

Income Inequality is revealed in Torontonians Apartment Maintenance*

Olaedo Okpareke

02 February 2022

Abstract

Housing maintenance benefits residents as it ensures proper functioning of amenities, maintains structural integrity, and adds value to homes. However, previous research has found that priority in maintenance for public housing apartments and lower income apartments is little compared to their more privileged counterparts. We use data from Open data Toronto to further analyze this by using plots and tables to break down apartment scores and other attributes by Ward and apartment type, using R (R Core Team 2021) and other packages. We find that apartments in higher income areas and privately owned apartments tend to have better evaluation scores than those in lower income wards and public housing respectively. This data shows the inequality in standards for housing in Toronto, as well as the need for better government accountability and housing justice. Improving the gap in apartment quality helps improve the life of poor and marginalized people.

BOXPLOTS TALK ABOUT MEDIAN < NOT MEAN. PUT THE TABLE FOR MEAN

1 1. Introduction

Toronto is a large Canadian city divided into 25 wards. In order to make sure building maintenance standards are upheld and to allow apartment residents to look up safety scores of their apartments, Rent Safe TO has provided an apartment evaluation dataset. According to a set criteria not revealed to public, this dataset rates apartment attributes according to a 5 point system, and then uses a formula to assign them an overall score over 100. In this report, we are interested in how apartment maintenance scores have changed over the years, and how attributes such as relative wealth of the ward and the type of apartment affects the total score given to an apartment.

We obtained the dataset from the City of Toronto Open Data Portal the address of this book with `opendatatoronto` (Gelfand 2020), and cleaned and analyzed the dataset using R (R Core Team 2021), `tidyverse` (Wickham et al. 2019) and `dplyr` (Wickham et al. 2021) packages. Figures and tables were created with `ggplot2` (Wickham 2016) and `kableextra` (Zhu 2021). We created tables that first showed the data and important summary statistics, then went into showing the data for area and apartment types, as well as box plots, linear model graphs and percentage bar charts, which highlighted inequality in housing evaluation scores.

Not all apartments are equal, as this report shows that although general housing maintenance has increased in recent years (OECD 2020), there is still a large difference in quality of apartments around the city. Social housing, both operated by the TCHC and by other institutions, have lower average maintenance scores and lower median security scores than their private counterparts. Apartments in some wealthier wards also have a higher average maintenance score than those in poorer wards. Private apartments also tends to have better security than their social housing counterparts. According to (OECD 2020) and (Desroches and Poland, 2021), these disparities are due to a decline in public investment.

*Code and data are available at: https://github.com/OlaedoOkpareke/starter_folder.

These scores are widely used to rate apartments ((Carter, 2004: 12)), however biased scoring and issues in measurement could create inaccurate property pricing, or otherwise give false impressions on the apartment block, and thus the type of people living there(Konadu-Agyemang et al. 1994, pg 146). Low scores for public housing also exacerbate the already low view of people who need them (Konadu-Agyemang et al. 1994, pg 146). This data and the statistics shows that vulnerable populations such as low income and minorities suffer systemic marginalization in facets of society, even in the housing sector as well maintained, secure housing is still unattainable for the city's most vulnerable, questioning the extent of the achievements of social housing. There are also concerns on the bias of the grading, and the impact of these scores on gentrification.

2 2 Data

The data for this report was gotten from the Apartment Building Evaluation (Rent Safe TO 2022) from the Toronto Open Data Portal (city of Toronto 2022) using the R package opendatatoronto (Gelfand 2020). This data is compiled by RentSafeTO and was last refreshed on January 22nd, 2022. Rent Safe TO collects data on apartments in the GTA that are 3 or more stories or had 10 more units, and response is mandatory for qualifying apartments, i.e the population. The raw data includes 9758 apartment blocks from the Greater Toronto Area and their categorical characteristics (40 variables in total) such as location and Type of property, as well as numeric characteristics such as Security, amenities and state of common areas. It also includes an overall score over 100 for each apartment block.

For data cleaning, we removed all rows with missing values as they indicated that the apartment did not have a certain room, amenity, or characteristic, but we wanted an “all things being equal” analysis, the assumption that attributes wise, they were alike. That is, we wanted to compare apartments that had poor/ lacking characteristic, not no characteristic at all. It was also helpful to see the full picture so we could see if there was bias in grading instead of it not being graded at all. The missing values seemed to be randomly dispersed about the dataset so removing them would not cause much deviation.

We included a variable if there was anything important that made it stand out for our analysis. We omitted most variables in the original set as their direct effects were not important to the analysis, and the total score variable encompassed all of their individual scores. we first removed `_id` and `RSN` as they were merely ID variables. We also excluded the ward number variable as the ward names variable were easier to identify. We included year built as we wanted to see the change in apartment evaluation between older and newer apartments. We included score as it was the main determinant of maintenance standards, as well as security to observe the the differences in security between different groups of apartments, and we wanted to see its direct effect. We then included property type, which states whether an apartment was privately owned, owned by the Toronto Housing Corporation (THCH), or owned by some other non-private institution as our analysis focuses on social/affordable housing. The final data frame included 493 observations with 5 variables.

```
rent_clean = read_csv(here::here("inputs/data/rent_clean.csv"))
```

```
## Rows: 493 Columns: 5

## -- Column specification -----
## Delimiter: ","
## chr (2): Ward, Type
## dbl (3): Score, Year_built, Security

##
## i Use `spec()` to retrieve the full column specification for this data.
## i Specify the column types or set `show_col_types = FALSE` to quiet this message.
```

We are interested in the relationships between property Type and Ward on Score and security, as well as the trend of Score over time. It shows a measure of income inequality in housing, and has possible implications for public policy. A preview of the data, (Table 1) and figure 1 using kablexrra(Zhu 2021) and ggplot2 (Wickham...) is shown below.

Table 1: Toronto Apartment Evaluation Data

Score	Ward	Year built	Type	Security
81	Toronto-Danforth	1996	SOCIAL HOUSING	5
78	Scarborough Southwest	1992	SOCIAL HOUSING	5
91	Eglinton-Lawrence	1965	PRIVATE	5
59	Davenport	1968	TCHC	4
88	Toronto-Danforth	1967	PRIVATE	5
86	York Centre	1983	SOCIAL HOUSING	5

```
x = head(rent_clean) %>%
  knitr::kable(col.names = c("Score", "Ward", "Year built", "Type", "Security"), caption = "Toronto Apartment Data", x
```

```
ggplot(rent_clean, aes(x=Score)) +
  geom_histogram(fill="black")+
  theme_minimal()
```

```
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
```

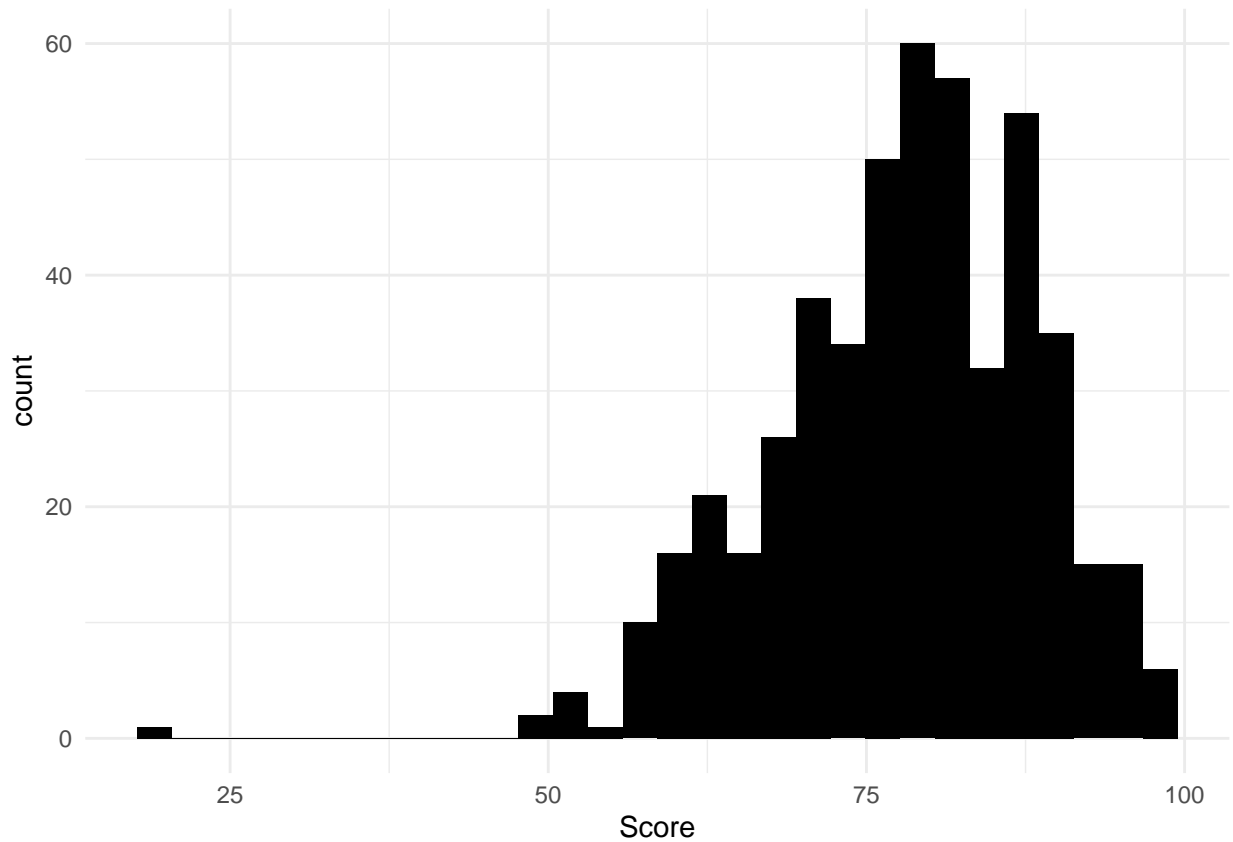


Figure 1: Table with summaries of the Score and Security for all apartments

(Figure 1) shows the the summary statistics for the buildings. We see that the data seems to be unimodal and normally distributed. There is an outlier which skews the data. We can see from the table that the average building has good standards and security, but is a bit old. The average apartment has a better score than

the median apartment, indicating that there are some apartments that are far above the average. However the mean security is lower the median, indicating some very low security apartments. Policy is needed to improve security for these apartments, and the observation that some scores are far higher than most are more thoroughly scrutinized later in the report.

(Figure 2), using ggplot2 (Wickham 2016), shows the average maintenance scores by ward. Cross referencing the income and demographics of each ward using (City of Toronto 2022) , we see that the higher income wards such as Toronto Saint Paul's and Toronto-Danforth (155,470 and 101,323 average yearly income respectively 72% white and 67% white respectively as of 2016) tend to have both higher means and medians than lower income wards, such as Davenport and Humber River-Black Creek (80,807 and 65,458 average yearly income as of 2016). An exception seems to be Scarborough North which has a low yearly average income and high racial minority population but a high average apartment maintenance score. However note that the range is very small, implying that there are very few samples, so they may not be representative of the true state of apartments in this ward. Although there are several outlier Scores for each ward, in general apartment housing is better in richer wards than in poorer wards. High income Wards also tend to have fewer racial minorities, with Toronto saint Paul and Toronto-Danforth having a 72% and 67% non-racial minority population respectively.

```
rent_clean %>%
  ggplot(aes(x = Ward,y = Score,group = Ward)) +
  theme_minimal() +
  geom_boxplot() +
  coord_flip()+
  labs(title = "Wealthier wards have better scores than poorer Wards",
       x = "Ward",
       y = "Apartment Score")
```

(Figure 3) shows the mean and median apartment scores by type. The results show that private buildings have the best average maintenance quality, and those from the Toronto housing Corporation have the lowest average. The spread of quality seems to be equal for all three housing types, so even accounting for variation private housing is consistently better than their public counterparts. This shows a disparity in priority for apartment maintenance and upkeep, as well as general housing inequality as those in social housing owned by the TCHC also deserve high average maintenance but do not, and reveals that the government are not on top of taking care of social housing. However it should be noted that (non TCHC owned) social housing maintenance is not very far behind private housing, showing a just upkeep of housing. Nevertheless according to (), people in social housing often have to take maintenance and upkeep into their own hands.

```
rent_clean %>%
  ggplot(aes(x = Type,y = Score)) +
  theme_minimal() +
  geom_boxplot() +
  labs(title = "Private Housing has better scores than social and TCHC housing ",
       x = "Apartment Type",
       y = "Apartment Score")
```

(Table 2) shows the differences in security between each housing type. We see that overall the average security scores are good for all types of apartment. However the median shows that both TCHC owned and non TCHC owned social buildings have less stellar security than that of the private buildings, so their average scores were driven up by a few very good apartments. This means that we can expect that a large number of social housing (which houses mainly low income populations) do not have the best security, which is counter-intuitive as poorer areas where this housing is often located are often in higher need of security due to correlation with crime rates. (cite poor housing live in crime areas). This could be alleviated with multi-income neighbourhoods, which places many different types of housing that can accomodate different income groups, but NIMBY-ism has slowed the progress of these neighbourhoods.

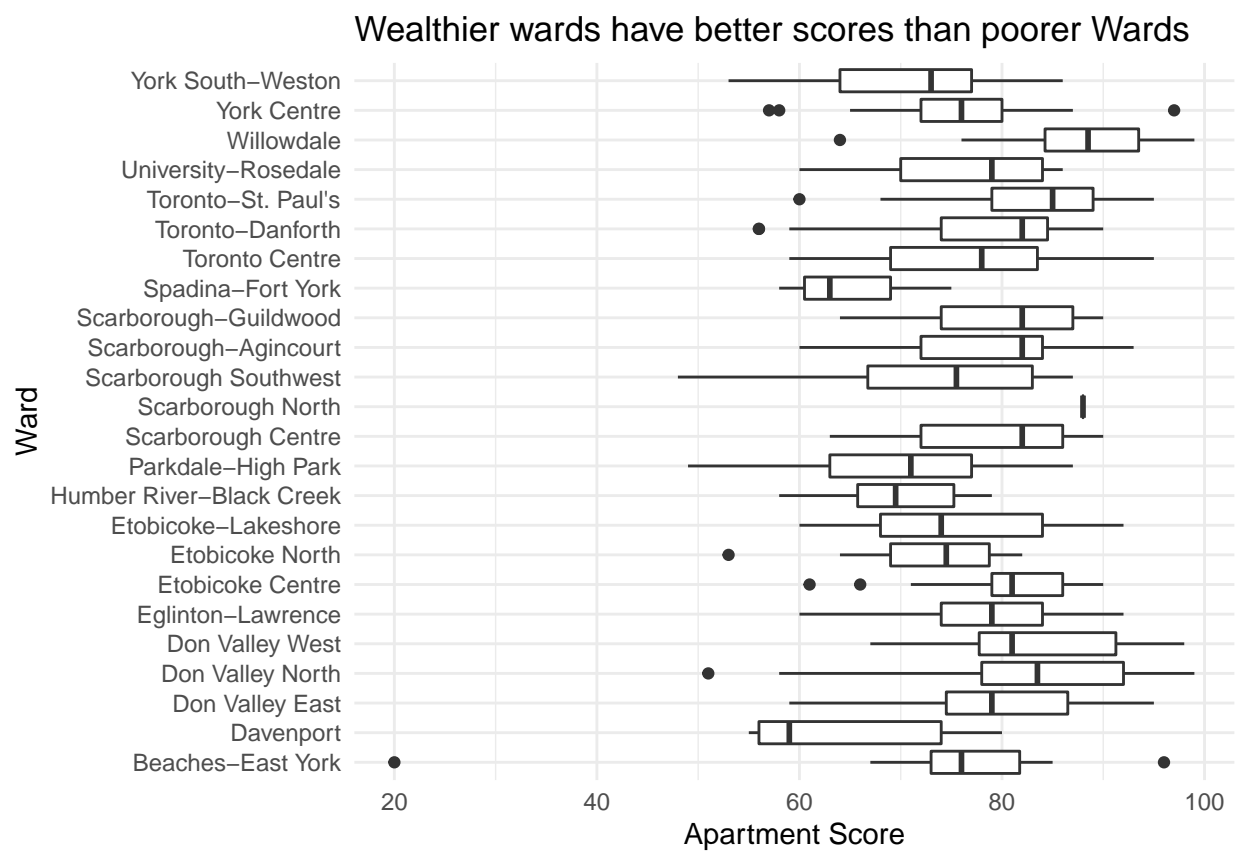


Figure 2: Boxplot of median apartment maintenance scores by ward

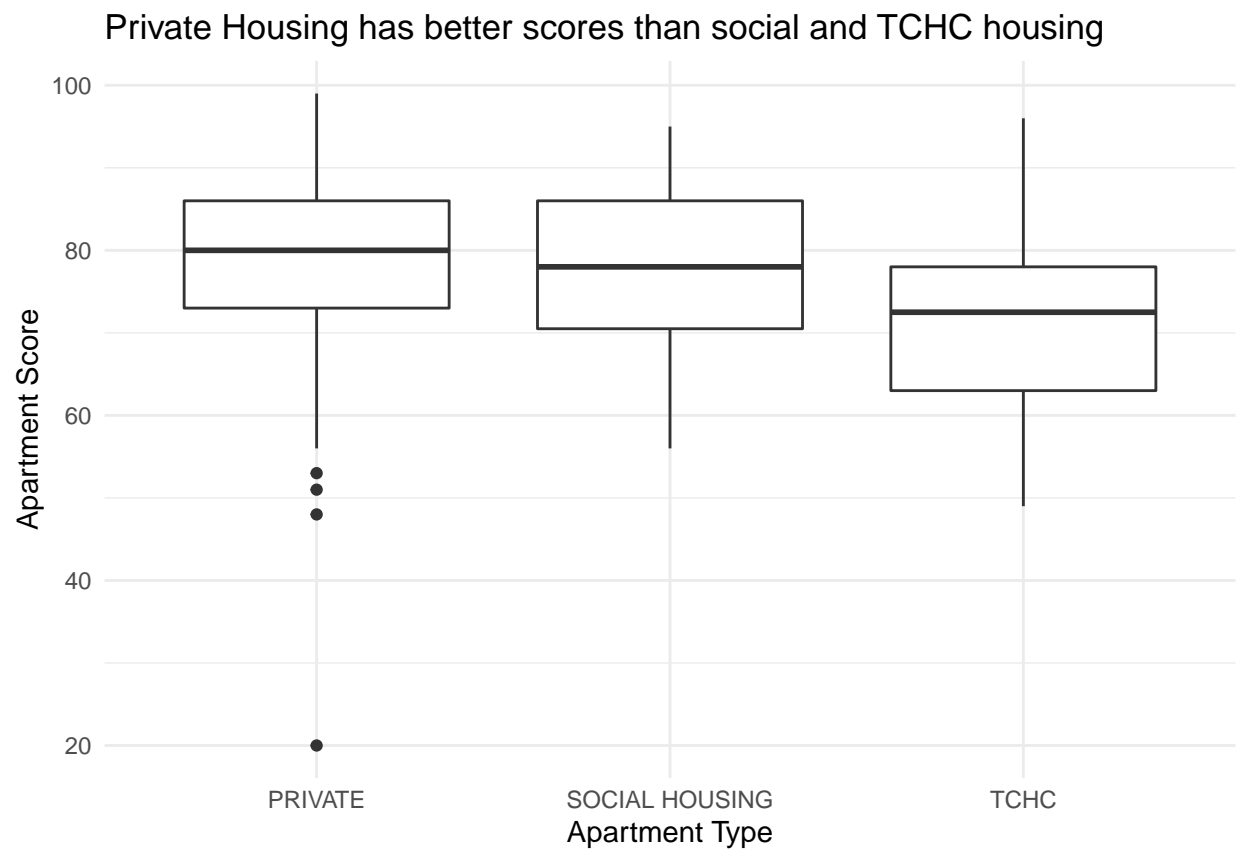


Figure 3: Boxplot of median apartment scores by Apartment Type

Table 2: Table of Mean and Median security for each group

Type	Mean	Median
PRIVATE	4.428947	5
SOCIAL HOUSING	4.237288	4
TCHC	4.111111	4

```
rent_clean %>%
  group_by(Type) %>%#
  summarize(Mean = mean(Security), Median = median(Security)) %>%
  knitr::kable(caption = "Table of Mean and Median security for each group")
```

(Figure @??fig:figfour)) shows the general trend of apartment maintenance scores for all types according to the year they were built. Even though for all three types newer buildings have better maintenance Scores than their older counterparts, private housing still gets consistently higher scores than the other two types for almost all the years. Most private apartments were built between the 1960s to 1970s while social housing was built between the 1980s to 2000s, so even the newer age of the social housing buildings is not enough to protect against lower rates of maintenance. This shows either the persistence in lack of government and public maintenance or persistent bias in scoring, which both tie into income inequality. It should be noted that the trend for private apartments may be heavily influenced by the points at the far right end, and that many private apartments also have low scores. We also removed an outlier, 1910 as it affected the accuracy of the trend for private apartments.

```
## `geom_smooth()` using formula 'y ~ x'
```

(Figure @??fig:figfive)) shows that the wealthier wards have a larger percentage of their apartments scoring above 70 compared to lower income wards. Some of the poorer wards such as Humber River Black creek or Davenport do not have very highly rated apartments at all. The racial minority population in these two wards as of 2016 was 74% and 43% respectively (City of Toronto 2022). In 2016 the income for these wards was 65,548 and 80,807 dollars respectively while in the richer wards like Toronto Saint Paul and Etobicoke Centre it was 155,470 and 128,448 (City of Toronto 2022). Many other wards follow the same trend, with the wards with higher income having larger percentages of their apartments scoring above 70, with some lower income wards even having some percentage of their apartments have a score less than 50. This is another way to see the inequality in housing standards, and communicates that housing maintenance disproportionately caters to higher income areas. This is possibly because lower income people pay less rent, are less likely to protest (August 2016), or less general funding compared to previous decades (Tsenkova 2021). note tsenkova states that less funding to build new housing

3 2.1 Results

4 Discussion

4.1 Societal Impact of these scores

As previously stated, the impact of these scores might lead to an inaccurate rating of social housing apartments or apartments in lower income areas as lower quality, thus perpetuating harmful views on low income people that further progress their marginalization in society by casting blame on social housing residents for poor conditions instead of the organizations supposed to keep maintenance (cite where poor are blamed for being dirty). Given the importance of these scores in rating apartments and housing choice, there is a need to understand the decisions that go into the scores and how they could be interpreted.

It also brings into discussion the impact of housing on well being and the needs of the low income that have less than standard housing or require lower income housing. It has been shown that public housing

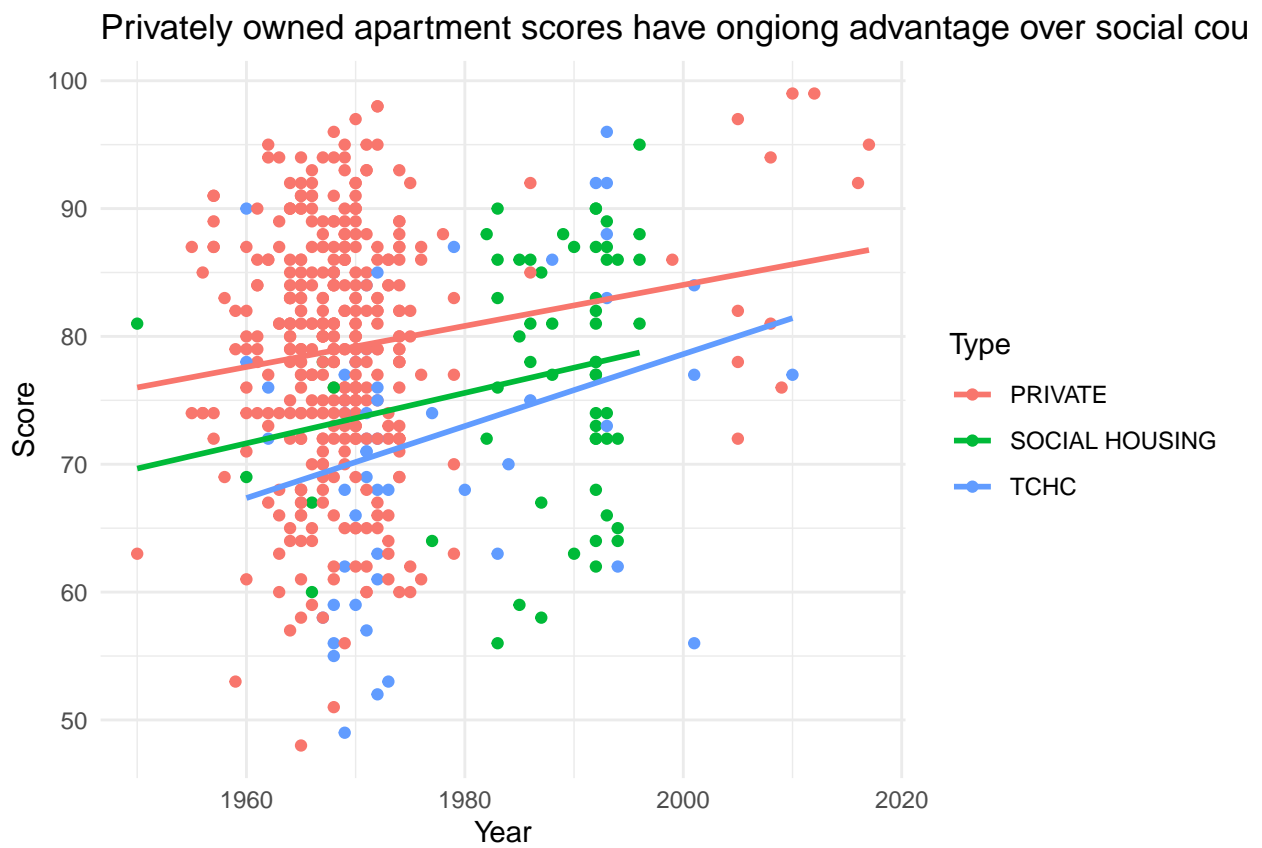


Figure 4: Scatterplots and regression line of trend of apartment scores by type over 1950 - 2021

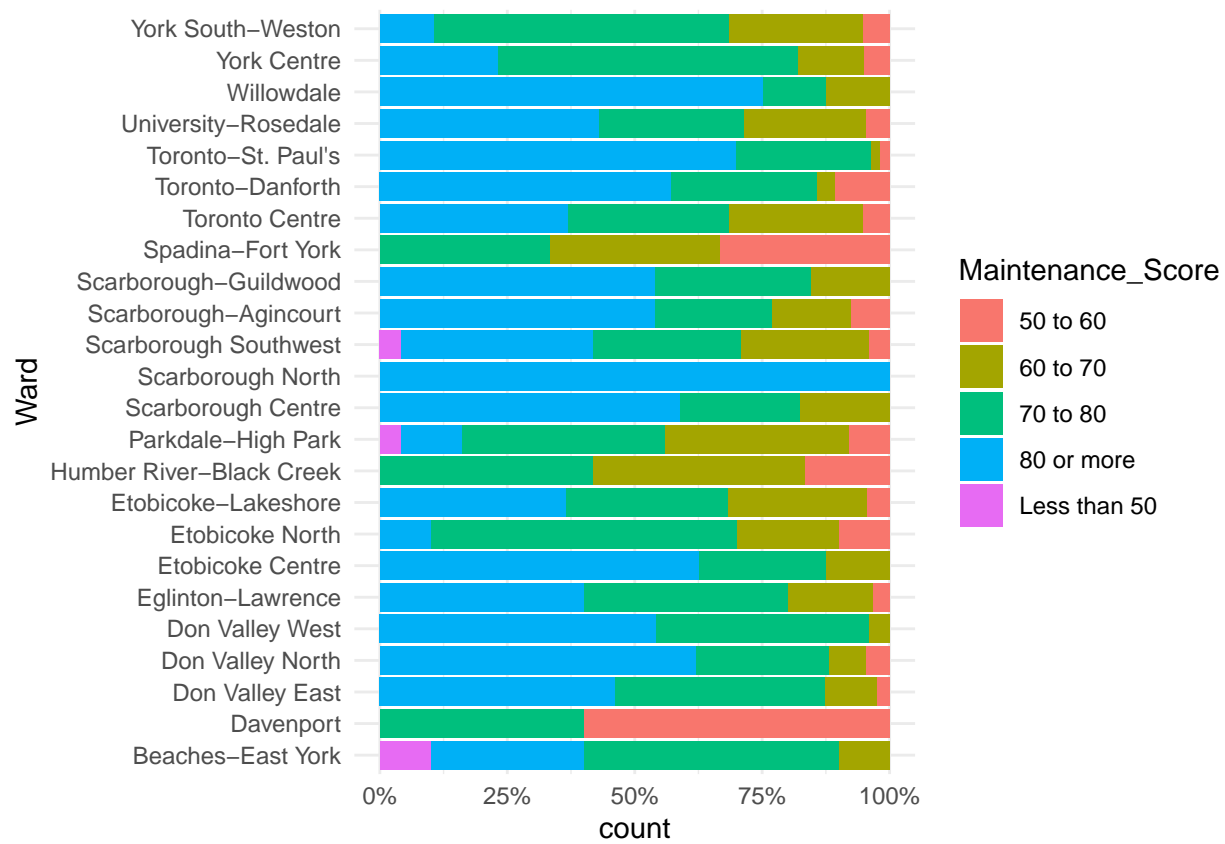


Figure 5: Percentage of apartments in each ward with that have a score in a certain group

apartments have lower quality amenities than their private counterparts, and these can have several effects on the residents. Previous research has shown that “Aspects of neighbourhoods that have been identified as having an impact on health are: ... “incivilities” such as the presence of graffiti, boarded up houses, garbage accumulation... and broken windows, and the signalling effect of these physical features that no one cares; the perception of neighbourhoods as “safe” and under effective informal social controls... Others have postulated that neighbourhood conditions directly influence habits of privacy, child rearing, house keeping, and study, which can have an indirect effect on health.”(Howden-Chapman 2004)

There is also the need to discuss the shortcomings of the Toronto Housing Corporation and why they have not been up to date with maintenance of their apartments, as well as public policy that could be formed to improve the maintenance and security of social housing apartments in low income areas. Housing for the lower income is often substandard because they are at lower priority in almost every other facet as life such as law enforcement, schooling, etc. Either because they are often seen as net negative tax payers or that they do not have as much political persuasion (August 2016) as the middle class. This intersects with a large number of poor being racial minorities, who are also often seen as less important in society. “Housing is merely one of many ways that systemic discrimination of poor folk and the lack of priority for them is shown. Lack of affordable or poor quality housing exacerbates challenges associated with low income, education, literacy, physical and mental disabilities, race, ethnicity and colour” (Carter, 2004: 25).

Regarding social housing apartments that have less security, it is paradoxical as these people often live in more crime ridden areas that needs the extra security, bringing into question if their own lives are considered and the failings of public landlords and the TCHC to provide security akin to those of private apartment. Lower income people often do not have money to protect themselves like rich, so they rely on publicly provided assistance. (cite?) This also highlights lack of security for poor in other areas like biased law enforcement. (cite)

The data showed that Scarborough North, which has a 90% visible minority population with an average income of 78,892 as of 2016, compared to the Toronto average of 100000 (cite toronto)) only has highly scored apartment buildings, an outlier. It is hopeful that average apartment scores for other wards with similar incomes increase in the future. However low income residents have concerns that a high level of security or good apartments in an otherwise lower income area can make the area more attractive to wealthier people, leading to gentrification, which pushes the original low income inhabitants out ((Evans, 2007: 6,7), Hulchanski,2006: Table – Change in Average Individual Income, 1970 to 2000, City of Toronto)(“It’s all about power and you have none:” The marginalization of tenant resistance to mixed-income social housing redevelopment in Toronto, Canada,), which is a problem with the already tenuous social housing for these marginalized groups.(Hay,2005: 1)(cite actual)

4.2 Bias and ethical Concerns

There are some areas for possible bias in the data which perpetuate preconceived notions on marginalized groups. As Rent Safe TO does not release its criteria for scoring, it is unknown what specific attributes they look for while grading. Measurement issues could occur as there is not a way to objectively measure an apartment (i.e it is unclear what distinguishes a rating of 4 from a rating of 5). It is thus possible to then give the an apartment attribute of the same quality a higher score in a wealthier ward/private building with other good amenities, whilst giving apartments in poorer wards/social housing a lower score due to bias against the general area or the social status of residents and not the apartment itself. This perpetuates biases against the lower income and minorities as people gets their previously held beliefs of the poor status of those communities confirmed (cite where people believe minority areas look worse). Lower scores are prevalent in poorer wards and wards with high visible minority populations, and biases against these people and the general “feel” of an area may influence people to rent a more expensive apartment with good scores when there’s a cheaper one that was scored wrongly. On the other side, being in a more affluent ward may artificially prop up a mediocre apartment’s score, thus inflating prices and making it less affordable even though its maintenance is lacking. There is thus a need for a more transparent grading system so that outsiders can cross check and make sure than no historic inequalities perpetuate themselves through scoring biased against poor or social housing folk.

4.3 Collection problem

There are possible data collection problems with this data set. RentSafeTO collects data only for apartment blocks with 3 or more stories and ten or more units(cite), meaning that some buildings that are in dire need of maintenance may be left behind as the organization simply doesn't collect data for them, even though most people may see them as apartments (i.e a three story block of 9 flats). This leaves out a big piece of the full picture on apartment quality, and may also be why some known wealthier wards are getting higher average scores (the low income housing there does not count) and the effect of income inequality as poor people are more likely to live in such types of "apartments" than the wealthy (cite). For apartments that did qualify, the grading is a scheduled event organized with the landlord ahead of time (cite where this is stated on the website). This means that scores presented may not be entirely accurate for general upkeep of the building as landlords who neglect buildings can hurriedly make affairs before the evaluator/enforcement officer arrives. It may thus be necessary for graders to come up semi unannounced to get the true maintenance levels of the building for a more accurate scoring.

4.4 Income

It is important to note that not all living in private apartments are wealthy. In fact, there are many private apartments inhabited by low income people so this disparity in scores between private and social housing may be due to reasons other than income inequality. As stated earlier, less governmental priority to social apartments may be a major reason (Tsenkova 2021). The average income for each ward was an important variable as it would put other information into context, as it determines the priority in government spending not just in apartments but in general as lower income areas get less spending but wealthier income are often full of funding in housing, surroundings, and other outside amenities that could have gone into the scoring bias. Unfortunately this variable was not included in the data set, so information on income had to come from a less up to date source.

(should these be in next steps?)Private housing is now in itself getting harder to afford due to rising housing costs, forcing many so called middle class people into the need for social housing, or forcing them into lower rated apartments. It would be interesting to see the effect of rising prices and the movement of people that would have been considered high income into apartments in lower income wards or social housing, on total maintenance scores and security, to see if the government would then divert funding to these.

4.5 Weaknesses and next steps

We should note that not everyone who lives in private housing is wealthy, and it is important not to overlook those who are struggling to pay rent to access those improved amenities. Not all those who live in private areas or wealthier wards are rich, wealth wards still have poor areas so their own experiences were neglected.

As average income was not provided for the data set, we had to collect external data which was set in 2016 and may not be accurate as it is not current to the modern day.

While the N/As were dispersed randomly throughout the data set, removing them may have caused bias in the data set so our means, medians and analysis of Scores may have been different if the N/A values had been filled. Also, removing variables even if they were already present in the Score variable may have caused incorrect analysis as there may be some important information in the rejected variables that either brings a new perspective or aids one of the accepted variable in explaining the relationships between apartment type, ward, and maintenance scores.

For next steps we could use the geo-location variable that was omitted for this report to see if the scores for apartments on boundaries between poorer or richer wards had the negative bias of the poor wards or positive bias of the rich.

We could also compare the income and scores of buildings including those that did not have all amenities to further observe if there was inequality as apartments in poor areas tend to have complete absence of some amenities.

5 References

1. OECD (2020), “Social housing: A key part of past and future housing policy”, Employment, Labour and Social Affairs Policy Briefs, OECD, Paris, <http://oe.cd/social-housing-2020>

It’s all about power and you have none:” The marginalization of tenant resistance to mixed-income social housing redevelopment in Toronto, Canada

<https://www.toronto.ca/city-government/data-research-maps/neighbourhoods-communities/ward-profiles/>

Howden-Chapman PHousing standards: a glossary of housing and healthJournal of Epidemiology & Community Health 2004;58:162-168.

https://link.springer.com/chapter/10.1007/978-3-030-69563-7_2