Do Lower-Income Wards and Older Apartment Buildings Obtain Lower Apartment Building Evaluation Scores? An Analysis of the 2023 Apartment Building Evaluation Scores.*

Hari Lee Robledo

January 25, 2024

The *RentSafeTO: Apartment Building Standards* and Ward Profiles (25-Ward Model) data sets are analyzed in this paper. The goal is to explore correlations between apartment building evaluation scores, ward income, and construction year. The analysis reveals a very weak correlation between lower-income wards and higher evaluation scores, challenging assumptions about the relationship between income and housing conditions. Contrarily, a strong correlation is revealed between older apartment buildings and lower evaluation scores. This analysis emphasizes the need for further research to inform regulations for healthier housing conditions.

Table of contents

*Code and data from this analysis are available at: https://github.com/hari-lr/RentSafeTO-Building-Evaluation

1 Introduction

In 2017, *RentSafeTO: Apartment Building Standards* was established as a regulatory initiative that aims to enforce bylaws to guarantee that apartment building owners and operators adhere to building maintenance standards. This initiative only applies to apartment buildings with three or more stories and/or a minimum of 10 units; excluding condominiums, townhomes, and units within private homes, such as basements or main-floor apartments (Gelfand 2022). This initiative is enforced by Bylaw Enforcement Officers who inspect apartment building common areas, mechanical and security systems, parking, and exterior grounds and proceed to evaluate based on the conditions of these aspects (Gelfand 2022). In 2023, the program increased the number of evaluation categories from 20 to 50 and "allocat[ed] a weight to the category" (Toronto 2023).

Although information and scores from *RentSafeTO* are public and accessible on multiple platforms (Anonen 2021), there is a gap in the research about apartment building evaluation scores and other factors. Filling this informational gap is important so that all tenants are equipped to access healthy housing and support (Phipps 2018). Thus, this paper aims to analyze RentSafeTO data. The goal is to explore and find a strong positive correlation between the apartment buildings and their scores, considering factors such as ward income and the year they were built.

According to the RentSafe Project Team research. individuals with low incomes or other marginalized groups are disproportionately affected by unfit housing conditions. These issues range from mental health problems to physical ailments (Phipps 2018). These problems are caused by mold, pests, lead, radon, and other unhealthy housing conditions (Phipps 2018), including aging housing stock. Nonetheless, the findings of this analysis indicate a very weak correlation between lower-income wards and higher evaluation scores and a strong correlation between older apartment buildings and lower evaluation scores.

The paper is divided into the following sections: Data, where I explain how I collected and cleaned the data; Results, where I present trends and correlations found from the data; Discussion, where I compare and evaluate the data; and Conclusion, where I summarize the findings from this paper.

2 Data

The data utilized in this paper was retrieved from the City of Toronto's Open Data Portal, specifically the Open Data Toronto library (Gelfand 2022). The data sets analyzed are titled Apartment Building Evaluations 2023 – current (Data 2024) and Ward Profiles (25-Ward Model)(Data 2021). Data was collected, cleaned, and analyzed using the open-source statistical programming software R (R Core Team 2023). This process involved various packages within R, including tidyverse (Wickham et al. 2019), `ggplot2` (Wickham 2016), `dplyr` (Wickham et al. 2023), `readxl` (Wickham and Bryan 2023), `tibble` (Müller and Wickham 2023), `janitor` (Firke 2023), `KableExtra` (Zhu 2021), `knitr` (Xie 2014), `ggbeeswarm` (Clarke, Sherrill-Mix, and Dawson 2023), and `ggrepel` (Slowikowski 2024). A comprehensive description of the data gathering and cleaning process is provided in the following subsections.

2.1 Apartment Building Evaluation Data

This data set, published by the Municipal Licensing & Standards (Data 2023), outlines building evaluation scores for buildings registered with RentSafeTO initiative. For each apartment building registered, the data set includes identification and description such as address, year of registration and evaluation scores, as well as data regarding the fifty possible areas of evaluation and their respective scores.

Based on previous research and data assumptions, columns related to the evaluated areas and their scores were deemed beyond the scope of this paper. Additionally, it is important to highlight that there are multiple building scores within the data set. However, due to recent changes in the evaluation process, all the scores reflect the same number. Thus, for this document, I have selected the *Proactive Building Score*, which represents the most recent evaluation of the common areas of a building (a score out of 100%) and reflects the weighted score of all applicable evaluation categories. Following this, I started with basic data cleaning to discard the unnecessary columns (see Table 1).

2.2 Average Building Scores per Ward

To determine a measure of central tendency that considers all the scores in the cleaned data set, I employed the summarise function to calculate the mean score per ward and then arranged the average scores per ward (see Table 2).

Table 1: Sample of Cleaned Apartment Building Evaluation Data

ID	Ward	Year Built	Ward Name	Building Score
84	01	1950	Etobicoke North	83
85	01	1960	Etobicoke North	87
86	01	1962	Etobicoke North	87
469	01	1960	Etobicoke North	90
470	01	1960	Etobicoke North	91

Table 2: Sample of Average Building Scores per Ward

Ward	Ward name	Average building score
01	Etobicoke North	89.02632
02	Etobicoke Centre	91.23214
03	Etobicoke-Lakeshore	89.21774
04	Parkdale-High Park	86.50000
05	York South-Weston	84.76522

2.3 Average Building Scores per Year Built

Similarly, to the Average Building Scores per Ward, I employed the summarise function to calculate the mean score per year and then arranged the average scores per year using the cleaned data set (see Table 3).

2.4 Ward Profiles (25-Ward Model)

This data set, published by City Planning (Data 2021), contains census data from the 2021, 2016 and 2011 Census of Population, including demographic, social and economic information for each Ward in the City of Toronto.

Table 3: Sample of Average Building Scores per Year Built

Year built	Average building score
1838	89.00000
1885	91.33333
1888	90.00000
1890	77.00000
1891	92.00000

Table 4: Sample of Average Total Income per Ward

Ward	Average Total Income
1	95200
2	146600
3	127200
4	127200
5	88700

Table 5: Average Building Scores per Average Total Income per Ward

Ward	Ward Name	Average Total Income	Average Building Score
01	Etobicoke North	95200	89.02632
02	Etobicoke Centre	146600	91.23214
03	Etobicoke-Lakeshore	127200	89.21774
04	Parkdale-High Park	127200	86.50000
05	York South-Weston	88700	84.76522

Based on previous research and data assumptions, the only variable relevant for this analysis is Income and Shelter Costs of 2021. Also, it is important to highlight that there are multiple income statistics and averages within the data set. For this paper, I have selected the *Average total income of households in 2020 ($)* row per ward and discarded all other columns and rows (see Table 4).

2.5 Merged Average Building Scores and Average Total Income of Households per Ward

Finally, I used the combine function to merge the average building scores per ward data and the average total income per ward (Table 5).

3 Results

3.1 Average Total Income of Households per Ward

The average household income level per ward is $120,096, with a standard deviation of $33,980.64. The wards with the highest household income are Don Valley West (Ward 15) at $ 224,800, Eglinton-Lawrence (Ward 8) at $ 176,400, and University-Rosedale (Ward 11) at $ 174,800. On the other hand, the wards with the lowest average household income are

Humber River-Black Creek (Ward 7) at $ 85,700, York South-Weston (Ward 5) at $ 88,700, and Toronto Centre (Ward 13) at $ 89,400 (see Figure 1).

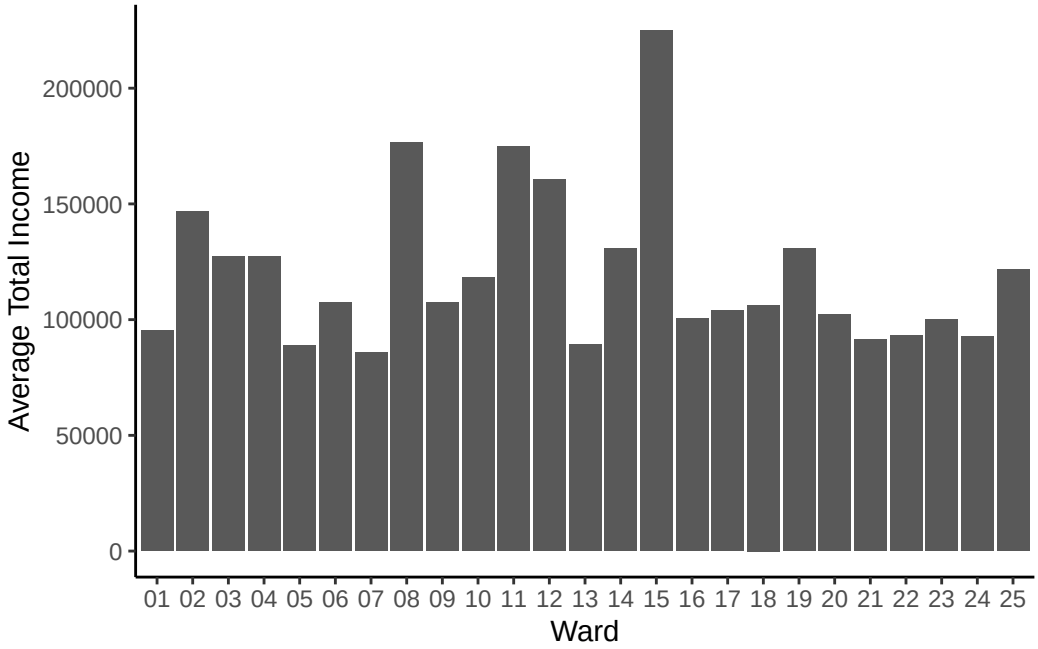

Figure 1: Average Total Income per Ward

3.2 Comparing Average Building Scores with Average Total Household Income per Ward

As seen on the scatter graph Average Building Scores vs. Average Total Household Income per Ward (see Figure 2), there is a slightly negative correlation between the average apartment building scores and the average total household income per ward, with a correlation coefficient of -0.1765207.

The top three wards with the highest average building scores are Scarborough-Rouge (Ward 25) with 95.75%, Scarborough North (Ward 23) with 91.68% and Willowdale (Ward 18) with 92.5%. And they have an average total household income, of
$121,800, $100,000 and $106,300 respectively.

Contrarily, the bottom three wards with the lowest average building scores are Beaches-East York (Ward 19) with 80.77%, University-Rosedale (Ward 11) with 83.65% and University-Rosedale (Ward 9) with 84%. And they have an average total household income, of $130,600, $174,800 and $107,300 respectively.

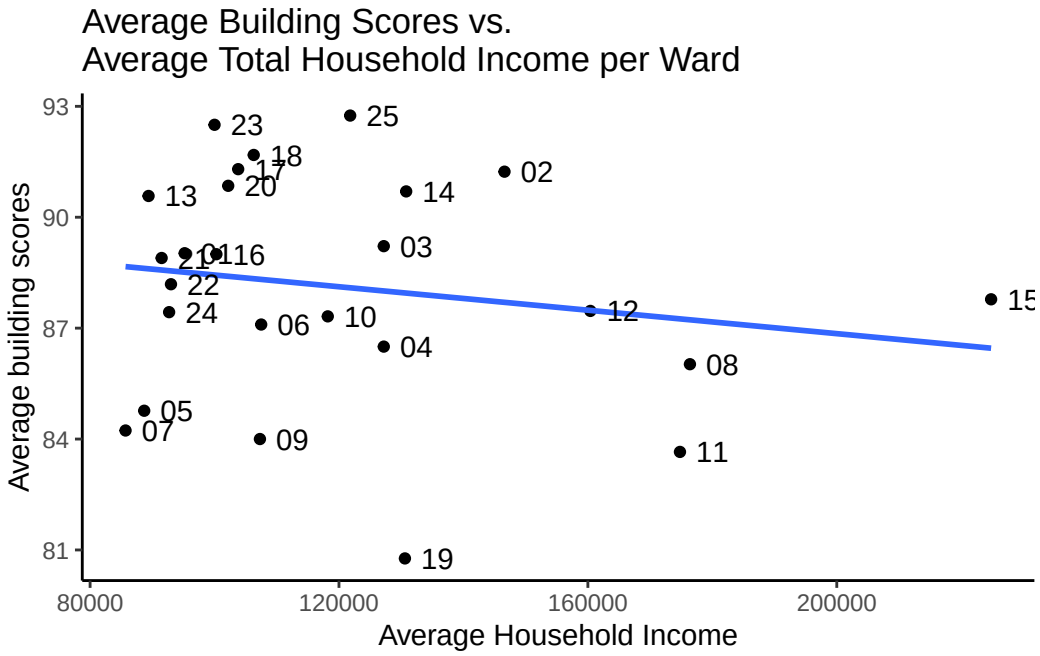

Figure 2: Correlations between average building scores vs. ward income

3.3 Average Building Scores per Year Built

The average building score per year built is 87.12%, with a standard deviation of 6.53%. The years with the highest average building score are 2017 with 99.75%, 2021 with 99.14% and 2020 with 99%. On the other hand, the years with the lowest average building score are 1922 with 64%, 1924 with 67% and 1898 with 68% (see Figure 3).

3.4 Comparing Average Building Scores with Year Built

As seen on the scatter graph Average Building Scores vs. Year Built (see Figure 4), there is a slight positive correlation between the average apartment building scores and the year built, with a correlation coefficient of 0.6075294.

4 Discussion

After analyzing Figure 2, the data reveals that wards with lower incomes do not necessarily correlate with lower apartment building evaluation scores. The negative correlation coefficient of -0.1765207, reveals that there is a very weak opposite correlation between ward incomes and apartment building evaluation scores, lower-income wards seemingly have higher average

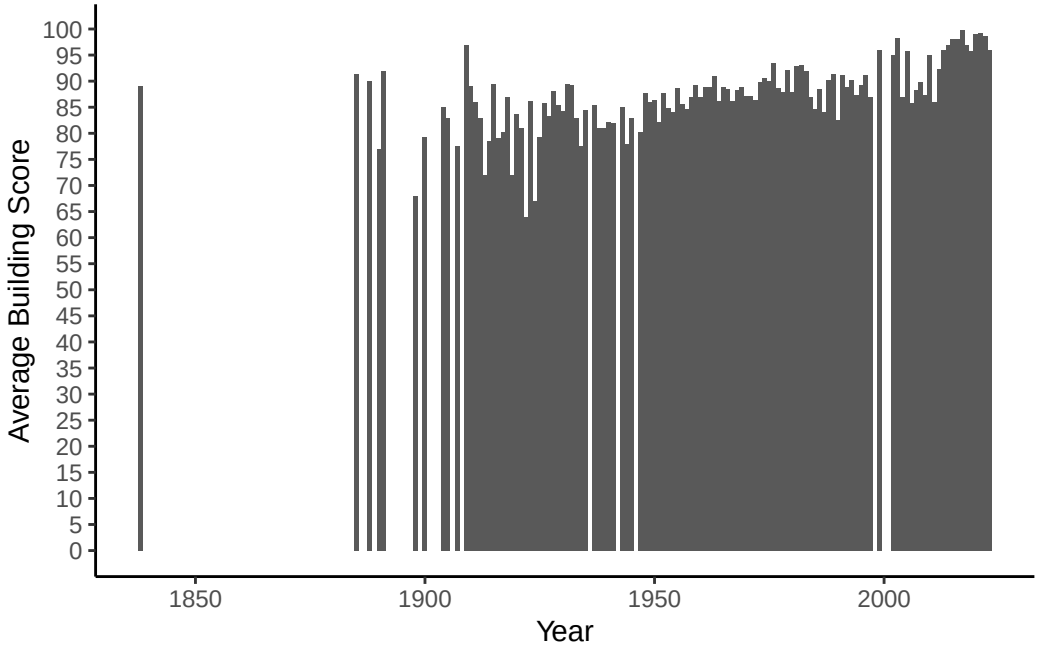

Figure 3: Average Building Scores per Year Built

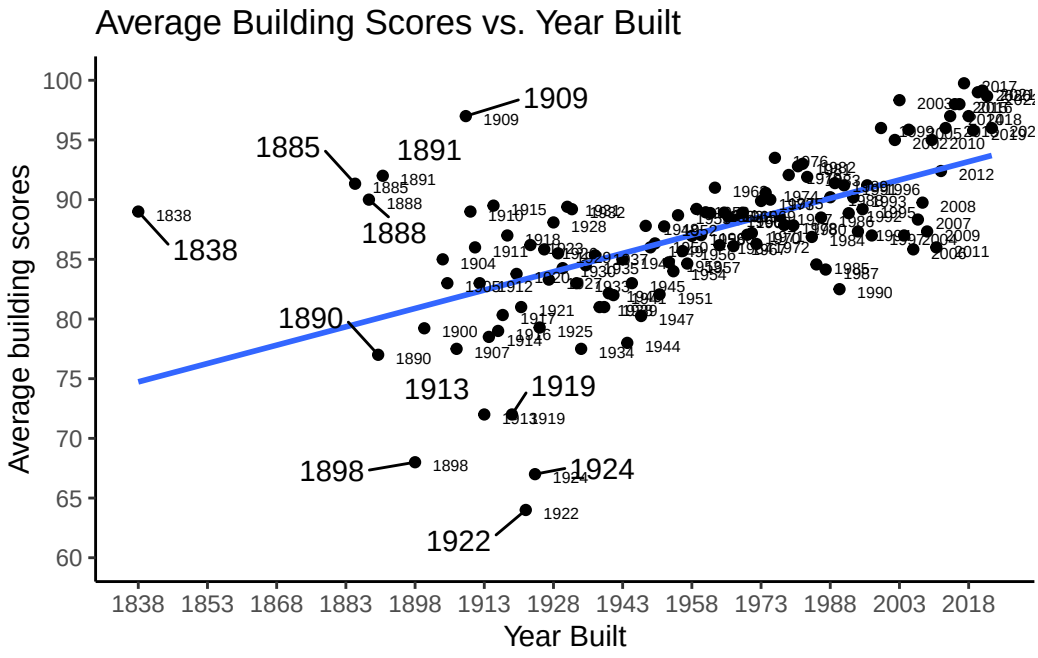

Figure 4: Correlations between average building scores vs. year built

building scores. The ward with the highest average building score Scarborough-Rouge (Ward 25) with 95.75% has an income of $121,800, which is lower than the ward with the lowest average building score Beaches-East York (Ward 19) with 80.77% and an income of $130,600. But as previously mentioned, this correlation is very weak according to BMJ (The BMJ: leading general medical journal. Research. Education. 2020) as the correlation coefficient remains in the 0-0.19 range.

Conversely, after analyzing Figure 4, the data reveals older buildings do correlate with lower apartment building evaluation scores. The positive correlation coefficient of 0.6075294, reveals that there is a strong correlation between the year the apartment buildings were built and apartment building evaluation scores. The year with the lowest average building score is 1922 with 64%, while the year with the highest average building score is 2017 with 99.75%. The BMJ (The BMJ: leading general medical journal. Research. Education. 2020) deems this correlation as strong since the coefficient remains in the 0.6-0.79 range.

Overall, the results of this analysis partly match the data assumptions initially made on how lower-income wards and older apartment buildings would obtain lower apartment building evaluation scores. This analysis shows that lower-income wards have a very weak tendency to get higher evaluation scores compared to higher-income wards. Yet, older apartment buildings have a strong tendency to get lower evaluation scores compared to newer apartment buildings.

The findings of this analysis potentially challenge the idea that people living on low income might be affected by unfit housing conditions (Phipps 2018). However, they potentially reinforce the idea that aging housing stock might be a causative factor of these unfit housing (Phipps 2018).

The analysis conducted in this essay merely aimed to establish a relationship between these variables. It is important to consider that correlation does not imply causation. Even though there may be a correlation, it does not necessarily mean that changes in one variable cause changes in the other. Additionally, correlation only measures linear relationships and may not capture more complex patterns or dependencies between variables. Therefore, this analysis is inadequate to provide potential causal factors for these relationships.

Likewise, this analysis presents certain limitations stemming from data selection and potential measurement errors. Firstly, the size of the property was not taken into consideration; some apartments may have 3 floors, while others may have 50, which can significantly impact the evaluation results as they could be categorized as different types of housing. Secondly, average score data from the average score per year-built data set had to be excluded during the correlation calculation because it analyzed scores for N/A years, representing unknown data.

5 Conclusion

In conclusion, the analysis of the 2023 Apartment Building Evaluation Scores in Toronto sheds light on the relationship between lower-income wards, older apartment buildings, and their respective evaluation scores. Contrary to initial assumptions, the data indicates a very weak correlation between lower-income wards and higher evaluation scores, challenging the notion that lower-income individuals consistently face unfit housing conditions. On the other hand, a strong correlation between older apartment buildings and lower evaluation scores supports the idea that aging housing stock is a contributing factor to suboptimal living conditions.

Future steps should involve more comprehensive research and analysis, considering additional variables like property size and specific characteristics of evaluated areas. This in-depth research could offer a deeper understanding of socioeconomic factors that interplay with housing conditions and could contribute to the development of regulations that ensure healthy housing for all.

References

Anonen, Laura. 2021. "City Launches Interactive Webpage to View Apartment Building Evaluation Results." https://www.councillorpaulafletcher.ca/city_launches_interactive_webpage_to_view_apartment_building_evaluation_results.

Clarke, Erik, Scott Sherrill-Mix, and Charlotte Dawson. 2023. *Ggbeeswarm: Categorical Scatter (Violin Point) Plots*. https://github.com/eclarke/ggbeeswarm.

Data, Toronto Open. 2021. "Ward Profiles (25-Ward Model)." https://open.toronto.ca/dataset/ward-profiles-25-ward-model/.

———. 2023. "Apartment Building Evaluation." Municipal Licensing & Standards. https://open.toronto.ca/dataset/apartment-building-evaluation/.

Firke, Sam. 2023. *Janitor: Simple Tools for Examining and Cleaning Dirty Data*. https://github.com/sfirke/janitor.

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

Müller, Kirill, and Hadley Wickham. 2023. *Tibble: Simple Data Frames*. https://tibble.tidyverse.org/.

Phipps, Erica. 2018. "Towards Healthy Homes for All: RentSafe Summary and Recommendations." RentSafe Project Team. https://rentsafecanada.files.wordpress.com/2018/10/rentsafe-summary-report_final.pdf.

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

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

The BMJ: leading general medical journal. Research. Education., The BMJ |. 2020. "11. Correlation and Regression." The BMJ | The BMJ: leading general medical journal. Research. Education.

Toronto, City of. 2023. "RentSafeTO – Evaluations & Audits." City of Toronto.

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

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

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

Wickham, Hadley, Romain François, Lionel Henry, and Kirill Müller. 2023. *Dplyr: A Grammar of Data Manipulation*. https://CRAN.R-project.org/package=dplyr.

Xie, Yihui. 2014. "Knitr: A Comprehensive Tool for Reproducible Research in R." In *Implementing Reproducible Computational Research*, edited by Victoria Stodden, Friedrich Leisch, and Roger D. Peng. Chapman; Hall/CRC. http://www.crcpress.com/product/isbn/9781466561595.

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