A Spatial look at Toronto Apartment Building External Maintenance*

Finn Korol

February 4, 2022

Abstract

The city of Toronto, Ontario features an eclectic building landscape. This paper seeks to analyze the spacial distribution of exterior qualities of apartment buildings relating to upkeep as they have been evaluated by Toronto bylaw. Visualizing subscores for exterior qualities reveals that upkeep is evaluated to be the worst in the Southwest corner of the city, and best in the Northeast. A relationship might exist between these subscores and building age with middle-aged buildings receiving the lowest subscores.

Table of contents

*Code and data are available at: https://github.com/korolodf/Apartments_Toronto

1 Introduction

Toronto architecture is difficult to describe outside of a "well, it's there" sort of way. In other words, architecture critics and laymen alike typically are unimpressed by Toronto's cityscape (Hume 2005). While rent in the city skyrockets, in part due to housing demand, it is interesting to look back. Toronto has responded to population booms by spitting out apartment buildings (Era 2018). The effect is that dating apartment buildings offers a rough estimate for a time in the city's history where the infrastructure was not meeting the needs of its residents. This slapdash approach to urban planning has contributed to a utilitarian building typology which is better appreciated for its eclecticism than for its beauty and innovative design (Era 2018). What becomes of the city built without a well thought-out plan? Is Toronto a "nice" looking city based upon how its buildings mature and are maintained?

This small realization of Toronto's approach to urban design, while not necessarily related to the features of apartment building data that this this paper seeks to explore, was motivational toward this exploration. As a case study of city featuring a utilitarian approach to residential urban design, I will explore data on evaluation of exterior conditions of apartment buildings. I aim to answer the question: How does the upkeep of apartment buildings vary spatially in Toronto? In addition, I will explore the distribution of ages of buildings in the city and plot upkeep against age to address a second question: have the exteriors of older buildings become more deteriorated? I utilize Gelfand (2022)'s code to view qualities of apartments across the city and my analysis is conducted using R Core Team (2022).

2 Data

2.1 Acquisition

The data used was provided by opendatatoronto, and was published by Municipal Licensing & Standards through RentSafeTO. Initial observations of the set begin in 2017 and have been completed through to the end of January, 2023. Data was loaded and saved using Wickham et al. (2022) and Wickham, Hester, and Bryan (2022), and directories were managed and called upon using Müller (2020).

2.2 Features

The data contains 11651 observations, beginning in 2017, with each row populated with the data collected upon an individual evaluation of an apartment building in Toronto. The buildings included are those under the jurisdictions of RentSafeTO, which are those either with greater than 3 storeys or 10 units. Evaluations are performed at least once every 3 years. A score out of 100 (included as a column in the set) was given to each building upon inspection, which was made up of subscores out of 5 (also included). The subscores represent evaluations of features including the building lobby, entrance door and windows, security system, stairwells, laundry rooms, handrails and guards, garbage chute rooms, garbage collection area, elevators, lockers and storage areas, walls and ceilings and floors, interior lighting levels, graffiti, exterior cladding, grounds, walkways, balcony guards, the external penetrative ability of water, parking area and unspecified other facilities. Other included data were the building RSN (an identification number), location specifiers (to be elaborated upon in 2.3), years built, registered and evaluated, the date of evaluation, the result of the evaluation (based upon the score, including audit and the length of time before next evaluation).

The original data set contained 40 column variables, which were reduced to 7 for relevance to this analysis. Many original data were redundant in this context, for example by including the apartment's ward numerically and by name in addition to grid ID locations and in a separate x and y set of coordinates. For compatability reasons with the tools used in this analysis, longitude and latitude were ideal, and because they were included as well, the rest of the location specifiers were removed from the set. For relevance to this analysis, subscores were reduced to just those related to the exterior of the building, also removing those which were not applicable to all buildings (e.g. parking lot) in the interest of keeping the set as large as possible and avoiding selection bias (where buildings with parking lots might represent only the wealthiest in an area, for example). Total score was also removed. The remaining subscores were thus the building grounds, cladding, graffiti and walkways.

These evaluation data are riddled with potential for biases, which I attempted to be aware of in this analysis, and the reader should consider as well. One source of bias arises due to the bylaw officers, who are human and capable of subconsciously miss-scoring, serving as building inspectors in the collection of this data. Examples of this might include judgement passed on buildings based upon their age or location in the city. This might present as an inspector assuming a building's condition will be inferior and carrying that assumption through their evaluation by nitpicking more relative to other buildings. Alternatively, inspectors might overscore a building that they believe is in better condition than would be expected of its age or location. Inspectors might also score more strictly upon a building that has failed past inspections, or might produce poor evaluation scores to an uncooperative landlord. My hypothetical examples are speculative and not meant to be accusatory of RentSafeTO. My identification of potential biases in the dataset is not exhaustive.

2.3 Cleaning

Before removal from the set, building RSN was used to prevent single buildings from representing multiple observations. It was observed that buildings were evaluated often 3 times, so only the latest observation of a building was kept to keep this analysis as current as possible. The remaining variables, all being numbers were converted to numerical types, and null values were removed. The year each building was built was also converted to an age for a more intuitive representation of a possible relationship with subscores. Unit tests were run to ensure that all score and subscore values fell within their intended bounds. The resultant clean data frame was reduced to 3333 observations. After cleaning, the earliest evaluation was completed in September, 2017 and the latest was January, 2023. The first 5 rows of the cleaned data frame are shown below in Table 1. Data was cleaned using Wickham and Girlich (2022) and Firke (2021), with help from Wickham et al. (2022) for simplifying syntax and providing shortcuts.

Table 1: A table showing example values of the cleaned data frame. Age is given in years, the graffiti, cladding, grounds and walkways subscores are out of 5 each, and longitude and latitude are in degrees. Each row is one observation representing a single building evaluation.

age	graffiti	cladding	grounds	walkways	longitude	latitude
4	5	5	5	5	-79.40681	43.67644
134	5	3	3	4	-79.37283	43.66012
138	3	2	4	3	-79.43511	43.64945
51	5	5	5	5	-79.38143	43.66234
5	5	5	5	5	-79.38400	43.65844
1	5	5	5	5	-79.39447	43.64496

2.4 Analysis

Exploratory analysis was performed through data visualization. Revelle (2022) was used for producing simple statistical tests. Xie (2014) was used for producing tables. Kahle and Wickham (2013) was used for creating the Toronto map and plotting over it. Wilke and Wiernik (2022) and Auguie (2017) were used for stylizing plots. A line plot was also made to visualize subscores across building age.

3 Results

The features of apartment buildings across the city are summarized in Table 2. Please note that coordinates are only rounded for display and calculations were passed on the full numbers. Below, in Figure 1, see the geography of the location of this analysis, Toronto, Ontario. For reference, the window included in Figure 1 is the same as in all further map plots in this paper.

Table 2: Mean and standard deviation of table columns.

	age	graffiti	cladding	grounds	walkways	longitude	latitude
mean	61.4	4.4	3.2	3.2	3.3	-79.4	43.7
sd	19.0	0.9	0.6	0.6	0.6	0.1	0.0

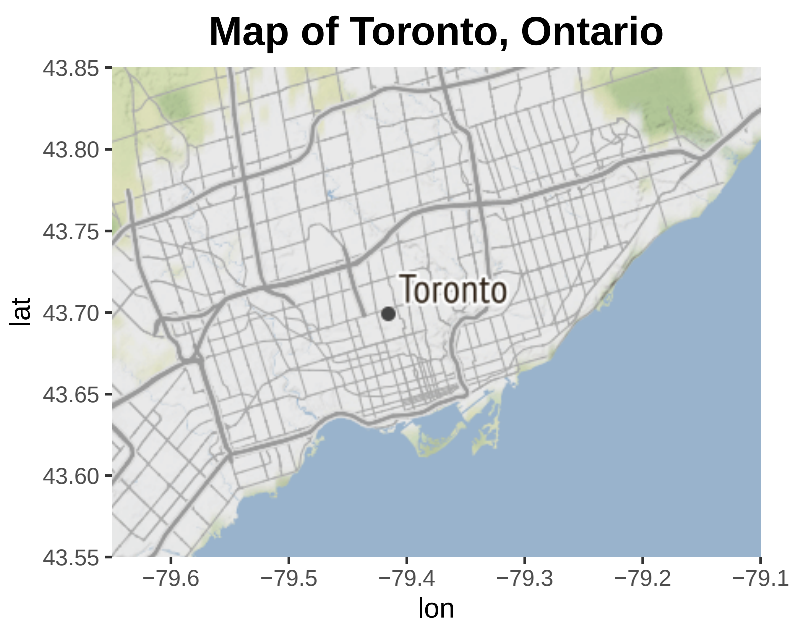

Figure 1: Map of Toronto, Ontario on longitude and latitude coordinates. The included window is rounded to the next 5 hundredths of a degree of the extreme-most values of observation.

To address my principal question (i.e. how does upkeep of apartment buildings vary spatially in Toronto), Figure 2 shows a grid of colour coded maps to show the distribution of buildings and each of the four subscores selected. This style of data representation offers a high-level view (literally), of the density of scores in an area. Figure 2 shows that all subscores are near their peak of 5 in Northeast Toronto. With the exception of graffiti, all subscores decrease toward Southwest Toronto, and graffiti remains close to peak subscore through all of Toronto.

Plotting buildings and colour-coding by age in Figure 3 revealed an approximately consistent distribution across Toronto, with ages near the mean (61.4) with an absolute range of approximately 50 years being typical. South Toronto along the waterfront appears to be the most densely populated with new builds.

Figure 4 shows the mean subscores of buildings at all ages. Graffiti remains the highest subscore at nearly all ages. A downward trend is observable in cladding, grounds and walkways before rising slightly and approximately flatlining. All subscores experience their lowest mean values for buildings between 100 and 150 years of age.

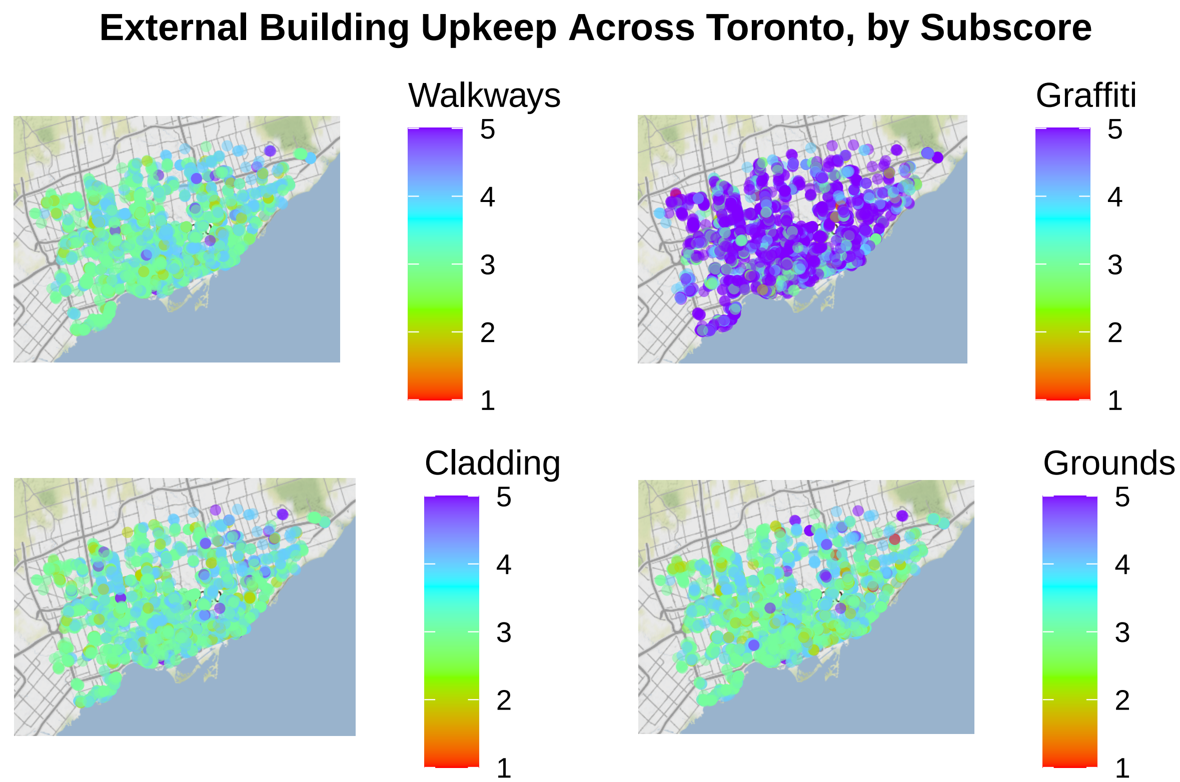

Figure 2: Toronto maps with superimposed locations of apartment buildings, colour-coded based on subscore.

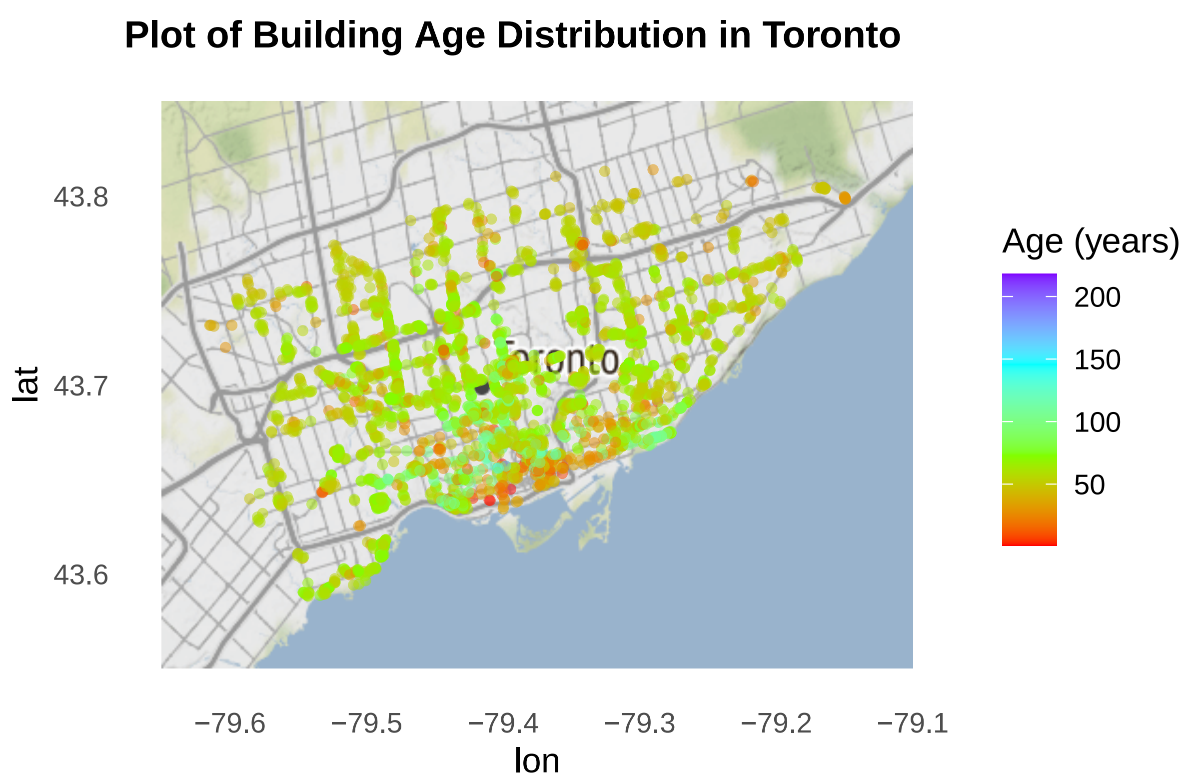

Figure 3: Toronto map with superimposed locations of apartment buildings, colour-coded by age in years.

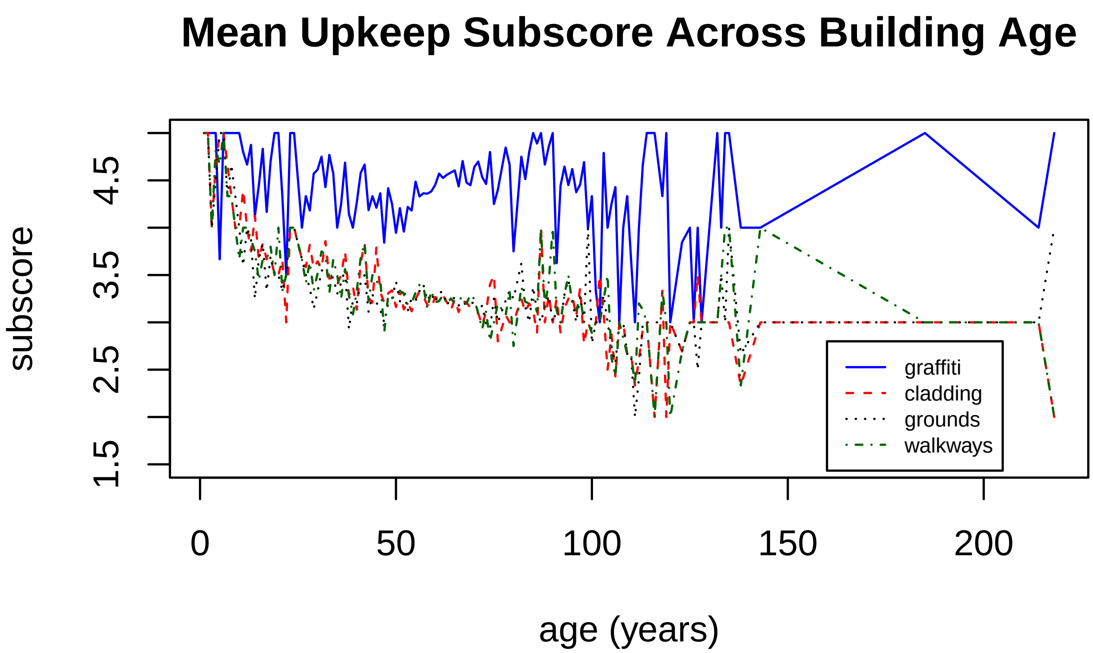

Figure 4: Mean subscores for apartment building upkeep plotted as functions of age, in years. Lines are plotted for each subscore of exterior building upkeep.

4 Discussion

4.1 Findings

Subscores other than for graffiti experience a common shared spatial trend, which might suggest a positive relationship between longitude and latitude with cladding, grounds and walkways. Graffiti subscores were high throughout the city compared with others (ranging typically from 4-5), representing the highest mean subscore at 4.4 (cladding, grounds and walkways were 3.2, 3.2 and 3.3, respectively). Reasons for these spatial trends might be due to a concentration of affluence to the Northeast or due to differential cultural densities throughout the city (these factors are likely to be related to each other). As previously mentioned, biases in collection might have resulted in this data being skewed if inspectors held judgements about parts of the city. Future research might look into these potential factors.

When mean subscores were plotted against building age, all hit troughs between 100 and 150 years. To speculate, a reason for this might have been because painting on a building that appears to be particularly new or old might be seen as high risk or undesirable to artists.

Finally, the visualization of subscores across age was flawed in that building age was concentrated near the mean. As a result, mean subscores at younger and older values were calculated using smaller samples, making them more prone to being skewed by extreme cases.

4.2 Limitations

Several limitations exist in this analysis. Due to the nature of the dataset used, certain types of apartment buildings were not represented based on RentSafeTO's guidelines; only apartments meeting the unit or storey criteria of RentSafeTO were included. Thus, a multi-apartment complex made up of small buildings might not be included. Some landlords might see this as a loophole, and not being held to municipal standards might mean that these buildings are of inferior quality due to lack of upkeep.

Another limitation was due to my chosen method of representing apartments by scatterplot. Over 3000 points were plotted over each map with visible overlap occurring. To mitigate overlooking these covered observations, I increased the transparency of points and decreased their size, however more representative methods likely exist and should be attempted in the future.

4.3 Future Research

Returning to the initial claim that Toronto's city layout is utilitarian, this exploration made me curious about the distribution of qualities in cities with more intentional urban layouts, for example in densely populated European countries. Future research might compare distributions between cities, and might examine them over time to view the trajectories of urban upkeep.

By taking the mean subscores across wards, a more readable map could be made as there would not be any overlap between polygons. In addition, viewing subscores by ward would create opportunities to explore political factors in scoring, both due to bias in evaluation (e.g. sentiments that certain groups prioritize upkeep over others) and values of these populations.

The evaluation year column in the set was null, however the evaluation date column remained intact. Future research might use the date of evaluation to observe how scores and subscores change over time. Due to the inclusion of RSN, repeat evaluations of individual buildings could be compared with one another to see if there are "good" and "bad" years for evaluation results, and could explore other features of buildings in relation to these scores.

5 References

Auguie, Baptiste. 2017. *gridExtra: Miscellaneous Functions for "Grid" Graphics*. https://CRAN.R-project.org/package=gridExtra.

Era. 2018. "Toronto Building Typology Study:" Typology Studies.

Firke, Sam. 2021. *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.

Hume, Christopher. 2005. "Toronto Architecture." *Internet Archive: Wayback Machine*. https://archive.org/web/.

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.

Müller, Kirill. 2020. *Here: A Simpler Way to Find Your Files*. https://CRAN.R-project.org/package=here.

R Core Team. 2022. *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing. https://www.R-project.org/.

Revelle, William. 2022. *Psych: Procedures for Psychological, Psychometric, and Personality Research*. Evanston, Illinois: Northwestern University. https://CRAN.R-project.org/package=psych.

Wickham, Hadley, Romain François, Lionel Henry, and Kirill Müller. 2022. *Dplyr: A Grammar of Data Manipulation*. https://CRAN.R-project.org/package=dplyr.

Wickham, Hadley, and Maximilian Girlich. 2022. *Tidyr: Tidy Messy Data*. https://CRAN.R-project.org/package=tidyr.

Wickham, Hadley, Jim Hester, and Jennifer Bryan. 2022. *Readr: Read Rectangular Text Data*. https://CRAN.R-project.org/package=readr.

Wilke, Claus O., and Brenton M. Wiernik. 2022. *Gridtext: Improved Text Rendering Support for 'Grid' Graphics*. https://CRAN.R-project.org/package=gridtext.

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.