...*

...

Julia Lee

December 3, 2024

By examining Toronto Neighbourhood Profiles data, the following analysis aims to understand what kinds of families drive average neighbourhood income to help support planning decisions for individual neighbourhoods along with local and general businesses. With both data visualization and the application of a linear regression model, this analysis finds that 2-person and 4-person families are key drivers of average neighbourhood income throughout Toronto. This implies that focusing neighbourhood or business planning decisions on these types of families may help to increase a neighbourhood or business' income position. Further, this analysis highlights that additional considerations for the unique characteristics of these key driver families are important to understand the interplay between families and neighbourhood income in more depth.

1 Introduction

The interactions families have with their surrounding environments (i.e. their neighbourhoods) can be influenced by both the characteristics of the environment and the kinds of resources that families have. Income and financial resources can be strong determinants of the types of opportunities (e.g. housing, education, employment, etc.) families are able to access within their neighbourhoods (Thomas et al. 2014). With this, to better understand the concept of income across the city of Toronto's neighbourhoods, this analysis aims to identify and determine the kinds of families that are driving average neighbourhood income levels.

This analysis can help support and inform planning decisions for individual neighbourhoods along with local and general businesses. With a better understanding of the livelihoods of families in Toronto, it may be possible to better allocate resources to the neighbourhoods that have few families that drive the average neighbourhood income. Further, planning decisions that center around effectively targeting or considering the families that drive neighbourhood

*Code and data are available at: https://github.com/jjlee-lee/Toronto_Neighbourhood_Income.git income may help to strengthen the income position of various neighbourhoods or local businesses. Thus, this analysis serves as a starting point to further understand the interaction between families and neighbourhood income.

By considering data from the 2021 Canadian Census, this analysis begins by investigating the distribution of different census family sizes (i.e. 2, 3, or 4-person families) across all 158 of Toronto's neighbourhoods. Next, it creates a multiple linear regression model that examines the relationship between average neighbourhood income and different family sizes. The results of this model demonstrate the family sizes that significantly impact neighbourhood income. Lastly, after identifying key drivers of average neighbourhood income through the model, the analysis considers the geographic locations of those drivers. As a result, the present analysis finds that 2 and 4-person families are key drivers of average neighbourhood income. Additionally, as there are varying numbers of 2 and 4-person families across Toronto neighbourhoods, it is important to also consider the characteristics of individual families to ensure effective neighbourhood or local business planning decisions.

The following provides an in-depth description of the data used within this analysis Section 2, an explanation of the analysis model Section 3, a detailed account of the findings Section 4, and a thorough discussion of the overall analysis Section 5.

2 Data

To simulate, test, download, and clean the Neighbourhoods Profile data, the statistical programming language R was used (R Core Team 2023). Specific libraries that assisted the analysis include `tidyverse` (Wickham et al. 2019), `opendatatoronto` (Gelfand 2022), `tinytex` (Xie 2019), `ggplot2` (Wickham 2016), `knitr` (Xie 2015), `testthat` (Wickham 2011), `here` (Müller and Bryan 2020), `arrow` (Richardson et al. 2024), `modelsummary` (Arel-Bundock 2022), and `sf` (Pebesma 2018).

2.1 Neighbourhood Profiles Data

Neighbourhood Profiles data for the city of Toronto is provided by the Social Development, Finance & Administration from Toronto's Open Data portal (Social Development, Finance & Administration 2024a). The dataset obtained from the city of Toronto's Open Data portal (City of Toronto 2024), is a collection of data from the 2021 Canadian Census from Statistics Canada (Statistics Canada 2024). Records of socio-economic data for 158 geographic regions (i.e. social planning neighbourhoods) in Toronto can be found within this dataset. From age to ethnocultural diversity, the Census data highlights both the socio-economic and demographic characteristics of Toronto residents across its individual neighbourhoods.

The Neighbourhood Profiles data is an extensive dataset with each row representing a demographic or socio-economic characteristic and each column reflecting a Toronto neighbourhood.

For each neighbourhood, there is information on its name, identification number, and status. Status refers to a neighbourhood's Toronto Strong Neighbourhoods Strategy (TSNS) designation, which indicates whether a neighbourhood is an emerging area, an area that needs improvement, or neither. Additionally, data for each neighbourhood's age population, income, education, and more from the 2021 Census is provided. Table 1 below offers a small preview of this dataset.

Table 1: Toronto Neighbourhood Profiles Data

Neigbourhood Name	West Humber-Clairville
Neighbourhood Number	1
TSNS 2020 Designation	Not an NIA or Emerging Neighbourhood
Total - Age groups of the population - 25% sample data	33300
0 to 14 years	4295
0 to 4 years	1460
5 to 9 years	1345

Table 1 displays the name, number, and TSNS designation along with data for the number of individuals within different age groups for a neighbourhood (West Humber-Clairville) in Toronto. By looking at Table 1, West Humber-Clairville is a neighbourhood that is not an emerging area and one that does not need improvement. It is also a region with a fairly large number of children ages 0 to 14 years old. The rest of the dataset follows the same format, displaying different characteristics (e.g. older age groups, education attainment, income levels, etc.) of all 158 neighbourhoods in the city of Toronto.

2.2 Analysis Data

For the present analysis, the variables of interest are the different types of census family sizes along with the average after-tax income for each of Toronto's neighbourhoods. The Neighbourhood Profiles data includes four different census family sizes: (1) 2-person families, (2) 3-person families, (3) 4-person families, and (4) five or more-person families. In simple terms, Statistics Canada defines a census family as one where all family members (related by blood marriage, common-law union, adoption, or a foster relationship) live together in the same dwelling (Statistics Canada 2023). Census families can also be referred to as economic families (Statistics Canada 2021). The data further provides the average after-tax income ($) recorded in 2020 for each neighbourhood.

Since the objective is to find out what kinds of families are driving average neighbourhood income in Toronto, the data used throughout this analysis reflects the average income level for all 158 neighbourhoods alongside their counts of different census family sizes. With minimal data wrangling, the analysis data is simply an extraction of the larger dataset with a focus on the different family sizes and average after-tax income. Table 2 below illustrates this analysis data for three neighbourhoods and summary statistics for this data can be found in the Appendix (Section A.1).

Table 2: Toronto Neighbourhood Profiles Analysis Data

Name	Number	2 persons	3 persons	4 persons	5 or more persons	Average Income
West Humber-Clairville	1	3635	2265	2025	805	101300
Mount Olive-Silverstone-Jamestown	2	2855	2145	1765	1290	85300
Thistletown-Beaumond Heights	3	1095	665	555	310	98100

In Table 2, the "Name" column represents each neighbourhood's name, and the "Number" column reflects each neighbourhood's identification number. The variables, "2 persons", "3 persons", "4 persons", and "5 or more persons" represent the number of census families in each neighbourhood with those particular family sizes. Lastly, the "Average Income" variable reflects the average after-tax income for all neighbourhoods in 2020 (recorded in dollars). Within the analysis data file provided by this analysis, there are additional variables that represent the total number of census families by family size, the average census family size, and the average number of children in census families for each neighbourhood. These variables are considered to obtain additional context about how families are made up and distributed across Toronto's neighbourhoods. It is important to note that while this additional information provided a better understanding of the variables of interest, they are not included in the model of the analysis. A detailed account of the model can be found in (Section 3).

2.3 Map Data

To further understand the type of families that drive average neighbourhood income, this analysis also uses Neighbourhoods data – a shapefile that contains geographic information of Toronto neighbourhood boundaries – to map the results of its model. The boundaries of each neighbourhood are defined using census tract information provided by Statistics Canada, and the shapefile itself is published by the Social Development, Finance & Administration from Toronto's Open Data portal (Social Development, Finance & Administration 2024b). The map within this analysis is created using ArcGIS Pro software (Esri 2024). By joining the analysis data shown in Table 2 with the shapefile based on neighbourhood names, a map that highlights the distribution of family sizes that drive average neighbourhood income can be created. The information contained in the shapefile is shown below in Table 3.

Table 3: Toronto Neighbourhood Location & Boundaries Data

Object ID	Neighbourhood Number	Neighbourhood Name	TSNS Designation	Geometry
1	174	South Eglinton-Davisville	Not an NIA or Emerging Neighbourhood	POLYGON ((-79.38635 43.6978…
2	173	North Toronto	Not an NIA or Emerging Neighbourhood	POLYGON ((-79.39744 43.7069…
3	172	Dovercourt Village	Not an NIA or Emerging Neighbourhood	POLYGON ((-79.43411 43.6601…

The Neighbourhoods shapefile data contains neighbourhood names, numbers, and TSNS designations similar to the Neighbourhood Profiles data (Table 1). It also includes geographic information (i.e. the geometry/coordinates of a neighbourhood) that defines the boundaries of each neighbourhood as shown in Table 3.

2.4 A Note on Measurement

The Neighbourhood Profiles data is a reflection of the 2021 Census for neighbourhoods in Toronto, meaning that it measures and represents the demographic and socio-economic characteristics of families across Toronto at this particular time period. This information is collected through the use of questionnaires – a short and long-form census. The short-form census is sent out to all households, and it asks questions about simple characteristics like age, gender, and household size. With this information, the short-form census attempts to enumerate all individuals within a geographic region and may impute any missing data. The long-form census is sent out to only 25% of households and is a deeper source of data that outlines information such as housing, education, and ethnicity.

So, various aspects of families' livelihoods are measured through the Census and aggregated to create an overall picture of particular geographic regions – in this case, Toronto neighbourhoods. It is important to note that through this measurement process, there is room for error as well as a need to be aware of this potential for error. Missing data values can be imputed and the aggregation of individual household characteristics or experiences can lead to incorrect assumptions about certain regions and their populations. Thus, this analysis presents its findings with this in mind.

3 Model

To understand the types of families that drive average neighbourhood income, this analysis constructs a multiple linear regression model using R (R Core Team 2023) that considers average neighbourhood income as the response variable and the different census family sizes as predictors. As mentioned in Section 2, the average neighbourhood income variable reflects the average after-tax income of all census families for each neighbourhood in 2020 (recorded in dollars). Different family sizes (i.e. 2 to 5 or more-person families) refer the number of family members that live in the same dwelling within each individual neighbourhood. With this, the final model within this analysis is structured as follows:

$$\hat{y} = b_0 + b_1 \cdot x_1 + b_2 \cdot x_2 + b_3 \cdot x_3 + \epsilon$$

where

- $\hat{y}$ represents the average neighbourhood after-tax income,
- b_0 represents the intercept of the regression model,
- b_1 represents the effect of two-person census families,
- x_1 represents the number of two-person census families,
- b_2 represents the effect of three-person census families,
- x_2 represents the number of three-person census families,
- b_3 represents the effect of four-person census families,
- x_3 represents the number of four-person census families,
- ϵ captures the error within the regression model

It is important to highlight that as a result of model selection and individual predictor t-tests, families with 5 or more members are not considered in the final model. The impact of 2, 3, and 4-person families on average neighbourhood income across Toronto neighbourhoods are examined through this model.

3.1 Model justification

As the objective of the analysis is to understand what might be driving the average neighbourhood income in terms of family size, a multiple linear regression will allow for an examination of the relationship between different census family sizes and average income. By striving to preserve the interpretability of the model, this analysis can further determine which family sizes are meaningful influencers of a neighbourhood's average after-tax income.

The underlying assumption of linearity that multiple linear regression models hold can be a potential limitation as these models will not be able accurately explain the influence of predictors on a response variable if these variables have a non-linear relationship with each other. The use of a multiple linear regression may not be the best choice in these cases. However, with average neighbourhood income and increasing census family sizes, the model's diagnostic plots provide evidence for linearity, meaning that the use of a multiple linear regression model can be suitable for this analysis. Further justification and validation for the model can be found in the Appendix (Section A.2).

4 Results

4.1 Distribution of Different Census Family Sizes Across Toronto Neighbourhoods

Figure 1, Figure 2, and Figure 3 illustrate the number of 2, 3, and 4-person families, respectively, across Toronto's 158 neighbourhoods. From this, 2-person census families appear to be the most common as almost all neighbourhoods have close to 1000 2-person families with several of them having over 3000, near 4000 2-person families. On the other hand, 3 and 4-person families seem to less common with more neighbourhoods having less than 1000 families of these sizes (i.e. under or just over 500 families) and with the maximum number of 3 or 4-person families being just under 3000. These observations seem to suggest that given the higher numbers of 2-person families across Toronto neighbourhoods, this family size may be a key driver of average neighbourhood income within the city of Toronto.

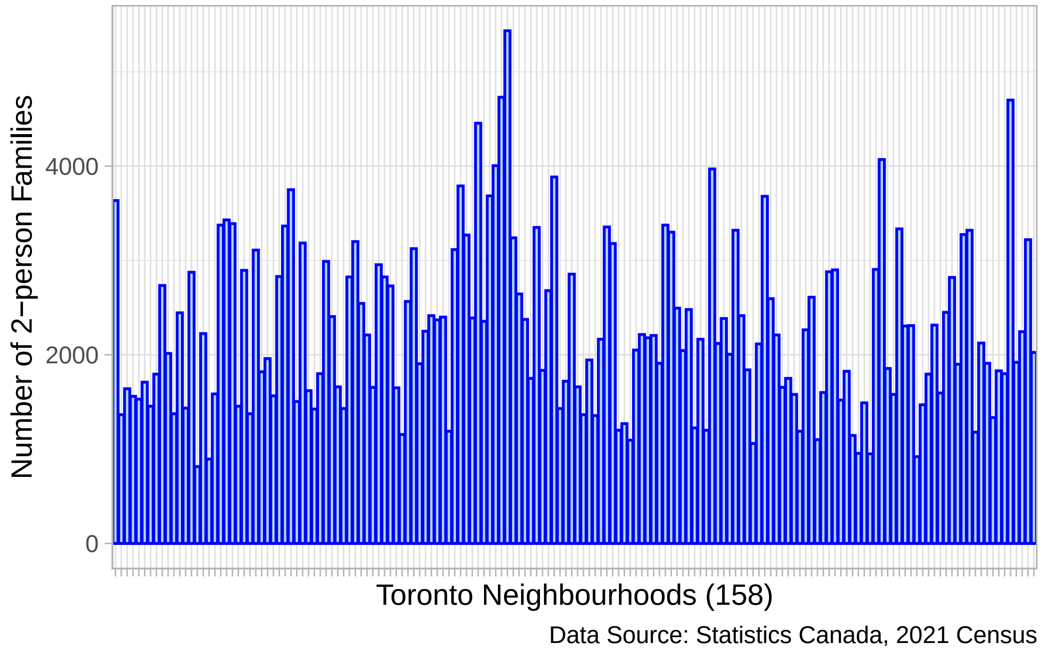

Figure 1: Number of 2-person Census Families Across Toronto Neighbourhoods

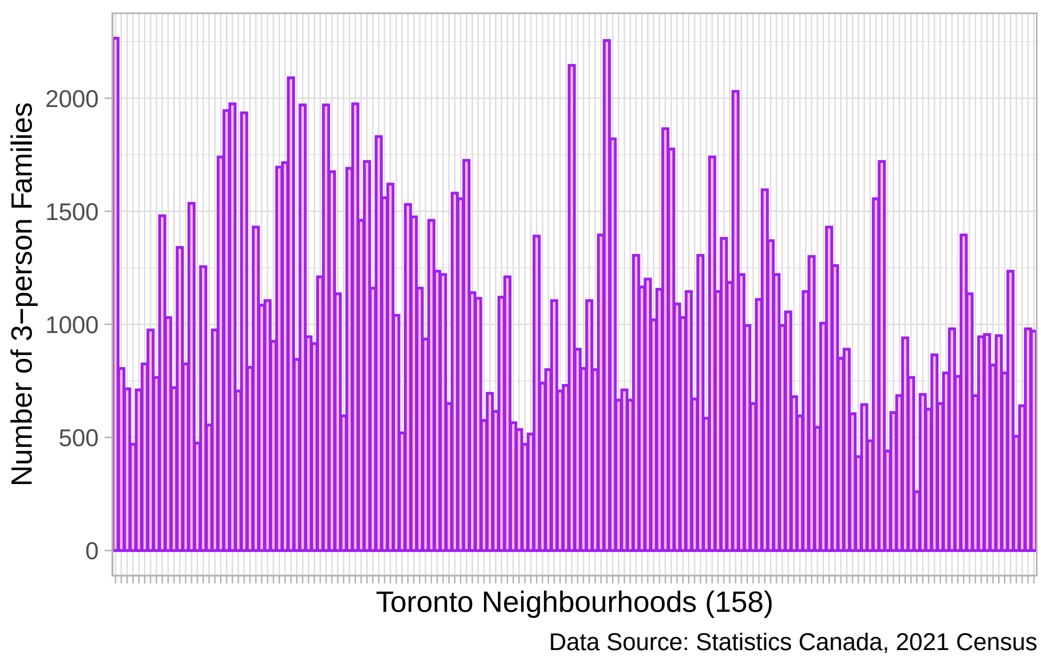

Figure 2: Number of 3-person Census Families Across Toronto Neighbourhoods

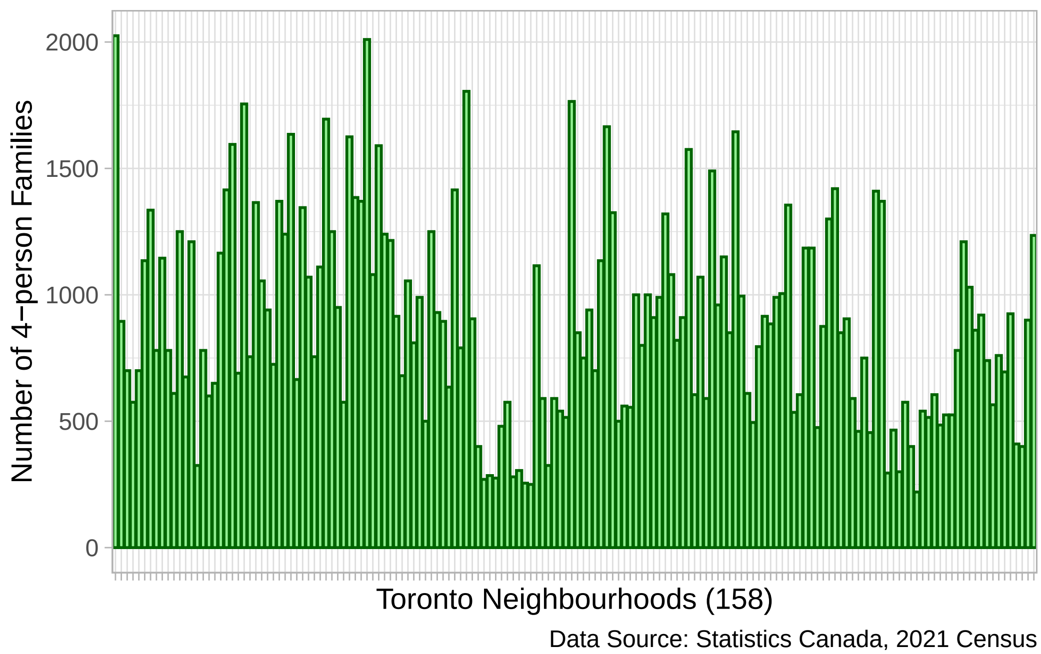

Figure 3: Number of 4-person Census Families Across Toronto Neighbourhoods

4.2 Model Results

The results from the model shown in Table 4 indicate significant relationships between average neighbourhood income and different census family sizes ($p < 0.001$ for each family size). This means that the addition of a 2-person census family to a neighbourhood is expected to increase the neighbourhood's average after-tax income by around \$28.59 ($\pm$ 4.76). Likewise, the addition of a 4-person family to a neighbourhood is expected to increase the neighbourhood's average after-tax income by approximately \$148.40 ($\pm$ 16.32). Interestingly, the addition of a 3-person census family to a neighbourhood is expected to decrease the neighbourhood's average after-tax income by around \$184.42 ($\pm$ 17.66). These findings seem to suggest that 4-person families are also be key drivers of average neighbourhood income. Additionally, this may imply that on average, while 2 and 4-person families can commonly have above-average incomes, more 3-person families may have below-average incomes across Toronto's neighbourhoods.

Table 4: Summary Results of Analysis Model

	(1)
(Intercept)	125 318.176
	8677.734 (<0.001)
size_two	28.593
	4.763 (<0.001)
size_three	−184.416
	17.657 (<0.001)
size_four	148.403
	16.318 (<0.001)
Num.Obs.	158
R2	0.424
R2 Adj.	0.413
AIC	3759.5
BIC	3774.8
Log.Lik.	−1874.736
RMSE	34 423.06

4.3 Map Visualization

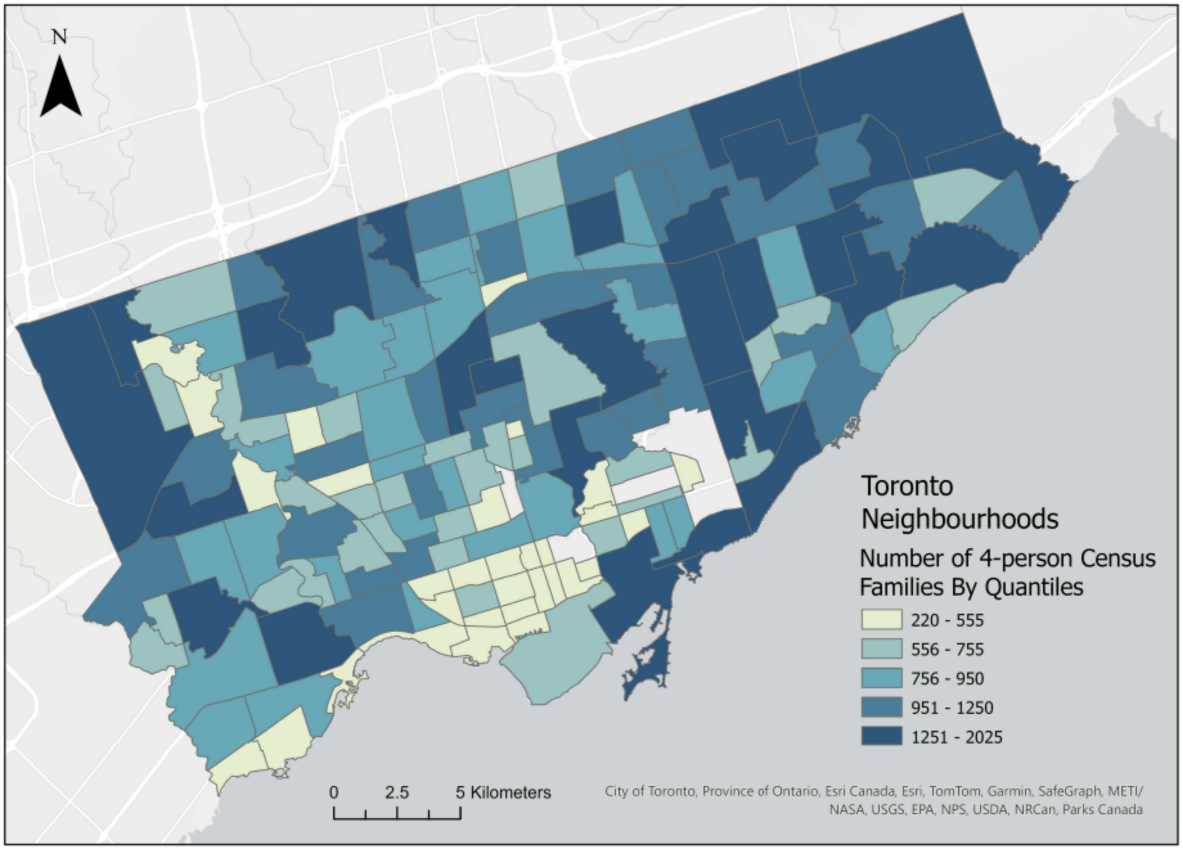

Figure 4: Distribution of 4-person Census Families Across Toronto

The initial analysis of the distribution of different census family sizes across Toronto neighbourhoods highlighted that 2-person families can be key drivers of average neighbourhood income. The results of the model further suggest that 2 and 4-person families positively contribute to average neighbourhood income. With this in mind, Figure 4 maps the distribution of 4-person census families throughout the city of Toronto. This shows how the majority of 4-person census families are found in the neighbourhoods towards the east end of the city with a few neighbourhoods near the city's centre and the Harbourfront area. Figure 5 demonstrates the distribution of 2-person families. Here, there appears to be a shift in that more 2-person census families seem to be in neighbourhoods closer to the city's centre and the west side of the Harbourfront area. These illustrations provide a better understanding of where the key drivers of average neighbourhood income are located.

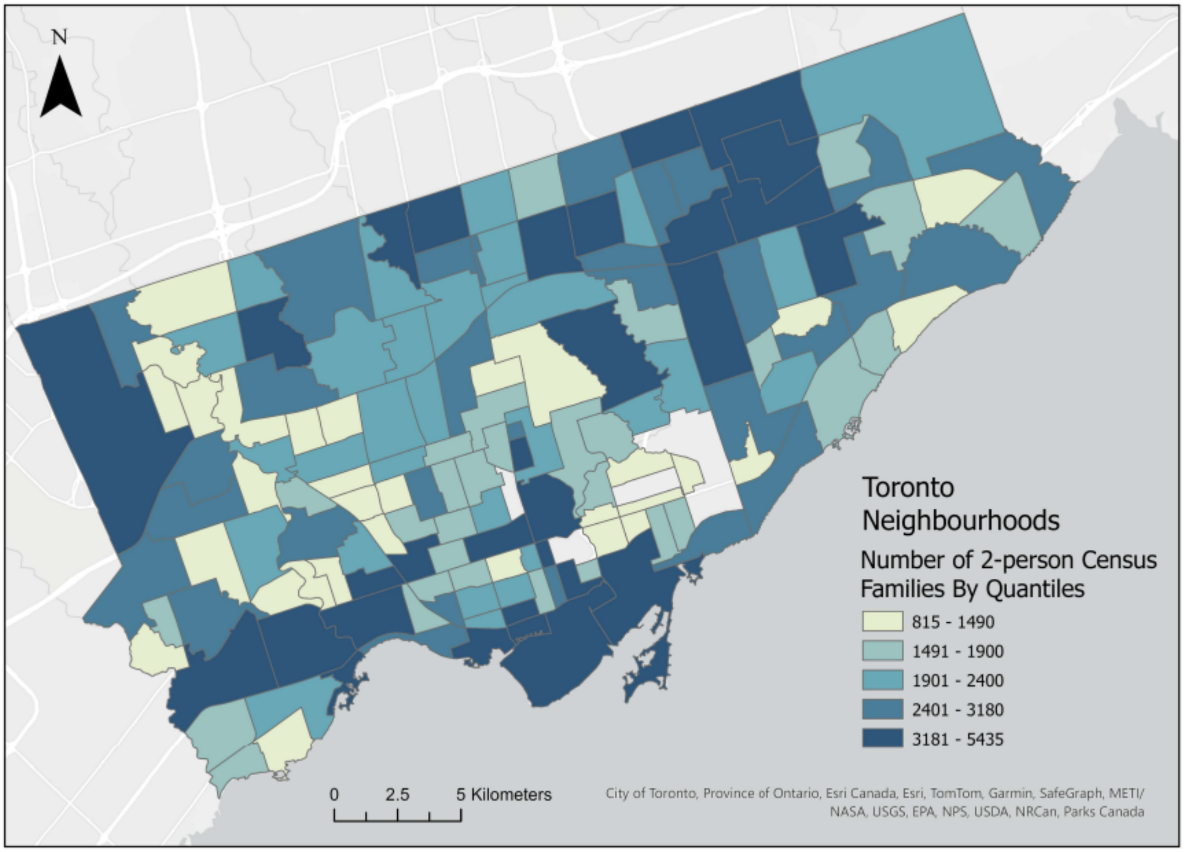

Figure 5: Distribution of 2-person Census Families Across Toronto

5 Discussion

5.1 Summary

The present analysis aims to understand the kinds of families that drive average neighbourhood income in Toronto. To learn more about the families that reside in Toronto neighbourhoods, the analysis observes the distribution of different census family sizes with Figure 1, Figure 2, and Figure 3. It also illustrates the geographic locations of Toronto families with Figure 4 and Figure 5. By constructing a multiple linear regression model that examines the relationship between average neighbourhood income and different census family sizes, the analysis provides further support for its observations and findings.

5.2 Analysis Findings

A study like the present analysis can help support planning decisions for individual neighbourhoods along with local and general businesses. The analysis finds that 2 and 4-person census families are key drivers of average neighbourhood income.

Considering the results of the model, the output shows that the baseline average neighbourhood income is approximately $125, 318 – meaning that Toronto neighbourhoods may have a healthy level of disposable income. This opens up opportunities for neighbourhoods as well as businesses. The question is, however, which neighbourhoods are more affluent and how can neighbourhoods raise the income level to become economically stronger? The answer to this can be found by looking at the impact of different census family sizes. To start, 2-person census families drive the average neighbourhood income by around $28 per family whereas 4-person homes drive income by approximately $148. While the amount of income that 2-person families contribute is smaller than that of 4-person families, the impact that 2-person families have on neighbourhood income can be just as large because this family size is most commonly found across Toronto's neighbourhoods. Further, 3-person families are found to decrease the income by around $184, suggesting that these families may have lower incomes compared to other family sizes.

With this, organizations could use this information to make planning decisions. For instance, neighbourhoods could target 2 or 4-person census families when planning for housing in the interest of attracting families that will stimulate the economy. Additionally, local and general businesses could focus on these even-numbered families as a means of effectively promoting their products or shaping their brand. This is not to say that 3-person census families should be ignored. However, it should be noted that for every 3-person family, it takes approximately one 2-person family plus a 4-person family to make up for the decrease in average neighbourhood income. The impact of 3-person families on neighbourhood could also be mitigated by reason of the large number of 2-person families that live throughout Toronto. In short, for planning decisions in the interest of increasing a neighbourhood or business' income position, there may be a benefit in considering the interests or need of 2 and 4-person families.

This analysis further visualizes the neighbourhoods where the most 2 and 4-person families live. Knowing the geographic regions in which the key drivers of average neighbourhood income are located can further help work towards a more equitable distribution of opportunities and resources. For example, in the neighbourhoods that have many 2 or 4-person census families, more opportunities such as higher quality schools or higher employment availability may be present and more resources could be allocated to these neighbourhoods to further support or maintain those opportunities. On the other hand, understanding where the drivers of neighbourhood income are also helps to identify the areas where there are more families such as 3-person families who are not able to contribute as much to the average neighbourhood income. By using this information, planning decisions could be made to provide these neighbourhoods with resources for better opportunities with education and employment, for example, to ensure equitable resource distributions across all Toronto neighbourhoods.

5.3 Limiations & Future Directions

Even with the deeper understanding that this analysis brings forth, there are some limitations to consider. The model within this analysis centers around interpretability, meaning that despite seeing some evidence against its underlying assumptions, transformations were not applied to the data to resolve these diagnostic issues. As the application of transformations make it challenging to interpret coefficient estimates and what they mean in the context of the analysis, the present analysis did not engage in this step of the model building process. However, this limits the model's ability to make more accurate predictions that capture what the average neighbourhood income would be with the addition of 2, 3, or 4-person census families to a given neighbourhood. So, findings that point to exact numbers of average neighbourhood income for different numbers of varying family sizes could not be included in the analysis.

In addition to this, the potential for error that exists within the Neighbourhood Profiles data from the 2021 Canadian Census (as briefly mentioned in Section 2.4) can be another limitation that hints at a compelling future direction for this analysis. As the Neighbourhood Profiles data is the aggregation of individual family or household characteristics and experiences, it can provide a better understanding of the overall picture – that is, the general patterns and trends within the city of Toronto. However, it is limited in capturing unique or individual characteristics which can be lost during its process of aggregating data. Thus, with this overall understanding about the key drivers of average neighbourhood income, future research could be focused on individual neighbourhoods (the neighbourhoods with many 4-person families, for instance). These future studies could look at how key families interact with the local resources (e.g. grocery stores, health services, etc.) available to them to understand how they contribute to their neighbourhood's average income. It may also be interesting to examine other characteristics (e.g. education attainment, ethnicity, etc.) that might explain why 2 and 4-person families are drivers of the average income for their neighbourhoods.

A Appendix

A.1 Analysis Data Summary Statistics

Table 5: Toronto Neighbourhood Profiles Data Summary Statistics

2-person Families	3-person Families	4-person Families	5 or more-person Families	Average Income
Min. : 815	Min. : 260.0	Min. : 220.0	Min. : 40.0	Min. : 76800
1st Qu.:1596	1st Qu.: 716.2	1st Qu.: 590.0	1st Qu.: 185.0	1st Qu.: 93450
Median :2172	Median :1030.0	Median : 867.5	Median : 310.0	Median :108150
Mean :2288	Mean :1099.4	Mean : 900.2	Mean : 351.4	Mean :121582
3rd Qu.:2879	3rd Qu.:1393.8	3rd Qu.:1180.0	3rd Qu.: 460.0	3rd Qu.:129500
Max. :5435	Max. :2265.0	Max. :2025.0	Max. :1290.0	Max. :351600

Table 5 presents the summary statistics for each census family size and the average after-tax income across all 158 neighbourhoods in Toronto.

A.2 Model Validation

The construction of a model for this analysis began by fitting an initial model with average neighbourhood income as the response variable and 2, 3, 4, and 5 or more-person families as predictors. As seen by Table 6 and the individual t-tests, 5 or more-person families do not appear to significantly impact a neighbourhood's average income with the largest p-value among all predictors ($p > 0.001$).

So, another model was created without this variable (i.e. a reduced model), and all predictors seem to have significant impacts ($p < 0.001$) on neighbourhood income as shown in Table 7.

Model selection was further conducted by comparing the BIC scores of both models (shown in Table 8). As the new model has a lower BIC score, this is the one that is used within the analysis and outlined in Section 3.

Table 6: Summary Table of the Initial Model

	(1)
(Intercept)	125 502.083
	8644.967 (<0.001)
size_two	26.444
	4.960 (<0.001)
size_three	−178.177
	18.084 (<0.001)
size_four	157.742
	17.431 (<0.001)
size_five_plus	−29.981
	20.206 (0.140)
Num.Obs.	158
R2	0.433
R2 Adj.	0.418
AIC	3759.2
BIC	3777.6
Log.Lik.	−1873.608
RMSE	34 178.04

Table 7: Summary Table of the Reduced Model

	(1)
(Intercept)	125 318.176
	8677.734 (<0.001)
size_two	28.593
	4.763 (<0.001)
size_three	−184.416
	17.657 (<0.001)
size_four	148.403
	16.318 (<0.001)
Num.Obs.	158
R2	0.424
R2 Adj.	0.413
AIC	3759.5
BIC	3774.8
Log.Lik.	−1874.736
RMSE	34 423.06

Table 8: BIC Scores for Initial and Reduced Models

Model	BIC Score
Initial	3777.6
Reduced	3774.8

The purpose of the model is to be able to interpret the coefficient estimates and what they mean in terms of influencing average neighbourhood income. For this reason, transformations (e.g. log transformations on predictor variables) were not applied to the data even with the presence of some evidence against the linear regression model assumptions.

Looking at the diagnostic plots (Figure 6), the lack of observational patterns in the Residuals vs Predictor plots for 2, 3, and 4-person families points to evidence for the linearity assumption of linear regression models.

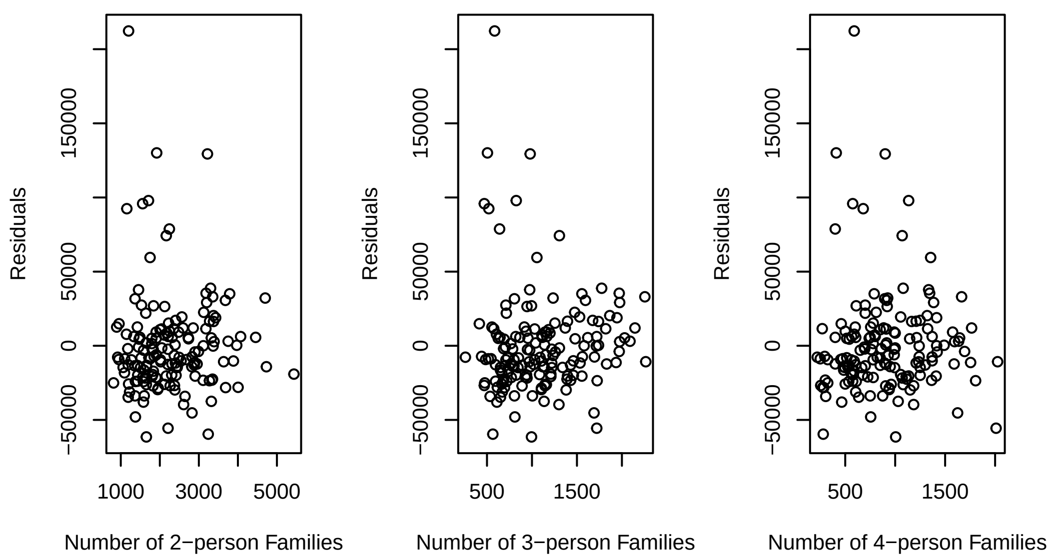

Figure 6: Showing the Residual vs. Predictors

The Residual vs Fitted Values plot (Figure 7) seems to have some clustering of points, indicating that the constant variance and independence assumptions are somewhat satisfied. However, because the overall pattern of the plot is still relatively spread out with no observable patterns, and as the objective of the model is interpretability, transformations were not applied to the variables in the model to better these assumptions.

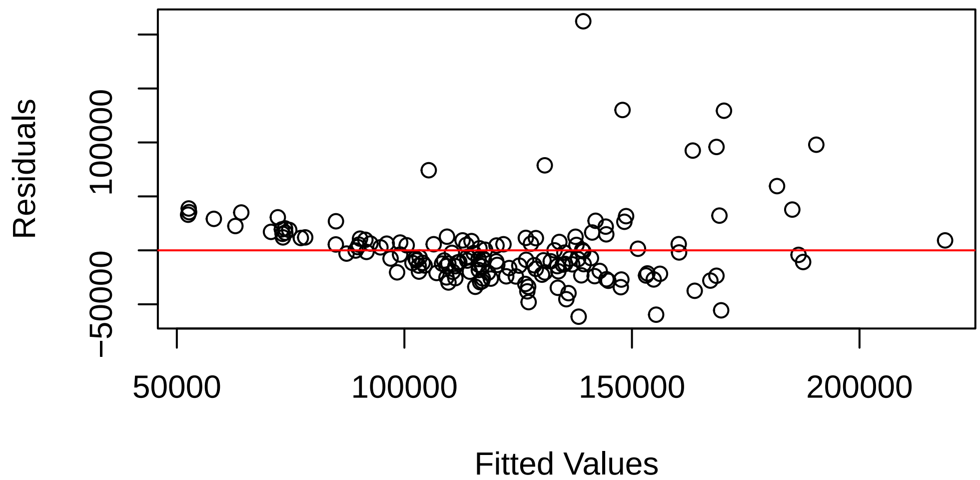

Figure 7: Showing the Residual vs. Fitted Values

Lastly, though the Normal Q-Q plot (Figure 8) shows some deviation from the Q-Q line, the normality assumption is somewhat satisfied with little evidence to suggest an alarming violation of this assumption. Putting all this together, while improvements can be made to the model for better, more accurate predictions, it is sufficient for the purpose of interpretability.

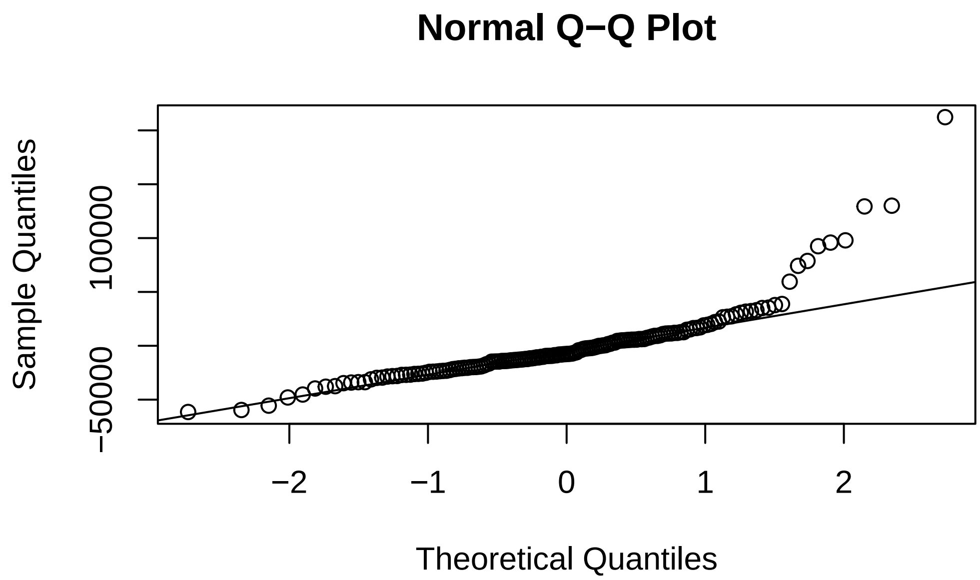

Figure 8: Showing the Normal Q-Q Plot

References

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

City of Toronto. 2024. "City of Toronto's Open Data Portal." https://open.toronto.ca/.

Esri. 2024. "ArcGIS Pro." https://www.esri.ca/en-ca/products/gis-mapping-products/arcgis-pro/overview.

Gelfand, Sharla. 2022. "opendatatoronto: Access the City of Toronto Open Data Portal." https://cran.r-project.org/web/packages/opendatatoronto/index.html.

Müller, Kirill, and Jennifer Bryan. 2020. *here: A Simpler Way to Find Your Files*. https://here.r-lib.org/.

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.

R Core Team. 2023. *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. 2024. *arrow: Integration to 'Apache' 'Arrow'*. https://github.com/apache/arrow/.

Social Development, Finance & Administration. 2024a. "Neighbourhood Profiles." https://open.toronto.ca/dataset/neighbourhood-profiles/.

———. 2024b. "Neighbourhoods." https://open.toronto.ca/dataset/neighbourhoods/.

Statistics Canada. 2021. "Economic family." https://www23.statcan.gc.ca/imdb/p3Var.pl?Function=Unit&Id=33863.

———. 2023. "Census family." https://www23.statcan.gc.ca/imdb/p3Var.pl?Function=Unit&Id=32746.

———. 2024. "Reference materials, 2021 Census." https://www12.statcan.gc.ca/census-recensement/2021/ref/index-eng.cfm.

Thomas, Hannah, Tatjana Meschede, Alexis Mann, Allison Stagg, and Thomas Shapiro. 2014. "Location, Location, Location: The Role Neighborhoods Play in Family Wealth and Well-Being." https://heller.brandeis.edu/iere/pdfs/racial-wealth-equity/leveraging-mobility/location.pdf.

Wickham, Hadley. 2011. *testthat: Get Started with Testing*. *The R Journal*. Vol. 3. https://journal.r-project.org/archive/2011-1/RJournal_2011-1_Wickham.pdf.

———. 2016. "ggplot2: Elegant Graphics for Data Analysis." 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.

Xie, Yihui. 2015. *Dynamic Documents with R and Knitr*. 2nd ed. Boca Raton, Florida: Chapman; Hall/CRC. https://yihui.org/knitr/.

———. 2019. "TinyTeX: A Lightweight, Cross-Platform, and Easy-to-Maintain LaTeX Distribution Based on TeX Live." *TUGboat* 40 (1): 30–32. https://tug.org/TUGboat/Contents/contents40-1.html.