives are measures to mitigate this effect.

A final consideration is social-desirability or conformity bias, particularly in contentious topics, as respondents might feel compelled to adapt their responses to fit societal norms rather than report the truth or their honest opinion. Behavioural and attitudinal questions are especially prone to this tendency, and designing the study to collect more than one reference point, perhaps asking more than one version of a similar question to then triangulate data, might be beneficial.

The kinds of measurement error and biases stated above, in addition to the lack of clarity on how the original study dealt with these in the design, collection, analysis, and reporting, compromise reliability. Because of this, results from both the original study and this secondary analysis should be considered an initial step towards a more comprehensive study, conducted with careful sampling and meticulous research design.

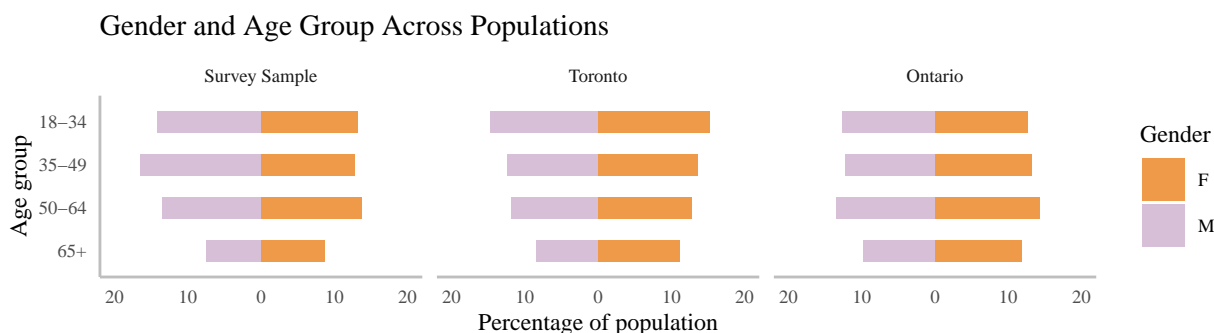


Figure 2: A graph comparing the survey sample to the city and province populations

3 Methodology

Because this study aims to find whether age has an effect on self-reported concern of climate change or not, a binary outcome, the model it relies upon is logistic regression. The model assumes that:

$$P(Y = 1|X_k) = \frac{\exp(\beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \beta_4 X_4)}{1 + \exp(\beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \beta_4 X_4)}$$

The model estimates the probability P of the outcome variable Y to be successful given the predictor variable(s) X_k . In this case, it elucidates the relationship between the dependent variable, level of concern, and the independent variables age, gender, education, income, and region estimating the parameter values for β via maximum likelihood.

Binomial logistic regression was selected due to its efficiency and popularity in assessing dichotomous outcomes, providing measures of appropriateness for each predictor as coefficients and indicating the direction of each association, either positive or negative. The model clearly indicates the importance of each predictor as compared to the others. However, the model assumes linear boundaries across variables – an assumption that might not as easily translate into the complexity of empirical evidence. Because of this, it is imperative to carefully consider, and include, the variables that might significantly affect the outcome. In this case, relevant demographic information is deliberately included. Logistic regression is also highly sensitive to anomalies in data. Therefore, 78 instances of invalid responses were omitted, as described in the Data section. Additionally - and not without careful consideration - the gender variable was reduced to a binary. Because a single observation leads to high instability in the model, “transgender” was removed, narrowing the observations used for each model to 325 from the original sample of 404.

Alternatively, an ordinal logistic regression model was considered as data on the level of concern of climate change (local and global) and the level of information about climate change were collected on a Likert scale. However, they were re-coded to binary variables, and a binomial logistic regression was performed, to create groups with counts as comparable as possible. For instance, while only two participants claimed to be “Not at all informed,” 89 self-described as “Not very informed.” Aggregating these two responses into a broader category, “Not informed” (and “Extremely informed” and “Very informed” into “Informed”) provided higher stability. To fit the model, all three ordinal dependent variables of interest were re-coded into binaries by evenly splitting the 4-point Likert scale.

An additional variable “region” was created for the models. The original report included a map of respondents’ approximate locations divided into four main regions: Toronto and East York, Etobicoke-York, North York, and Scarborough. Based on the first three digits of the postal codes as listed in the original dataset, the variable “region” was created following this division. Postal code information, as well as longitude and latitude lists, were primarily sourced from GeoNames (GeoNames 2019).¹ Bar charts of all independent variables arranged by each of the dependent variables can be found in Appendix 2. They contain all 325 observations as used in each of the models.

As often stated, all models are wrong but some are useful. As noted in the Data section, the original study design and resulting data have some limitations that even a perfectly calibrated model cannot counteract. Particularly, models operating on observational data might lead to a false level of confidence in drawing conclusions or making claims, as it is easy to fail to recognize the myriad decisions required at every step of the study design, data collection, and analysis. Each of these decisions is an opportunity for bias and human error to be introduced in the study, tainting the results. With this in mind, results should be interpreted as very small insights into the data as it fits the model, and not as unquestionable truths about the world from which the data was collected. Although reproducibility and replicability are rare in social sciences research, a practice of transparency, accountability, and openness can ensure, at the very least, that the results be critically interpreted, disputed when invalid, and augmented or confirmed with further research.

¹More information about how regions were created, including information on coordinates and region subdivisions, can be found in the script *map_regions.R*.

4 Results

The reference group for all models is male, aged 18-34, with an education level of high school or less, a household income of under \$40,000, and living in the Toronto/East York region. Because the model does not discriminate between ordinal or nominal data, as all independent variables were recoded as dummy variables, the results indicate whether a change in one of the conditions affects the outcomes of interest. For age, education, and income, the coefficient estimate indicates the effect of a one-unit increase in these groups - from “some university” to “completed undergraduate degree,” for example. For gender and region, however, the estimated coefficients simply indicate the effect between identifying as male or female or living in one region versus another.

The coefficients in Table 3, which presents a summary of the three models, are exponentiated and indicate the odds ratio rather than log odds. According to the first model, gender, education, and income have a statistically significant effect on the likelihood of participants self-describing as concerned about climate change as it affects Toronto and the Greater Toronto Area. When it comes to concern about climate change as it affects the rest of the world, only gender and education are important. Similarly, education and gender are statistically significant as it pertains to whether participants self-describe as informed about climate change or not.

According to the first model, as compared to males, females have 3.37 times higher odds of self-describing as concerned about the local impacts of climate change, keeping other variables constant. An increase in education, based on the pre-established education group categories, leads to 1.30 times higher odds of self-describing as concerned about the local impacts of climate change, keeping other variables constant. An increase in the income group, as per the established categories, is associated with 0.84 times higher odds of self-describing as concerned about the local impacts of climate change, keeping other variables constant.

The second model indicates that as compared to males, females have 3.43 times higher odds of self-describing as concerned about the global impacts of climate change, keeping other variables constant. An increase in education, following the education groups determined in the study, leads to 1.25 times higher odds of self-describing as concerned about the global impacts of climate change, keeping other variables constant.

Finally, the third model shows that females, as compared to males, have 0.64 times higher odds of self-describing as informed about climate change. An increase in education, based on the established education groups, leads to 1.36 times higher odds of self-describing as informed about climate change, keeping other variables constant.

This paper sought to answer two research questions: (1) how is age related to concern and knowledge of climate change?, and (2) how do intersecting identities (gender, class) affect concern and knowledge of climate change?

The hypotheses for this study were: (1) in keeping with some of the popular rhetoric in the youth climate movement, age is negatively correlated with self-reported concern and knowledge about climate change. And (2), informed by an intersectional environmentalism approach to climate action, systemically oppressed peoples (non-hegemonic gender, low income, and low education) are more likely to be concerned about climate change.

Based on the results from three multivariate logistic regressions, this study fails to reject the null hypothesis that age does not affect the level of concern and information about climate change. Conversely, the study rejects the null hypothesis that gender, income, and education do not affect the level of concern and information about climate change.

5 Discussion

The results provide a small window into Torontonians’ perceptions about climate change. To complete the picture, it is necessary to assess these results within a broader context. The following sections provide further detail on the relevance of the results and are subdivided into Findings, Ethics & Implications, and

Table 3: Summary table of three multivariate models

	Concerned: Local	Concerned: Global	Informed
	OR	OR	OR
Age	1.169 (0.148)	0.896 (0.168)	0.876 (0.124)
Gender	3.368*** (0.335)	3.433*** (0.402)	0.637* (0.262)
Education	1.296** (0.105)	1.255* (0.119)	1.364*** (0.088)
Income	0.839* (0.100)	0.840 (0.118)	1.081 (0.086)
Region	0.977 (0.131)	1.020 (0.152)	0.890 (0.115)
Num.Obs.	325	325	325
AIC	298.8	241.7	367.2
BIC	321.5	264.4	389.9
Log.Lik.	-143.409	-114.859	-177.578

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Limitations & Future Work. The first section focuses on the predictors that proved to be most significant when it comes to concern and information about climate change: gender, education, and income. In Ethics & Implications, a brief overview of the moral principles implicated in both the original study and this secondary analysis are discussed, as well as the consequences and real-world impacts of the insights gleaned from the study. The Limitations & Future Work section, similar to the critique in the Data section, lists further considerations that the main study might have missed, as well as shortcomings from this secondary analysis, providing some suggestions for potential research to expand upon this study.

5.1 Findings

Although age does not have a statistically significant effect on level of concern and information about climate change, it is worth mentioning that according to the three models, there is a very slight increase in the likelihood of self-describing as concerned and informed as age increases. This may help inform the rhetoric of the youth climate movement locally.

Also, while region does not have a statistically significant effect either, it is important to consider what living in one area of Toronto versus another means, and why a finer-grained look at neighbourhoods instead of amalgamated regions might be necessary for future research. Toronto, one of the most multicultural cities in the world, is home to an incredibly diverse population that is not evenly distributed across geography but rather seems to cluster by cultural and ethnic background in localized areas (Clark 2013). With more than 50% of residents belonging to a visible minority group (Statistics Canada 2019), understanding how each community perceives the threat of climate change, and how informed they are, will be imperative for developing programs and measures.

5.1.1 Gender

When it comes to climate change perceptions, gender matters. Figure 3 shows how reports on concern about climate change differ for men and women in Toronto. Gender has an important effect on whether an individual is concerned or not about climate change as it affects both Toronto and the rest of the world, with female-identifying Torontonians having a higher likelihood of self-describing as concerned. Pearse’s review

of the literature on gender and climate illustrates how “gender relations are an integral feature of social transformations associated with climate change” (Pearse 2016), warning against gender-blind social research on the topic.

Gender also matters in how informed people are about climate change. In Toronto, female-identifying people are more likely to be informed. As stated by Arora-Jonsson, women across the world are and will continue to experience a higher negative impact from the effects of climate change (SeemaArora-Jonsson 2011). For women, knowledge about responses to climate change, both mitigation and adaptation, is vital. Arora-Jonsson advocates for a feminist response to climate change that carefully questions gender polarities (SeemaArora-Jonsson 2011).

For establishing goals, priorities and metrics for success - the motivation for the original study - gender has to be considered. Programs that target female-identifying people and address the compounded negative impacts of climate change, including the economic, health-related, and psychological harms, will become increasingly necessary. Additionally, programs designed to increase the understanding of climate issues for men will also be worthwhile. Establishing gender-aware priorities towards not only informing but also protecting and supporting Torontonians is crucial.

Gender and Concern About Climate Change

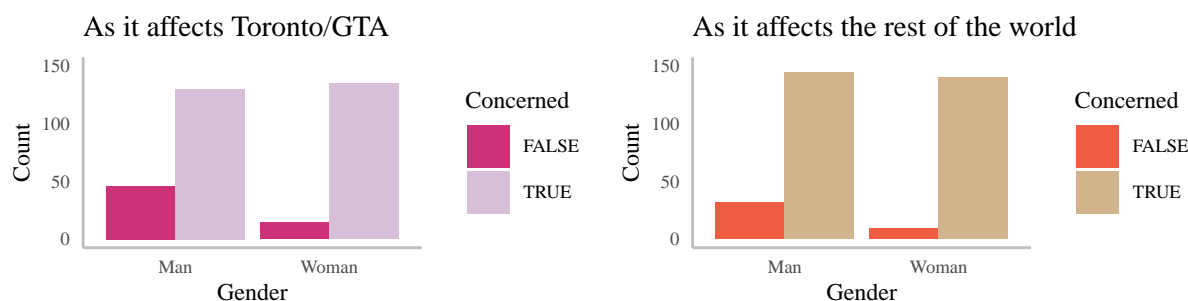


Figure 3: Histograms of Torontonians’ gender and concern about climate change.

5.1.2 Education

Unsurprisingly, an individual’s level of education has an effect on how informed they are about climate change. The higher level of education achieved, the greater likelihood to self-describe as informed about climate change. This means that the city should consider educational campaigns and interventions that target residents with lower education levels, providing avenues for them to stay informed.

Level of education also affects whether a person describes as concerned or not about climate change. Education proved to be particularly important for expressing concern about the local impacts of climate change, but it also influences the level of concern about global impacts. Figure 4 shows participants’ incomes and whether they describe as concerned about climate change. Education serves as a social equalizer providing residents with the ability to make informed decisions. For climate action, while the City of Toronto’s current interventions through TransformTO emphasize individual action, education is imperative for residents to go beyond and to participate in regulation and policy, demanding accountability from government and industry.

5.1.3 Income

Interestingly, level of income has an effect on whether participants self-describe as concerned about climate change as it affects Toronto and the Greater Toronto Area. Wealthier groups are more likely to express concern. Further research is required to understand why this is the case, but it seems like providing relevant information on climate change and its impacts, particularly as they affect those with low incomes, might be beneficial. It is well established that around the world, the poorest people are the most impacted (Islam and Winkel 2017). This can help inform response on a local level.

Education and Concern About Climate Change

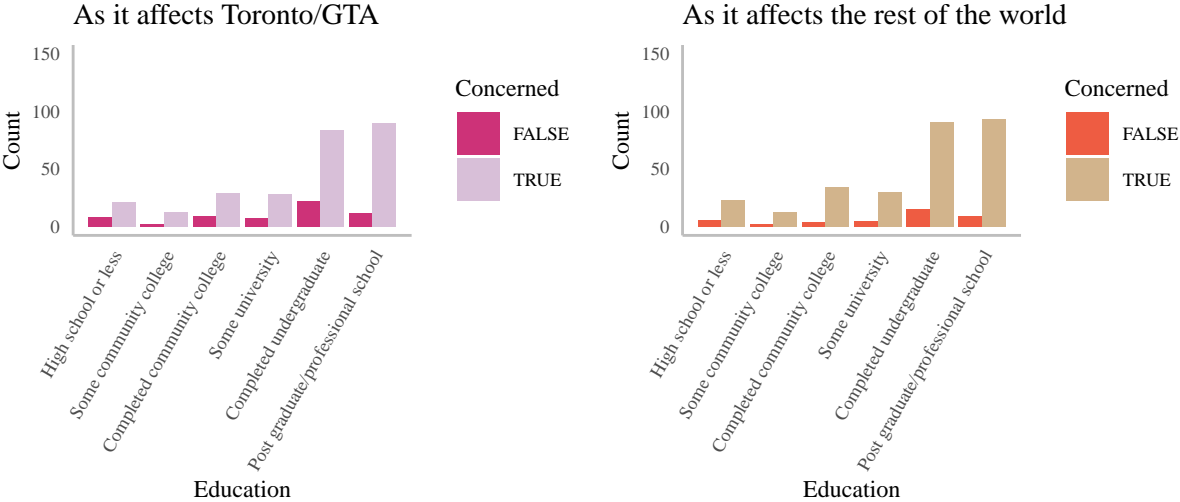


Figure 4: Histograms of Torontonians' level of education and concern about climate change.

5.2 Ethics & Implications

Concerning the original study, and acknowledging that the published questionnaire might have been different from the one implemented, there are key items to consider in relation to ethics. First, the original survey introduction does not explain whether anonymity was ensured nor how. It does not mention confidentiality either, only the final question does, asking about level of income (included in Appendix 1). Additionally, it is not clear whether data protection measures were taken nor if participants were informed about the intent to publish their responses as a publicly accessible dataset through the Open Data Portal. However, the questionnaire indicates that name, address, email address, phone number, or other personally identifiable information was not collected. Because the survey did collect quasi-identifiers or linkable information like year of birth, level of education and income, postal code, number of residents in the household, and number of children living in the household, participants should have been informed about the measures taken to protect their privacy, a key feature to informed consent. The purpose of the survey and the reiteration of voluntary participation is not evident in the survey introduction either. Ethics standards are paramount to any study involving human participants, and appropriate disclosure and assurance would have strengthened the study overall.

Based on the limited sample and sampling method, which led to the over- and under-representation of certain groups, and other limitations mentioned in the Data section, the author would advise caution in developing city-wide interventions. Rather, the results should be referenced as a point of departure for future studies, employing robust methods, or for small-scale pilot programs. It is unclear whether the information from the original study has already informed the many climate change-related programs the City of Toronto runs, but improved methods and design would lead to a more comprehensive understanding and consequently, better-informed initiatives.

The TransformTO strategy stands to gain significantly from expanded research. TransformTO “includes short and long-term actions to meet the City’s greenhouse gas reduction targets of 65 per cent by 2030, and net zero by 2050, or sooner” (City of Toronto 2021c). It includes several programs and initiatives, promoting “4 Big Steps” to fight climate change: reducing energy, reducing transportation footprint, minimizing waste, and getting involved in community activities. The original study and survey are included as part of TransformTO (City of Toronto 2021c).

From the programs that the City currently operates, the results can inform a few in particular. Currently, the Women4Climate Mentorship Programme through Live Green Toronto aims to “contribute to the next generation of climate leaders through a mentorship programme dedicated to women who are working to

address the climate crisis” (City of Toronto 2021d). As the results from this study show, gender plays an important role in the level of concern and level of information about climate change both locally and globally. Plans to expand the program, as it currently only invites 12 women to participate, should consider results from this and similar studies.

Additionally, the Climate Action Fund “supports community-led projects, activities and events that directly or indirectly reduce the harmful emissions that contribute to climate change” (City of Toronto 2021a). As the information site explains, funding is available on a geographic or population-level cluster basis (City of Toronto 2021a). The initiative would benefit from robust studies that take a closer look at the neighbourhood-level intricacies of Torontonians’ perceptions of climate change, rather than aggregated regions.

Finally, the Neighbourhood Climate Action Champions program, enabling local leaders to inspire climate action, has eligibility criteria based on age among other factors (City of Toronto 2020). It prioritizes applications from people of colour, newcomers, LGBTQ2S+ persons, women, Indigenous peoples, and persons with disabilities / disabled (City of Toronto 2020). A deeper understanding of intersectionality as it relates to climate change, based on evidence, would inform the program criteria and provide further specificity.

5.3 Limitations & Future Work

As noted in the Data section, the original study had some limitations that directly impact the data. A fundamental flaw is the treatment of gender as three categories, which inevitably was reduced to a binary in this paper. The author’s conviction is that this binary is artificial not only in data but in society, and in the context of this paper it is merely utilitarian, elucidating whether differences in self-reported gender affect the outcomes of interest. With this, it is absolutely imperative to acknowledge that while the original survey appropriately included “transgender” as an option, only one individual self-identified. Although the questionnaire might have been inclusive of transgender people, not including comprehensive options for gender is problematic. Additionally, collecting a single data point from a population that continues to face discrimination that in many cases stems directly from under-representation is unacceptable. A robust study design, particularly one that takes place in Toronto - arguably home to Canada’s largest transgender, gender-fluid, nonbinary, and gender non-conforming population - would deliberately collect their perspectives with a representative sample.

Additionally, the survey failed to incorporate key factors known to influence an individual’s perspectives on climate change, which impact construct validity. Three specific items come to mind: political affiliation, field of knowledge, and race. Work by Berbeco and Hamilton states that political affiliation seems to be a major factor in whether people accept the scientific evidence of climate change (Berbeco and Hamilton 2013). Understanding how political affiliation may be influencing perspectives in Toronto is fundamental to determine how leaders should communicate to their constituents to encourage climate action towards achieving TransformTO goals.

As the results revealed, level of education has an effect on the level of concern and the level of information about climate change, with highly educated individuals self-describing as more concerned and more informed. However, the study did not inquire about specific fields of study. As seminal research by O’Riordan states, environmental attitudes are dependent on an individual’s level of knowledge (O’Riordan 1971). Work by Srichai found that significant differences in perceptions and attitudes towards environmental issues stem from the field of study in secondary education (Srichai 1988). Given the high percentage of university-educated residents of Toronto, learning about their field of study would better inform climate action programs.

With the ongoing fight against race-based discrimination, it is necessary to understand how visible minorities are affected by climate change to ensure their concerns are directly addressed by proposed interventions. In November 2018, just a few months after the original study was conducted, the City of Toronto published a report on Indigenous Strategic Discussions on City of Toronto Climate Strategies. The report acknowledges the fundamental role Indigenous communities play in the development of a climate just future (City of Toronto 2018). The discussants agreed that “a more in-depth consultation process would be necessary to support a more formal set of recommendations or action items (let alone policies) in ways that represented meaningful regional Indigenous consultation on these important matters” (City of Toronto 2018). With this

in mind, the deliberate inclusion of the race variable in any future studies would help elucidate how the impacts of climate change are distributed across Toronto’s peoples. The work of the Indigenous-led Environmental Data Justice Lab looks at the relationships between data, pollution, and colonialism, developing initiatives that are aware of the uneven negative effects of environmental racism (Technoscience Research Unit 2021). Environmental racism, as defined by Chavis and quoted by Beech, describes the “racial discrimination in environmental policy-making, the enforcement of regulations and laws, [the deliberate targeting of communities of colour for toxic waste facilities, the official sanctioning of the life-threatening presence of poisons and pollutants in our communities [. . .], and the history of excluding people of colour from leadership of the ecology movements” (Beech 2020). Future research that accounts for this form of systemic oppression is paramount.

As evidenced by their participation in hundreds of strikes and actions around the globe, youth have played an instrumental role in the climate movement, and understandably so. After all, their future depends on effective mitigation and adaptation strategies. The decision to exclude people under 18 for the study is shortsighted at best, irresponsible at worst. While conducting research with minors requires an additional layer of scrutiny and ethical considerations, including ensuring parental consent, this particular topic certainly merits the effort. Neither the report nor the source disclose why youth were excluded. In Toronto, with a younger population compared to the province, understanding youth’s perceptions of climate change is absolutely necessary. As research by Onur et al. found, elementary school students express high levels of concern about the environment (Onur, Sahin, and Tekkaya 2012). The researchers also found that gender has an impact even at such a young age, suggesting that girls were significantly more concerned than boys (Onur, Sahin, and Tekkaya 2012). Future research should ensure their voices are heard, intentionally enacting intergenerational equity.

As briefly acknowledged in the Methodology section, a statistical model operating on observational data cannot make predictions with absolute accuracy. However, careful experimental research design can provide appropriate bases for discussing causality even just with observational data. While the data gathered from the original study is insufficient to do so, modifications to the study design would have allowed for interpretation beyond correlation.

Importantly, a very deliberate assessment of potential confounding should take place. Developing a visual representation of the variable interactions that might be at play, through a directed acyclic graph, can be beneficial. In this case, the study would have benefited from including the variables that are proven to affect perceptions of climate change, as earlier described. Additionally, the sampling method would have had to allow for an accurate representation of the population to accurately estimate a counterfactual. Because the population of interest is residents of Toronto but the sample included a higher proportion of males in the 35 to 49 category, causality statements would be compromised. As mentioned in the Data section, systematic sampling and providing incentives for participation would have helped mitigate selection bias. Measurement bias, as well, limited the possibilities of discussing causality. Although there is no missing data in the dataset, in the sense that all observations contain values for all columns, it is difficult to assess whether participants provided honest answers or whether they were victims of social-desirability bias. A survey that accounted for this, particularly regarding income and attitudinal questions, by providing more than a single opportunity for individuals to share their answers would have been valuable. A thoroughly tested survey, and one that is delivered through the most appropriate means based on expected results rather than convenience, is fundamental for establishing causation. Lastly, there is the risk of Berkson’s paradox occurring when looking only at Toronto data. Berkson’s paradox refers to a perceived relationship that arises from a narrow look at data and can be addressed by collecting data from a larger population group (Berkson 1946). While region might seem unimportant when analyzing Toronto data, for example, this might not be the case when analyzing provincial data, as we may find that highly urban regions differ greatly from rural and suburban regions and that this distinction does have an effect on climate perceptions. With a large enough sample, and having data on all relevant variables, propensity score matching would have been possible. Propensity score matching “involves assigning some probability to each observation” (Alexander 2021). In other words, it uses each individual’s propensity to receive a treatment when control and treatment groups have not been randomly assigned, and matches individuals based on similarities. In this case, individuals would have been matched on age, gender, income, and education to try to uncover whether the place where they live, hypothetically, affects their level of concern about climate change. We could then find, for example, that a

highly educated male with a high income has a higher propensity to describe as concerned about climate change, compared to someone with the opposite characteristics, illustrating that the region itself might not affect. However, this approach requires including all variables that might be having an effect, and as stated before, the survey did not account for those.

Regardless, the results provide further information on how future studies can be operationalized. More broadly, the results continue to expand our understanding of a challenging issue of an immensely complex nature through the use of open data. Research that embraces open science and open data practices is fundamental to accelerate the pace of problem discovery, empower groups and individuals, and foster collaboration towards actionable solutions (Vasa and Tamilselvam 2014). This is especially true for climate change research. Effective action, as Jackson and De Coninck emphasize, depends in part on active participation at a local level (Jackson and Coninck 2019), led by passionate citizens and not just governments, academic institutions, or charitable organizations. Open data is vital for civic engagement and grassroots organizing, as it allows organizers to make data-informed choices when determining their outreach efforts, resource allocation, program design, and other strategic decisions. The hope is that the results listed in this paper, and the general practice of producing the paper itself, will motivate others to engage in data analysis with the use of open data², free software, and the wealth of information available online.

²The author would like to kindly acknowledge Megan MacLean, Project Lead of Outreach & Engagement at the Environment & Energy Division of the City of Toronto and listed data owner for the dataset used in this study, for their support in providing additional resources and promptly addressing any difficulties.

6 Appendix 1

Excerpt of relevant questions from original survey³

- **S1.** Please enter the six digits of your home postal code below. This is to ensure that the survey is representative of respondents from across the city. (Enter postal code below)
- **S2.** Your gender: (Select one option only)
 - 01 – Woman
 - 02 – Man
 - 03 – Transgender
 - 04 – Prefer not to say
 - 05 – Not listed above
- **S3.** In what year were you born?
 - 01 – 18-34 (C=1984-2000)
 - 02 – 35-49 (C=1969-1983)
 - 03 – 50-64 (C=1954-1968)
 - 04 – 65+ (C=1953 OR EARLIER)
- **Q1.** For each of the following, how concerned are you about climate change? (Select one response for each)
 - **Q1.1** As it affects Toronto / the GTA
 - * 04 – Extremely concerned
 - * 03 – Very concerned
 - * 02 – Not very concerned
 - * 01 – Not at all concerned
 - * 99 – Don't know
 - **Q1.2** As it affects the rest of the world
 - * 04 – Extremely concerned
 - * 03 – Very concerned
 - * 02 – Not very concerned
 - * 01 – Not at all concerned
 - * 99 – Don't know
- **Q2.** To what extent do you consider yourself to be informed about what causes climate change? (Select one response only)
 - 04 – Extremely informed
 - 03 – Very informed
 - 02 – Not very informed
 - 01 – Not at all informed
- **D5.** What is the highest level of education that you have completed? (Select one response only)
 - 01 – High school or less
 - 02 – Some community college, vocational, trade school
 - 03 – Completed community college, vocational, trade school
 - 04 – Some university
 - 05 – Completed undergraduate degree

³The survey is not publicly accessible but was made available upon request by the City of Toronto's Environment & Energy Division.

- 06 - Post graduate/professional school
 - 99 - Prefer not to answer
- **D6.** For statistical purposes only, we need information about your income. All individual responses will be kept confidential. Please indicate which category applies to your total household income before taxes for 2017. (Select one response only)
 - 01 - Under \$40,000
 - 02 - \$40,001 to \$60,000
 - 03 - \$60,001 to \$80,000
 - 04 - \$80,001 to \$100,000
 - 05 - \$100,001 to \$150,000
 - 06 - More than \$150,000
 - 99 - Prefer not to answer

7 Appendix 2

Torontonians and Information About Climate Change

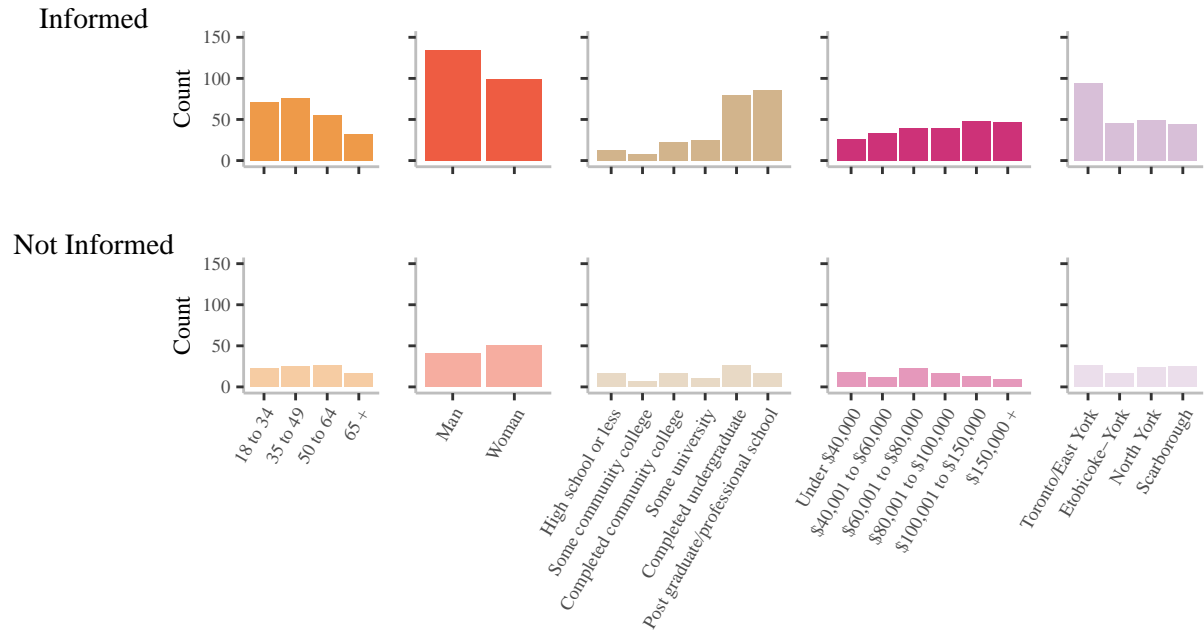


Figure 5: Bar charts of participants' responses about whether or not they are informed about climate change classified by age, gender, education, income and region.

Torontonians and Concern About Local Impacts of Climate Change

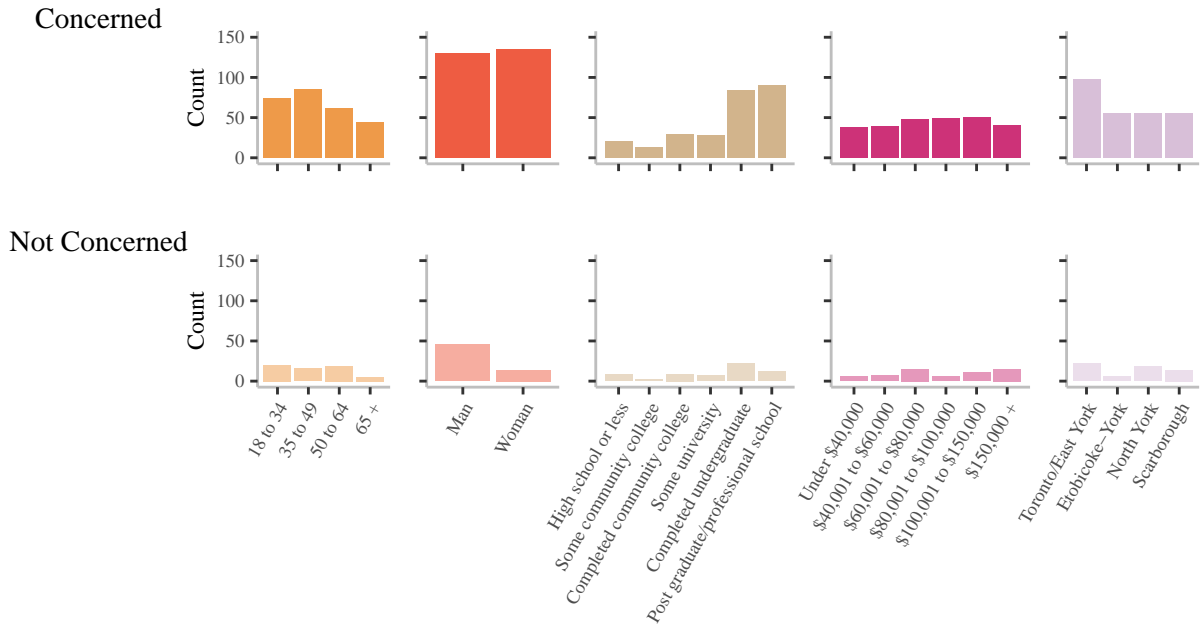


Figure 6: Bar charts of participants' responses about whether or not they are concerned about climate change as it affects Toronto and the Greater Toronto Area, classified by age, gender, education, income and region.

Torontonians and Concern About Global Impacts of Climate Change

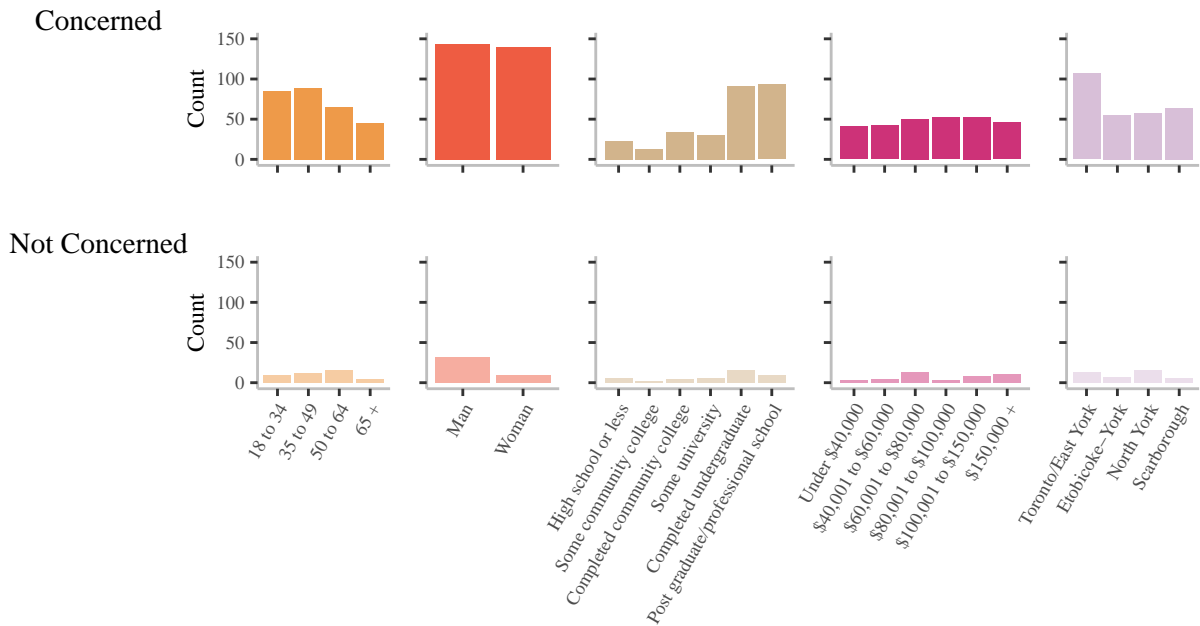


Figure 7: Bar charts of participants' responses about whether or not they are concerned about climate change as it affects the rest of the world, classified by age, gender, education, income and region.

References

- Alexander, Rohan. 2021. “Telling Stories with Data.” <https://www.tellingstorieswithdata.com/causality-from-observational-data.html>.
- Arel-Bundock, Vincent. 2021. *Modelsummary: Summary Tables and Plots for Statistical Models and Data: Beautiful, Customizable, and Publication-Ready*. <https://CRAN.R-project.org/package=modelsummary>.
- Beech, Peter. 2020. “What Is Environmental Racism?” <https://www.weforum.org/agenda/2020/07/what-is-environmental-racism-pollution-covid-systemic/>.
- Berbeco, Minda, and Lawrence C. Hamilton. 2013. “Political Bias Meets Climate Bias: Overcoming Science Denial in a Politically Polarized World.” https://scholars.unh.edu/cgi/viewcontent.cgi?article=1483&context=soc_facpub.
- Berkson, Joseph. 1946. “Limitations of the Application of Fourfold Table Analysis to Hospital Data.” <https://www.jstor.org/stable/3002000?seq=1>.
- City of Toronto. 2018. “Indigenous Climate Action Summary Report.” https://www.toronto.ca/wp-content/uploads/2019/05/8eb4-2019-03-25_Indigenous-Climate-Report_final.pdf.
- . 2020. “Neighbourhood Climate Action Champions.” <https://www.toronto.ca/services-payments/water-environment/live-green-toronto/neighbourhood-climate-action-champions/>.
- . 2021a. “Climate Action Fund.” <https://www.toronto.ca/services-payments/water-environment/environmental-grants-incentives/neighbourhood-climate-action-grants/>.
- . 2021b. “Climate Perception Study.” <https://open.toronto.ca/dataset/climate-perception-study/>.
- . 2021c. “TransformTO.” <https://www.toronto.ca/services-payments/water-environment/environmentally-friendly-city-initiatives/transformto/>.
- . 2021d. “Women4Climate Mentorship Programme.” <https://www.toronto.ca/services-payments/water-environment/live-green-toronto/women4climate-mentorship-program/>.
- Clark, Jeff. 2013. “Toronto Visible Minorities.” <http://neoformix.com/Projects/DotMaps/TorontoVisMin.html>.
- EnviroNics Research. 2018b. “City-Wide Climate Perceptions Study.”
- . 2018a. “City-Wide Climate Perceptions Study.” https://www.toronto.ca/wp-content/uploads/2019/03/97b6-City-of-Toronto_City-wide-Climate-Perceptions-Study_Full-Report.pdf.
- Firke, Sam. 2021. *Janitor: Simple Tools for Examining and Cleaning Dirty Data*. <https://CRAN.R-project.org/package=janitor>.
- Gelfand, Sharla. 2020. *Opendatatoronto: Access the City of Toronto Open Data Portal*. <https://CRAN.R-project.org/package=opendatatoronto>.
- GeoNames. 2019. “GeoNames Data.” <http://www.geonames.org/export/>.
- Healy, Kieran. 2020. “The Kitchen Counter Observatory.” <https://kieranhealy.org/blog/archives/2020/05/21/the-kitchen-counter-observatory/>.
- Hirsi, Isra, Haven Coleman, and Alexandria Villaseñor. 2019. “Adults Won’t Take Climate Change Seriously. So We, the Youth, Are Forced to Strike.” <https://thebulletin.org/2019/03/adults-wont-take-climate-change-seriously-so-we-the-youth-are-forced-to-strike/>.
- Islam, S Nazrul, and John Winkel. 2017. “Climate Change and Social Inequality.” https://www.un.org/esa/desa/papers/2017/wp152_2017.pdf.
- Jackson, David, and Sophie De Coninck. 2019. “How Supporting Climate Action on a Local Level Can Transform the World.” <https://www.weforum.org/agenda/2019/09/local-climate-change-adaptation-good-governance-authorities-undp/>.

- Kahle, David, and Hadley Wickham. 2013. “Ggmap: Spatial Visualization with Ggplot2.” *The R Journal* 5 (1): 144–61. <https://journal.r-project.org/archive/2013-1/kahle-wickham.pdf>.
- Lüdecke, Daniel, Mattan S. Ben-Shachar, Indrajeet Patil, Philip Waggoner, and Dominique Makowski. 2021. “Assessment, Testing and Comparison of Statistical Models Using r.” *Journal of Open Source Software* 6 (59): 3112. <https://doi.org/10.31234/osf.io/vtq8f>.
- Lykke, Nina. 2010. *Feminist Studies: A Guide to Intersectional Theory, Methodology and Writing*. Routledge New York.
- Madson, Diana. 2020. “What Is Intersectional Environmentalism?” <https://yaleclimateconnections.org/2020/09/what-is-intersectional-environmentalism-leah-thomas/>.
- Marlar, Jenny. 2018. “Why Phone and Web Survey Results Aren’t the Same.” <https://news.gallup.com/opinion/methodology/233291/why-phone-web-survey-results-aren.aspx>.
- Moore, Jeffrey C., Linda L. Stinson, and Edward J. Welniak-Jr. 1997. “Income Measurement Error in Surveys: A Review.” <https://www.census.gov/content/dam/Census/library/working-papers/1997/adrm/sm97-05.pdf>.
- Müller, Kirill. 2020. *Here: A Simpler Way to Find Your Files*. <https://CRAN.R-project.org/package=here>.
- O’Riordan, Timothy. 1971. “Public Opinion and Environmental Quality: A Reappraisal.” <http://myaccess.library.utoronto.ca/login?qurl=https>.
- Onur, Arzu, Elvan Sahin, and Ceren Tekkaya. 2012. “An Investigation on Value Orientations, Attitudes and Concern Towards the Environment: The Case of Turkish Elementary School Students.” <https://search-ebshost-com.myaccess.library.utoronto.ca/login.aspx?direct=true&db=eue&AN=73326398&site=ehost-live/>.
- Oxford Public International Law. 2013. “Intergenerational Equity.” <https://opil.ouplaw.com/view/10.1093/law:epil/9780199231690/law-9780199231690-e1421>.
- Padgham, Mark, Bob Rudis, Robin Lovelace, and Maëlle Salmon. 2017. “Osmdata.” *The Journal of Open Source Software* 2 (14). <https://doi.org/10.21105/joss.00305>.
- Pearse, Rebecca. 2016. “Gender and Climate Change.” <https://doi-org.myaccess.library.utoronto.ca/10.1002/wcc.451>.
- Pebesma, Edzer. 2018. “Simple Features for R: Standardized Support for Spatial Vector Data.” *The R Journal* 10 (1): 439–46. <https://doi.org/10.32614/RJ-2018-009>.
- Pedersen, Thomas Lin. 2020. *Patchwork: The Composer of Plots*. <https://CRAN.R-project.org/package=patchwork>.
- SeemaArora-Jonsson. 2011. “Virtue and Vulnerability: Discourses on Women, Gender and Climate Change.” <https://doi.org/10.1016/j.gloenvcha.2011.01.005>.
- Srichai, Naiyana. 1988. “A Study of Environmental Perceptions and Attitudes of Selected University Students in Thailand.” <http://myaccess.library.utoronto.ca/login?qurl=https>.
- Statistics Canada. 2019. “Census Profile, 2016 Census.” <https://www12.statcan.gc.ca/census-recensement/2016/dp-pd/prof/details/page.cfm?Lang=E&Geo1=CSD&Code1=3520005&Geo2=PR&Code2=35&SearchText=Toronto&SearchType=Begins&SearchPR=01&B1=All&GeoLevel=PR&GeoCode=3520005&TABID=1&type=0>.
- Technoscience Research Unit. 2021. “Environmental Data Justice Lab.” <https://technoscienceunit.org/people/lab/>.
- Temple Lang, Duncan. 2021. *RCurl: General Network (HTTP/FTP/...) Client Interface for r*. <https://CRAN.R-project.org/package=RCurl>.
- Turk, A., C. Heneghan, and D. Nunan. 2019. “Non-Response Bias.” <https://catalogofbias.org/biases/non-response-bias/>.

- Unicef. 2020. “Youth for Climate Action: Elevating the Voices of Young People to Protect the Future of Our Planet.” <https://www.unicef.org/environment-and-climate-change/youth-action>.
- Vasa, M, and Srikanth Tamilselvam. 2014. “Building Apps with Open Data in India: An Experience.” <https://www.semanticscholar.org/paper/Building-apps-with-open-data-in-india>.
- Wickham, Hadley. 2016. *Ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag New York. <https://ggplot2.tidyverse.org>.
- . 2020. *Httr: Tools for Working with URLs and HTTP*. <https://CRAN.R-project.org/package=httr>.
- . 2021. *Rvest: Easily Harvest (Scrape) Web Pages*. <https://CRAN.R-project.org/package=rvest>.
- 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>.
- Wickham, Hadley, and Jennifer Bryan. 2019. *Readxl: Read Excel Files*. <https://CRAN.R-project.org/package=readxl>.
- Wickham, Hadley, Jim Hester, and Jeroen Ooms. 2020. *Xml2: Parse XML*. <https://CRAN.R-project.org/package=xml2>.
- Winston Chang. 2014. *Extrafont: Tools for Using Fonts*. <https://CRAN.R-project.org/package=extrafont>.
- Xie, Yihui. 2020. *Bookdown: Authoring Books and Technical Documents with RMarkdown*. <https://github.com/rstudio/bookdown>.
- . 2021. *Knitr: A General-Purpose Package for Dynamic Report Generation in r. R Package Version 1.31*.
- Zhu, Hao. 2020. *kableExtra: Construct Complex Table with ‘Kable’ and Pipe Syntax*. <https://CRAN.R-project.org/package=kableExtra>.