Evaluation of Apartment Building Conditions in Toronto: Ratings and Complaint Trends Across Neighborhoods

Xuanang Ren

2024-09-26

This paper explores the Toronto Apartment Building Evaluation dataset to assess the state of apartment buildings across the city. The analysis focuses on the relationships between building evaluation scores, complaints, and building age, examining how these factors vary by neighborhood. There are significant variations in complaints based on the buildings' age, with older buildings generally having more complaints. Additionally, there are differences in neighborhood-level, observed in evaluation scores, suggesting areas of focus for city housing improvement..

Table of contents

1 Introduction

Toronto's housing infrastructure has key components such as apartment buildings whose conditions directly impact the well-being of its residents. Therefore, it is essential to understand how factors like neighborhood, building age, and maintenance affect apartment conditions in improving housing quality and ensuring safe living environments.

The dataset used in this analysis is the Toronto Apartment Building Evaluation dataset. It contains evaluation scores, building ages, and complaint data for over 3,400 buildings. This analysis aims to explore how these variables relate to one another and the varying insights they provide about housing conditions in different parts of the city. Specifically, the analysis investigates how the number of complaints varies between newer and older buildings and whether there are significant differences in building evaluation scores across neighborhoods.

The rest of the paper has the following structure: Section 2 describes the dataset used and major variables. Section 3 contains the analysis and results compounded with various visualizations of trends in evaluation scores and complaints. Section 4 entails a discussion of the findings and their implications. Lastly, Section 5 briefly discusses the limitations and remedies based on the analysis.

2 Data

2.1 Raw and Cleansed Data

The dataset used in this analysis is the Toronto Apartment Building Evaluation dataset, obtained from Open Data Toronto. The dataset contains detailed information on apartment buildings' conditions across Toronto (Gelfand (2022)). This information was collected through inspection evaluations conducted by city inspectors. Cleanliness, structural integrity, and maintenance are some of the factors used in rating each building. Furthermore, the dataset consists of 3,454 observations based on the following key variables:

Building age: The age of the building is the difference between the year built and the year evaluated.

Evaluation Score: This is a numeric score representing the building's overall condition. A higher evaluation score indicates better conditions.

Neighborhood: This variable represents the location of each building categorized by wards in Toronto.

Complaints: This variable represents the number of issues reported by tenants for each building.

The data analysis was performed using R (R Core Team (2023)). The analysis also entailed the use of the following R packages: tidyverse (Wickham et al. (2019)), ggplot2 (Wickham (2016)), janitor (Firke (2023)), and knitr (Xie (2023)).

Some of the data points had missing attributes whereby an "NA" was put in place of the true value. Before any analysis, these entries were removed entirely during the process of data cleaning to simplify the process.

2.2 Summary Statistics

Variable	Mean	Std.Dev	Min	Max
Building Age (Years)	45.6	23.8	10	100
Evaluation Score (0-100)	87.2	6.5	65	100
Number of Complaints	1.8	2.3	0	10

These variables provide essential information about the condition of apartment buildings across Toronto. The average age of the buildings in the dataset is 45.6 years (SD = 23.8 years), indicating a wide variation in building ages. The evaluation score which measures the overall condition of each building, has an average score of 87.2 (SD = 6.5). Lastly, the average number of complaints per building is 1.8 (SD = 2.3). The high variability in complaints suggests that some buildings may be facing chronic problems, while others are well-maintained with few or no issues reported.

# 3 ANALYSIS AND RESULTS

3.1 Evaluation Scores by Neighborhood

Evaluation scores vary across Toronto's neighborhoods. Below is a box plot of evaluation scores by neighborhood.

Apartment Ratings by Neighborhood
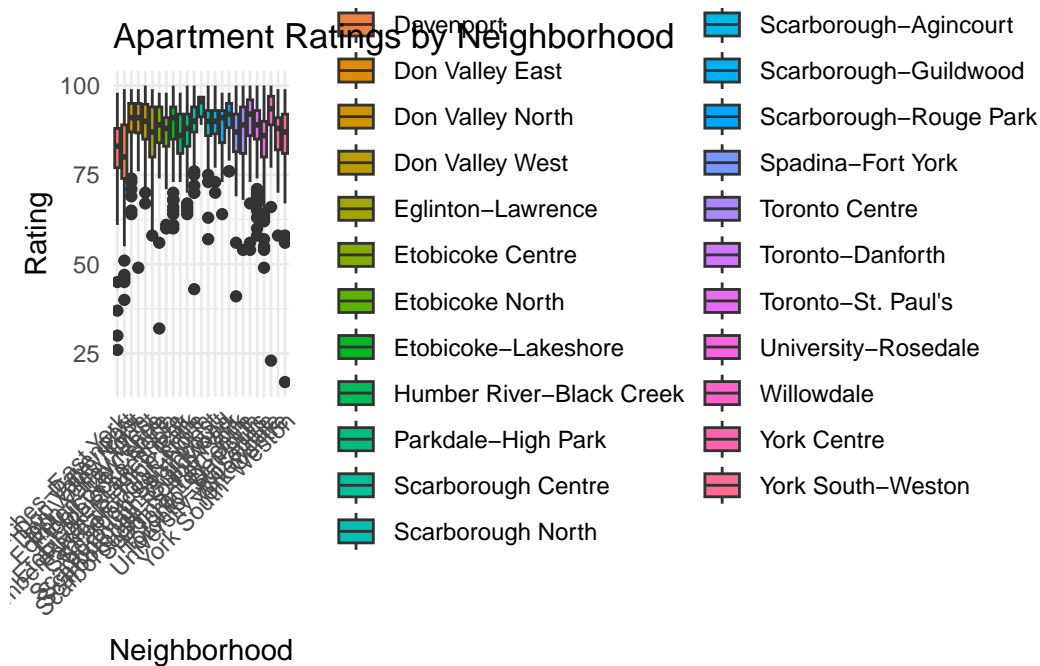
Beaches-East York, Davenport, Don Valley East, and Don Valley North appear to have the highest median ratings and relatively small IQRs, suggesting consistently high-rated apartments in these areas. On the other hand, Scarborough-Rouge Park, Scarborough Centre, Scarborough North, and Scarborough Southwest have lower median ratings and larger IQRs, indicating more variability in ratings within these neighborhoods. There is also a greater range of apartment quality evident from the outliers common in neighborhoods with larger IQRS.

### 3.2 Distribution of Evaluation Scores

The distribution of apartment ratings in Toronto suggests that most apartments are of good quality. The histogram below shows that the majority of ratings are concentrated around 95.

Distribution of Apartment Ratings
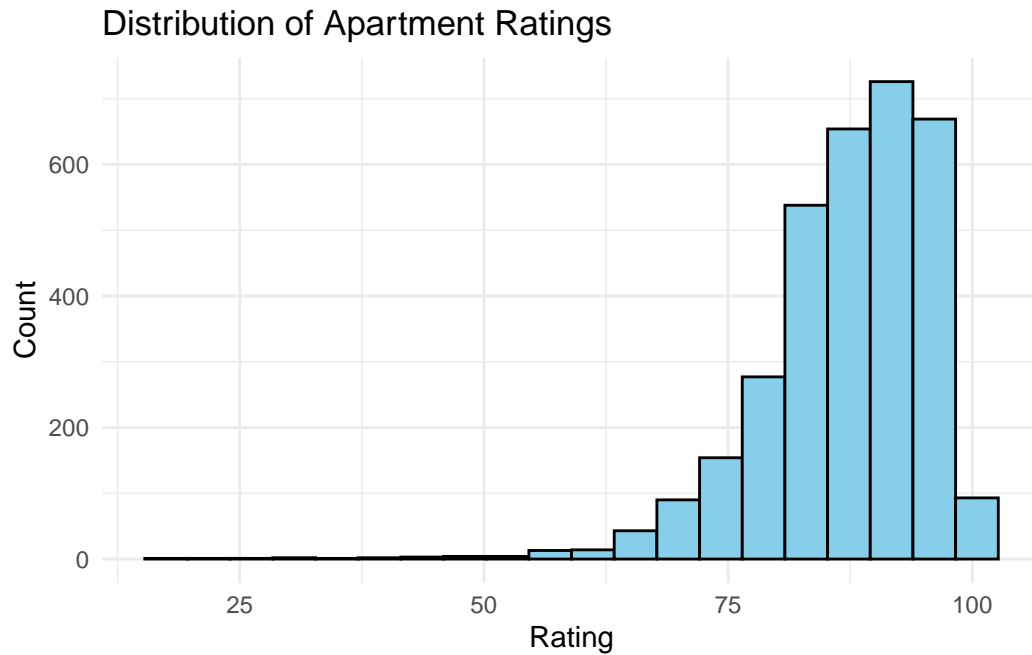
The figure suggests that apartment ratings tend to be high (concentrated around 95). The rightward skew indicates that a majority of apartments received ratings above the average (between 75 and 100) suggesting they are considered of good quality. However, the tail in Figure 2 also indicates the presence of apartments with lower scores

3.3 Relationship Between Building Age and Complaints

There is a positive correlation between building age and the number of complaints. Older buildings generally receive more complaints than newer ones.

Relationship Between Building Age and Number of Complaints
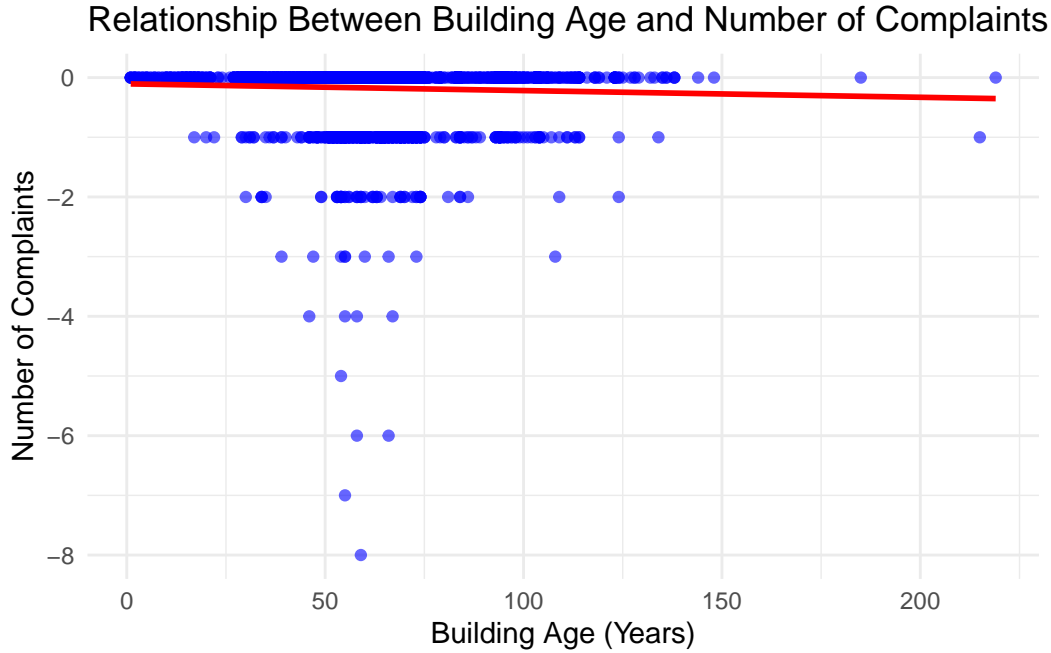
There is a positive correlation between the two variables plotted in Figure 3. Figure 3 shows that older buildings receive more complaints than the newer ones. For instance, buildings over 50 years old exhibit a higher frequency of complaints compared to newer buildings.

4 DISCUSSION

The Toronto Apartment Building Evaluation dataset analysis gives details into the relationship between the number of complaints and building age. It suggests that older buildings face more maintenance issues highlighting the need for increased attention to aging apartments in Toronto wards.

The evaluation scores differences between neighborhoods suggest that some areas have better-maintained buildings than others. Neighborhoods with higher scores suggest possible greater investments. In contrast, the neighborhoods with lower scores, suggest that they may require more attention to improve housing conditions from policymakers.

The evaluation scores of apartments vary, as shown in Figure 1. Neighborhoods like Beaches-East York, Davenport, Don Valley East, and Don Valley North have the highest median ratings with smaller interquartile ranges (IQRs), indicating consistently high-rated apartments. In contrast, neighborhoods such as Scarborough-Rouge Park, Scarborough Centre, and Scarborough North show lower median ratings and larger IQRs, reflecting more variability in apartment quality. Figure 2, a histogram, reveals that most apartments have high ratings, with a concentration around 95, indicating good overall quality, though a small number have lower scores.

Figure 3 highlights a positive correlation between building age and complaints, showing that older buildings, particularly those over 50 years old, tend to receive more complaints compared to newer buildings, suggesting potential maintenance or quality issues in aging structures.

Other factors such as renovation history and building management practices are important to note since they also likely contribute to the differences in complaints and scores. For a deeper understanding of building conditions, future studies could incorporate these variables.

5 LIMITATIONS AND REMEDIES

One limitation of this study is that it does not account for all potential factors affecting building conditions, such as renovation history and quality of management. Additionally, the dataset lacks granular information about the types of complaints received, which could further enlighten the specific issues faced by tenants.

Future research could address these limitations by including additional data sources, such as building management records or tenant satisfaction surveys. Furthermore, a longitudinal analysis of changes in building conditions over time would inform on the long-term effects of maintenance and renovation efforts.

REFERENCES

Firke, Sam. 2023. "Simple Tools for Examining and Cleaning Dirty Data." https://cran.r-project.org/package=janitor.

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

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

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 McGowan, Romain François, Garrett Grolemund, et al. 2019. "Welcome to the Tidyverse." *The Journal of Open Source Software* 4 (43): 1686. https://doi.org/10.21105/joss.01686.

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