Toronto apartment building evaluationanalysis*

Annie Yan

February 3, 2023

Many people encountered problems with apartment selection. In order to provide beneficial suggestions in apartment selections, this report analyzed apartment building evaluation scores under the examination of RentSafeTO relate with location, building age and number of place evaluated. Evaluation data used in this report are fully collected by RentSafeTO.

1. Introduction

According to Statistics Canada study, 1.9 million Canadian households lived in condominium units. (Scrinko (2023)) The number of apartments increased by 30% to 430,080 households. Apartment buildings accounted for 68% of newly occupied units in Toronto.(Toronto (2021)) The demand for apartment buildings is significantly growing at the time. Many people encountered problems with apartment selection. In order to make beneficial suggestions in apartment selections, this report analyzed apartment building evaluation scores under the examination of RentSafeTO. Several other factors are also considered, including location, building age, and the number of areas evaluated.

Apartment Building Standard is a bylaw enforcement program established in 2017 that ensures that owners and operators of multi-family dwellings with 3 or more floors or 10 or more units comply with building maintenance standards. ("Open Data Dataset" (n.d.)) The system RentSafeTO disclosed its evaluation data on the Toronto open data portal. It is valuable for us to interpret.

*Code and data are available at: https://github.com/AnnieYan0807/Toronto_Apartment_Building_Evaluation_Analysis.

2. Data

2.1 Data Source

The majority of the data used in this report are drawn from the "Apartment building evaluation" database. This database was collated and published by Toronto's open data portal under Municipal Licensing & Standards. The latest update of the database drawn in this report is on Jan 31, 2023.

2.2 Data Collection

The database includes evaluation scores for buildings registered with RentSafeTO. RentSaftTO is a system works to ensure that owners and operators of apartment buildings in Toronto meets building maintenance standards through initiatives like evaluation. (Toronto (2023)) This database contains building registered earliest as 2017 and latest until 2023.

Evaluation scores recorded in this database are fully collected by RentSafeTO. Bylaw Enforcement officers will inspect several items including common areas, mechanical and security systems, parking and exterior grounds. A score of one to five will be assign with one being the lowest and five being the highest. These detailed scores is also available in the database. If an item is not applicable to the building at the time of inspection, the score will appear blank in the dataset. All buildings must undergo evaluation at least once every three years. (Toronto (2023))

2.3 Data Analysis

To analysis this dataset, R programing language (R Core Team (2020)), tidyverse (Wickham et al. (2019)), janitor (Firke (2023)), dplyr (Wickham et al. (2023)), and opendatatoronto (Gelfand (2022)) are used. ggplot2 (Wickham (2016)) has been used in order to produce graphic reports. The use of knitr (Xie (2014)) helped me to generate tables.

Credit to RentSafeTO and Opendatatoronto, this dataset allows me to take a pike into the correlation between building evaluation scores and several other aspects. These findings, though still have limitations, may be able to help for people who want to find a nice apartment in the future.

The dataset has 40 variables in total that described apartment building conditions. But, for the purpose of this report, we will only utilize a small amount of them. The official explanation of these variables are listed below: ("Open Data Dataset" (n.d.))

- YEAR BUILT: This is the year that the building was built in. Information is provided by the Building Owners/Managers.

- PROPERTY TYPE: This field informs users of whether a building is owed privately, by Toronto Community Housing Corporation (TCHC) or another assisted, social or supportive housing provider.
- WARDNAME: This is the name of the ward. All data is provided based on the 25 ward system.
- SCORE: This is the overall score of the building. The score is the sum total of each item that was evaluated. The formula to calculate scores is as follows: sum of all assigned scores during the evaluation / (number of unique items reviewed *5)
- NO OF AREAS EVALUATED: This is the number of items that were evaluated during a single evaluation.

Evaluation scores, being one of the essential variables in this database, best illustrate the condition of the apartment buildings. As all the correlation analyses will be conducted related to evaluation scores, I created a graph to give an overview of the entire score distribution. Figure 1, the count of scores in each ward, presents the score distribution with color-coded wards. From figure 1, we can see the highest pike of score count is around 75, with a repetition of approximately 500 times. The lower bond ends at around 37.5, excluding outliners. The higher bond is 100.

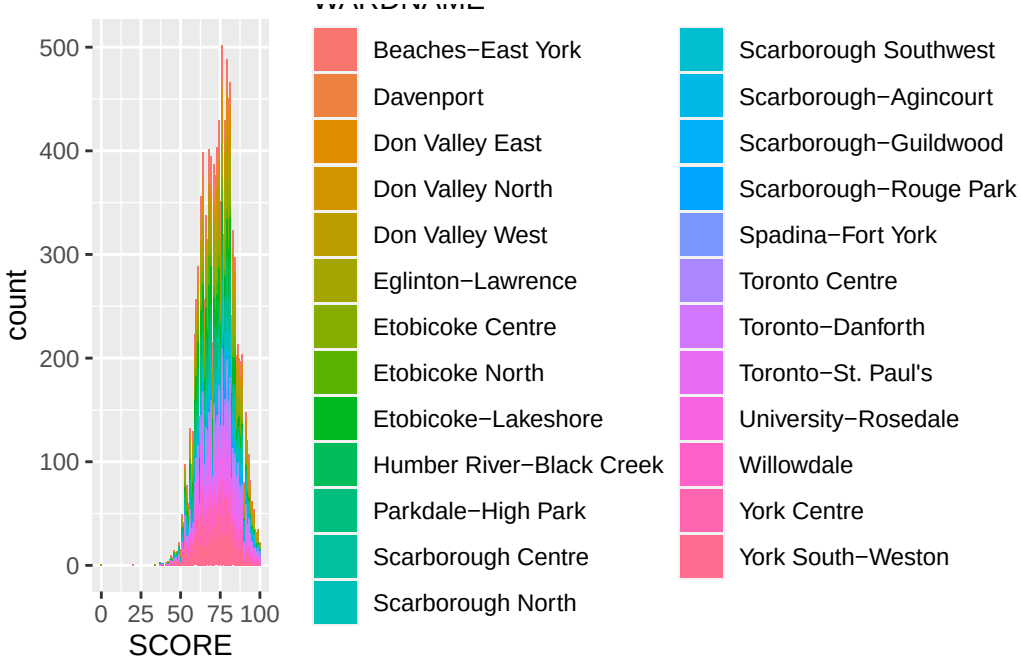

Figure 1: The count of score in each wards

However, due to a large amount of data, the measurement is not very clear in the graph. Thus, I calculated the mean, median and standard deviation of scores. Similar to figure 1 shown, the

median of all scores is 74. The mean of all scores is 73.77, with a standard deviation of 10.55. Above are some measurements of scores of the entire population. These measurements will be helpful in later analysis.

Table 1: Measurement of entire population

MEAN	MEDIAN	STANDARD_DEVIATION
73.77084	74	10.54963

The first aspect I examine with its relationship to building evaluation scores is ward location. As shown in Table 2 below, there are 14 wards in total. Scarborough North has the highest mean evaluation score of 84.56. That is 10.79 above the mean score of the entire population. Don Valley North, in second place, has an average building evaluation score of 82.19, which is 8.42 above the mean score of the population. Ward Scarborough-Agincourt has the third-highest mean score of 80.44. From this table, we may suspect that these three wards have a relatively higher standard of apartment buildings.

In contrast, Humber River-Black Creek (69.85), Etobicoke North (69.92), and Davenport (69.97) are the three wards that have the lowest average scores. They are all significantly lower than the mean score of the entire population of 73.77. That may suggest apartment buildings in these three areas will be a bigger chance of having a low evaluation score.

Table 2: Mean score of different wards

WARDNAME	MEAN
Beaches-East York	73.52333
Davenport	69.97394
Don Valley East	77.44356
Don Valley North	82.18817
Don Valley West	78.16279
Eglinton-Lawrence	73.37678
Etobicoke Centre	73.79029
Etobicoke North	69.91935
Etobicoke-Lakeshore	72.39573
Humber River-Black Creek	69.84553
Parkdale-High Park	70.62267
Scarborough Centre	76.08200
Scarborough North	84.56000
Scarborough Southwest	73.80000
Scarborough-Agincourt	80.43885
Scarborough-Guildwood	74.51968
Scarborough-Rouge Park	76.35000

WARDNAME	MEAN
Spadina-Fort York	76.66667
Toronto Centre	73.33982
Toronto-Danforth	74.27842
Toronto-St. Paul's	75.03517
University-Rosedale	73.18345
Willowdale	79.49306
York Centre	72.25320
York South-Weston	71.53828

Apart from location, the age of the building is also a factor that may affect evolution scores. Due to the distribution of figure 2: Evaluation score related to year apartment built, we have a reason to suspect that there is a correlation between the year buildings were built and the evolution scores. A reasonable hypothesis for the relationship between these two variables is the early the apartment buildings were built, the lower the evolution scores, and vice versa. However, this is an untested hypothesis. A later test is needed to confirm the such correlation.

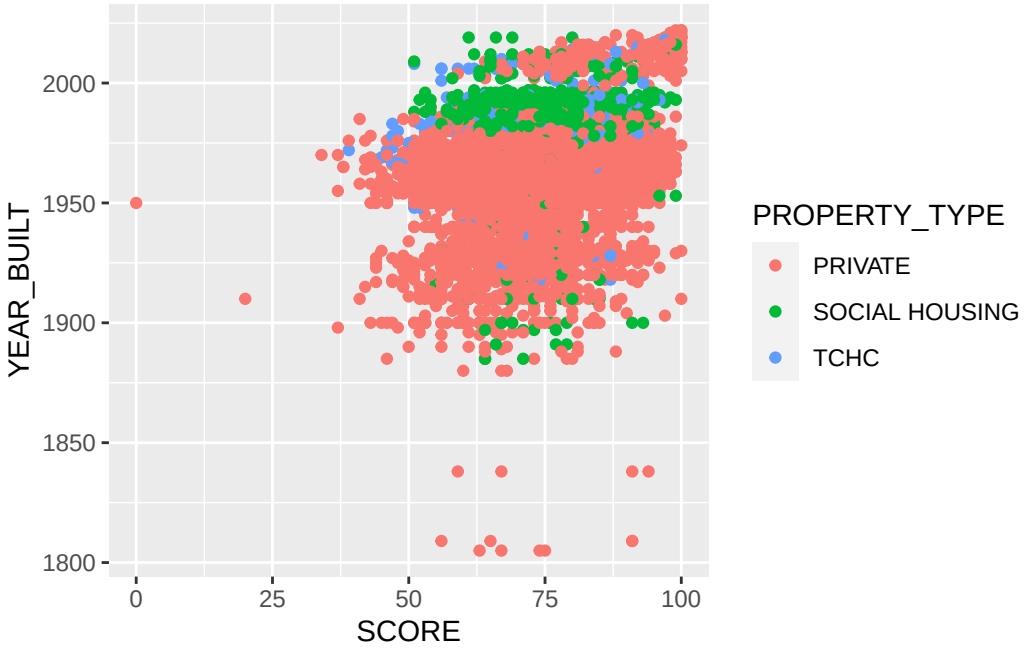

Figure 2: Evaluation score relate to year apartment built

Last but not least, I examine the no of areas evaluated. According to figure 3, apartment buildings will most likely have 18 areas for evaluation, including the entrance lobby, security,

laundry rooms, garbage chute rooms, and others. The majority of the apartment building will have 15 to 19 areas that need to be evaluated. Only a small number of apartment buildings will have 0 to 12 areas to evaluate.

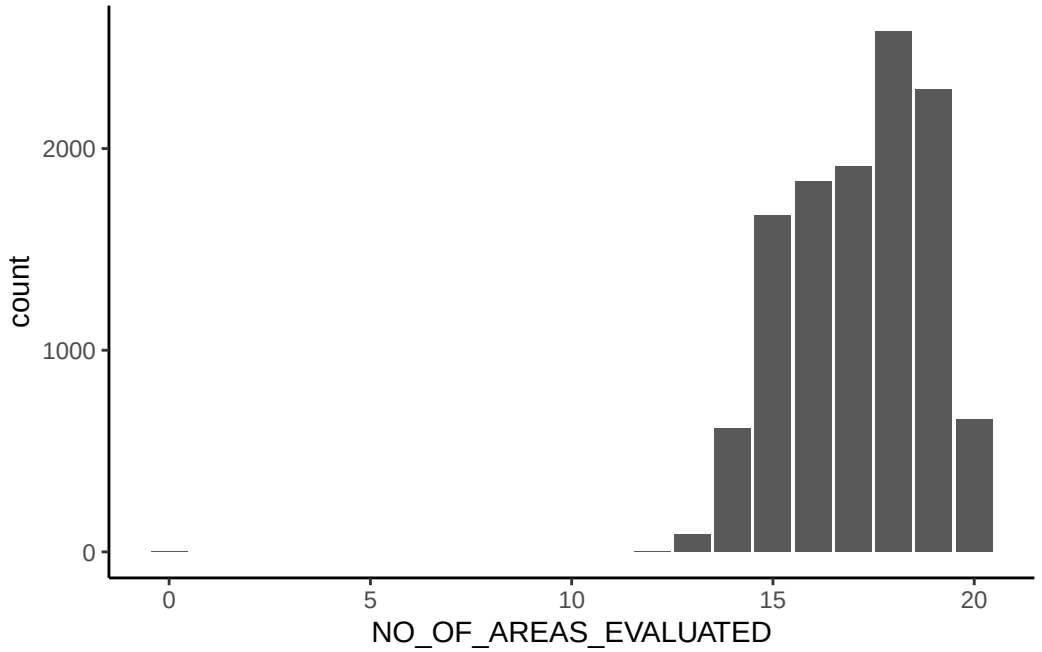

Figure 3: Count of number of areas evaluated

3. Result

Overall, the dataset "Apartment building evaluation" is helpful in this report to analyze some potential factors that could affect an apartment building evaluation score. This information, though further testing is still needed, is worth considering in future apartment selections. In conclusion, among the 14 wards in Toronto, Scarborough North, Don Valley North, and Ward Scarborough-Agincourt will most likely have a higher quality apartment building under the standard of RentSafeTO apartment building evaluation. On the other hand, apartment buildings located in Humber River-Black Creek, Etobicoke North, and Davenport may have a larger chance of having lower quality. Other than that, building age may have a negative correlation with buildings evaluation scores. Apartment buildings that have been built in recent years tend to have higher evaluation scores. Lastly, an apartment building that has around 15 to 18 areas in the list of areas RentSafeTO evaluating is at a normal range. Apartment buildings that have a lower number of areas evaluated may have fewer facilities. This information will be helpful in selecting future apartments.

4. Limitation and improvement

As frequently mentioned above, this report does have limitations and restrictions. Firstly, the dataset only includes registered from 2017 until 2023. Bias may exist in the data collection process. Secondly, this dataset is solely provided by the RentSafeTO system. The score does not fully represent the quality of the building. The score only explains the performance of this apartment building under the standard of the RentSafeTO examination. Thirdly, further testing is needed for the results of the analysis. To testify the correlation between the year the building was built and the evolution scores, correlation coefficient calculation, and T-test needs to be conducted. Fourthly, due to the lack of information, we can not estimate the relationship between the number of areas RentSafeTO evaluated and the level of facilities. Lastly, correlation does not mean causation. The variables mentioned in this report may or may not affect the evolution scores of the apartment buildings. To confirm causation, an experiment needs to be conducted in a controlled environment. In the end, although these limitations and restrictions exist, the analysis results of this report are still worth considering in future apartment selections. It may be helpful for people to choose a high-quality apartment building under the standard of RentSafeTO.

References

Firke, Sam. 2023. *Janitor: Simple Tools for Examining and Cleaning Dirty Data*. https://CRAN.R-project.org/package=janitor.

Gelfand, Sharla. 2022. *Opendatatoronto: Access the City of Toronto Open Data Portal*. https://CRAN.R-project.org/package=opendatatoronto.

"Open Data Dataset." n.d. *City of Toronto Open Data Portal*. https://open.toronto.ca/dataset/apartment-building-evaluation/.

R Core Team. 2020. *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing. https://www.R-project.org/.

Scrinko, Jordon. 2023. "What Percentage of People Live in Condos in Toronto? ([y])." *Precondo*. https://precondo.ca/what-percentage-of-people-live-in-condos-in-toronto.

Toronto, City of. 2021. "Housing Occupancy Trends." *City of Toronto*. https://www.toronto.ca/city-government/data-research-maps/research-reports/planning-development/housing-occupancy-trends/#:~:text=The%20number%20of%20households%20in,a%20total%20of%2098%2C150%20households.

———. 2023. "Rentsafeto for Tenants." *City of Toronto*. https://www.toronto.ca/community-people/housing-shelter/rental-housing-tenant-information/rental-housing-standards/apartment-building-standards/rentsafeto-for-tenants/.

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, Romain François, Lionel Henry, Kirill Müller, and Davis Vaughan. 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.