Toronto Apartment Evaluations Are Hard to Predict*

Ritvik Puri

06 February 2022

Abstract

Building evaluations are always important to conduct, not only from a safety perspective, but also to promote people buying and renting more property. The RentSafeTO program started in 2017, and the data is updated daily on Open Data Toronto, and we will be looking at the various factors that play in the role of giving a better score. Over the last 5 years, we have seen a generally upward trend in building scores. We also find that there are certain features, like multiple stairwells and higher security, that tend to procure a extra points.

1 Introduction

The city of Toronto introcued a new bylaw enformenent program in July 2017 that is used to ensure that the apartment building owners and operators comply with the neccesary building maintanence standards, known as RentSafeTO. These standards are applied to all apartment buildings with 10 or more units or with 3 or more stories. (REF)

Owners of such apartments are required to register with RentSafeTO as well as mainitan the standards defined by this program. Tentants need to conatct their respective landlord incase they face an issue. These issues could either be vital, such as heat or hydro fault, or service requests, such as window flaws or common area cleaning. If the landlord does not comply to these requests, then according to the bylaws of RentSafeTO program, legal action can be taken against them.

Each property that falls under the program gets insepected by an officer and receives an evaluation score, which is made available to not only the landlord and tenant, but also to the potential tenants. If the score of a building is 86 or above, it will be evaluated again in the next three years. If it is between 66 and 85, it will be evaluated again within two years. If it is between 51 and 65, it will be evaluated again within a year. If a building gets a score of 50 or below, then then the full building will undergo a comprehensive inspection. (REF)

Home safety evaluations are neccesary to protect residents from potential hazards which may lead to personal injury if left unchecked. In a city like Toronto, where a lot of people do not own their personal property and live in rented apartments, it is critical that the landlord is kept informed of their building's condition so that they can ensure a tenants safety.

What has been the results of introducing such a system? Has it been beneficial? Are there any specific features in a building that can give it more points and result in a better score?

2 Data

We have taken Apartment Evaluation Score Data from Open Data Toronto(REF) and we will using this to gain a better understading of whether there have been anay improvements in the apartment building

*Code and data are available at: www.github.com/ritvikpuri/sta304-paper-1 conditons. The data almost 10,000 entries spread over 40 columns. These columns give us a variety of information about each evaluation conducted, ranging from year built, year evaluated, property type, score, building address, wardnames, confirmed storeys, confimred units,stairwells, security rooms, laundry rooms, number of entrances, exits and a lot more. (REF)

To analyze our dataset, I will be using R (REF), tidyverse (REF), dyplr (REF). There were multiple steps I took to clean and extract the neccesary data. The first thing I did was change the neccesssary column types from character to numeric, because initally columns such as SCORE and WARD were character type and it would not have been possible to graph them. Since we have a lot more entries of PRIVATE property type than PUBLIC and TCHC property types, I filtered out all entries that were not PRIVATE. Thus we will only be assesing the building evaluations for private apartment buildings.

First I decide to take a look at the average scores of all the buildings based on the categories defined by the RentSafeTO Program (Ref). This meant categorizng scores according to the ranges they mentioned (Figure 1). To cateogirze the scores I added another column to the database that gave the cateogry based on their scores. I also reported the max, min, mean and standard deviation of all the scores in our dataset. (Table 1). Along with these two plots, I also made a bargraph based on the avereage scores of each of the 25 wards the apartment buildings are classified into. (Figure 2)

The next four graphs are to see what helps buildings get a higher score than the rest. Since we have a lot of columns, I decided to focus on the most basic ones. Namely, Stairwells, Security, Laundry Rooms and Elevators. I made scatterplots using ggplot(REF) to see if any of these factors promote scores.

```
package <- show_package("4ef82789-e038-44ef-a478-a8f3590c3eb1")
package
```

```
## # A tibble: 1 x 11
##   title       id          topics civic_issues publisher excerpt dataset_category
##   <chr>       <chr>       <chr>  <chr>        <chr>     <chr>   <chr>
## 1 Apartment ~ 4ef82789-e~ <NA>   <NA>         <NA>      <NA>    <NA>
## # ... with 4 more variables: num_resources <int>, formats <chr>,
## #   refresh_rate <chr>, last_refreshed <date>
```

```
# get all resources for this package
resources <- list_package_resources("4ef82789-e038-44ef-a478-a8f3590c3eb1")

# identify datastore resources; by default, Toronto Open Data sets datastore resource format to CSV for
datastore_resources <- filter(resources, tolower(format) %in% c('csv', 'geojson'))

# load the first datastore resource as a sample
data <- filter(datastore_resources, row_number()==1) %>% get_resource()

sapply(data, class)
```

```
##                       _id                          RSN
##                 "integer"                  "character"
##           YEAR_REGISTERED                YEAR_EVALUATED
##               "character"                  "character"
##                YEAR_BUILT                 PROPERTY_TYPE
##               "character"                  "character"
##                      WARD                      WARDNAME
##               "character"                  "character"
##              SITE_ADDRESS             CONFIRMED_STOREYS
##               "character"                  "character"
```

```
##              CONFIRMED_UNITS     EVALUATION_COMPLETED_ON
##                  "character"                 "character"
##                        SCORE            RESULTS_OF_SCORE
##                  "character"                 "character"
##         NO_OF_AREAS_EVALUATED              ENTRANCE_LOBBY
##                  "character"                 "character"
##       ENTRANCE_DOORS_WINDOWS                    SECURITY
##                  "character"                 "character"
##                   STAIRWELLS               LAUNDRY_ROOMS
##                  "character"                 "character"
##    INTERNAL_GUARDS_HANDRAILS         GARBAGE_CHUTE_ROOMS
##                  "character"                 "character"
##     GARBAGE_BIN_STORAGE_AREA                   ELEVATORS
##                  "character"                 "character"
##        STORAGE_AREAS_LOCKERS INTERIOR_WALL_CEILING_FLOOR
##                  "character"                 "character"
##     INTERIOR_LIGHTING_LEVELS                    GRAFFITI
##                  "character"                 "character"
##            EXTERIOR_CLADDING            EXTERIOR_GROUNDS
##                  "character"                 "character"
##            EXTERIOR_WALKWAYS              BALCONY_GUARDS
##                  "character"                 "character"
## WATER_PEN_EXT_BLDG_ELEMENTS                PARKING_AREA
##                  "character"                 "character"
##             OTHER_FACILITIES                        GRID
##                  "character"                 "character"
##                     LATITUDE                   LONGITUDE
##                  "character"                 "character"
##                            X                           Y
##                  "character"                 "character"
```

```
data[,1:5] <- sapply(data[,1:5], as.numeric)
data[,7:7] <- sapply(data[,7:7], as.numeric)
data[,10:11] <- sapply(data[,10:11], as.numeric)
data[,13:13] <- sapply(data[,13:13], as.numeric)
data[,14:35] <- sapply(data[,14:35], as.numeric)
```

```
## Warning in lapply(X = X, FUN = FUN, ...): NAs introduced by coercion

## Warning in lapply(X = X, FUN = FUN, ...): NAs introduced by coercion

## Warning in lapply(X = X, FUN = FUN, ...): NAs introduced by coercion

## Warning in lapply(X = X, FUN = FUN, ...): NAs introduced by coercion

## Warning in lapply(X = X, FUN = FUN, ...): NAs introduced by coercion

## Warning in lapply(X = X, FUN = FUN, ...): NAs introduced by coercion

## Warning in lapply(X = X, FUN = FUN, ...): NAs introduced by coercion

## Warning in lapply(X = X, FUN = FUN, ...): NAs introduced by coercion

## Warning in lapply(X = X, FUN = FUN, ...): NAs introduced by coercion
```

```
## Warning in lapply(X = X, FUN = FUN, ...): NAs introduced by coercion

## Warning in lapply(X = X, FUN = FUN, ...): NAs introduced by coercion

## Warning in lapply(X = X, FUN = FUN, ...): NAs introduced by coercion

## Warning in lapply(X = X, FUN = FUN, ...): NAs introduced by coercion

## Warning in lapply(X = X, FUN = FUN, ...): NAs introduced by coercion
```

```
data <- data %>%
  filter(PROPERTY_TYPE == "PRIVATE")

mean_data <- data %>%
  group_by(WARDNAME) %>%
  mutate(mean_scores = mean(SCORE))

a <- unique(mean_data$WARDNAME)
b <- unique(mean_data$mean_scores)
mean_wards <- data.frame(a, b)
view(data)
head(mean_wards)
```

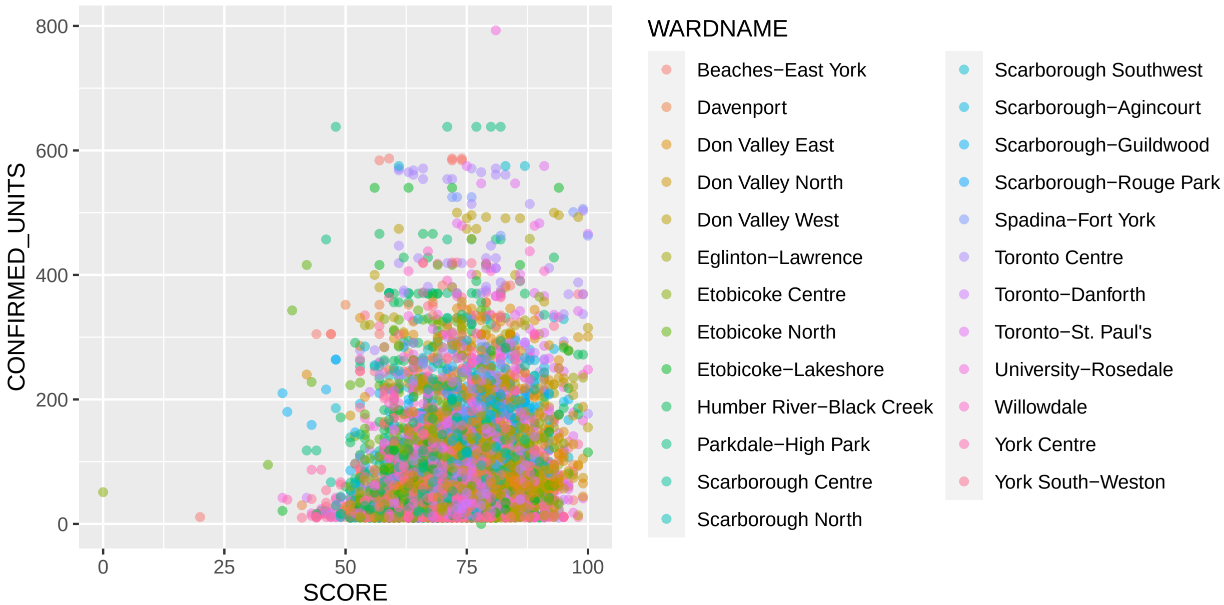
```
##                       a        b
## 1 Scarborough Southwest 72.11444
## 2      Toronto-Danforth 73.23695
## 3   Etobicoke-Lakeshore 71.27854
## 4      Etobicoke Centre 72.27666
## 5        Toronto Centre 73.31627
## 6           York Centre 71.18634
```

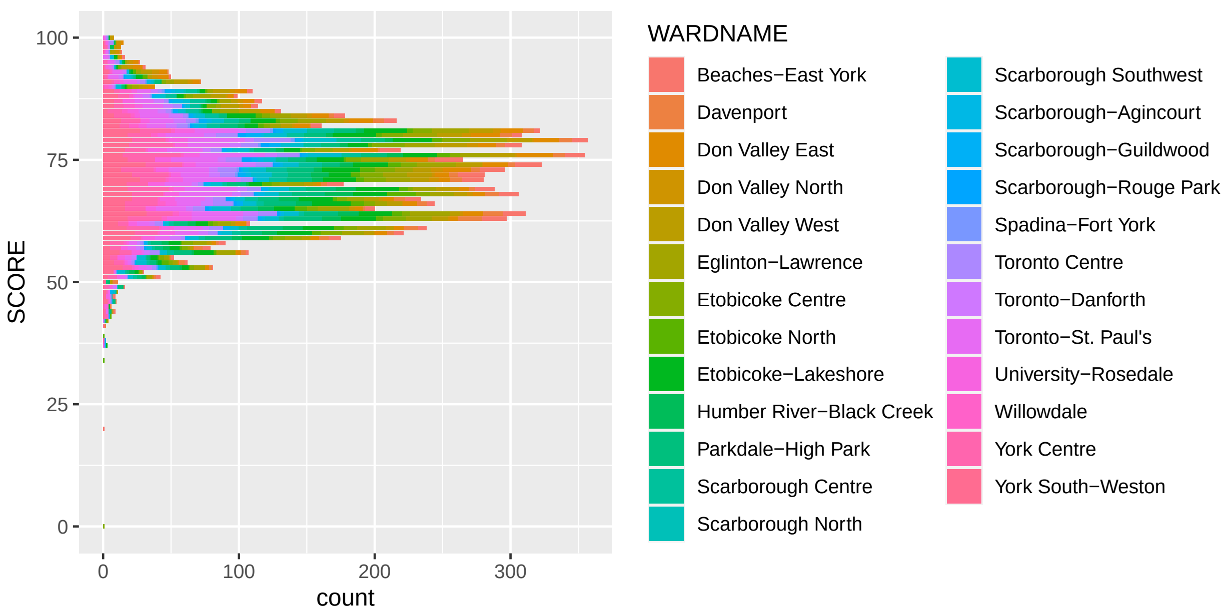
```
view(mean_wards)
```

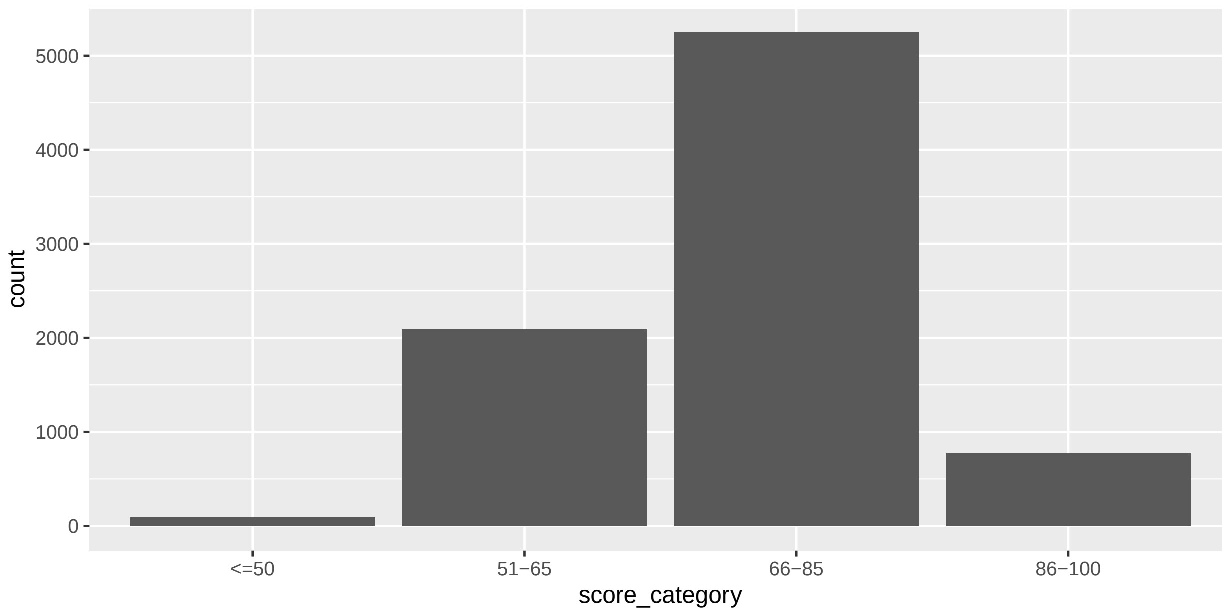
```
compare_scores <- data %>%
  as_tibble() %>%
  select(WARD, SCORE) %>%
  summarize(
    The_min = min(SCORE),
    The_max = max(SCORE),
    Mean = mean(SCORE),
    Std_dev = sd(SCORE)) %>%
  arrange(desc(Mean))

compare_scores %>%
  knitr::kable(caption = "Variation of Ward Scores",
               col.names = c("Score Min", "Score Max", "Score Mean", "Score Standard Deviation"),
               align = c('l', 'l', 'l', 'l'),
               booktabs = T)
```

Table 1: Variation of Ward Scores

Score Min	Score Max	Score Mean	Score Standard Deviation
0	100	72.41891	10.13594



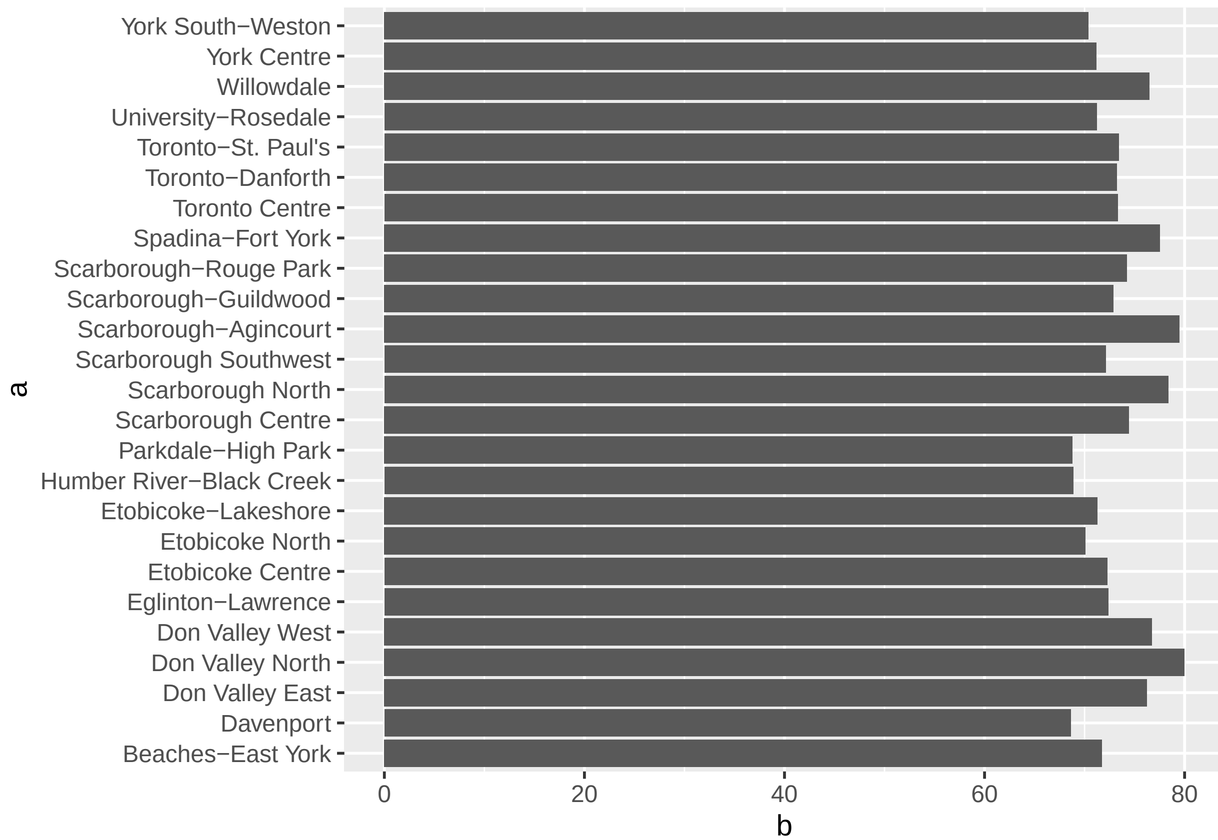




We can see how there are a lot of apartment building with a score thats between 66 and 85, which means they wont be evaluated for the next 2 years and are consiered to be in good condition overall. But what makes the score of these buildings higher than the rest? And what do the few buildings with a score of more than 86 do that separtes them from the rest?

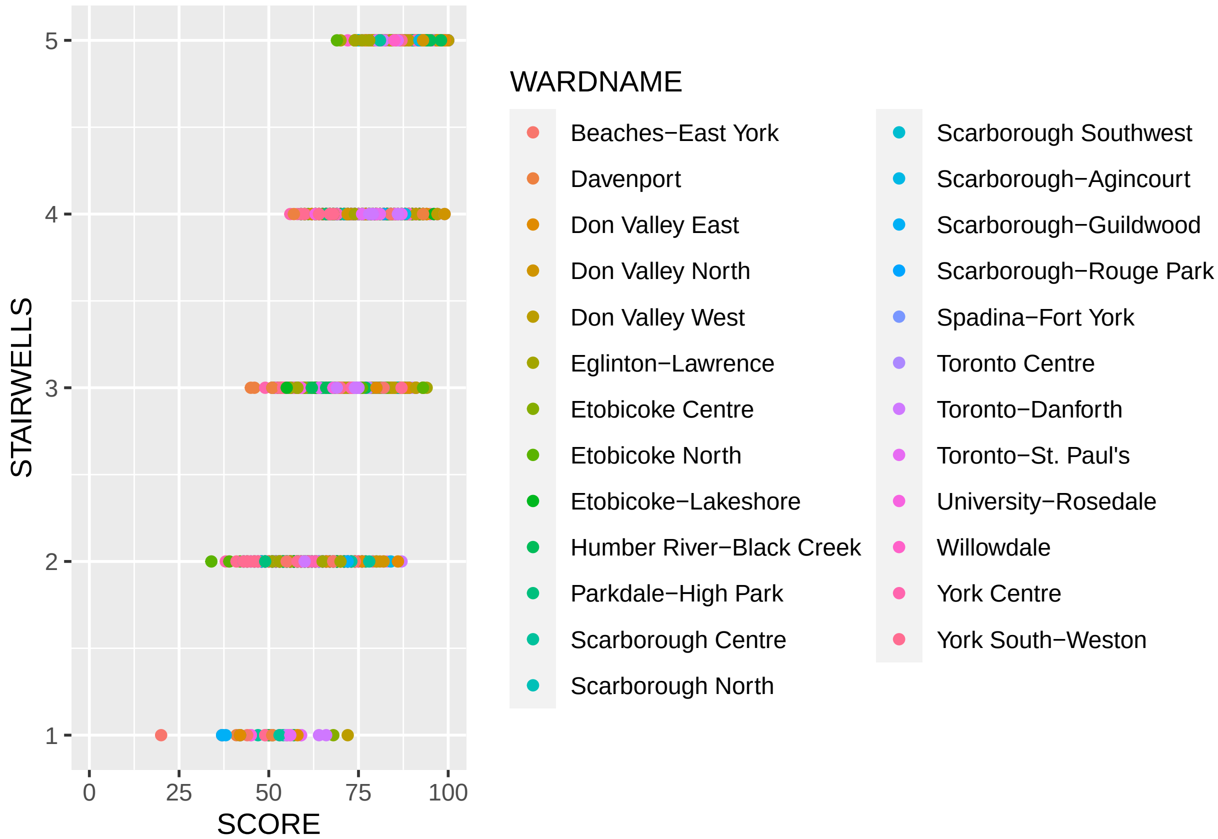
3 Model

```
#Figure 2
mean_wards %>%
  ggplot(mapping = aes(x = b, y = a)) +
  geom_bar(stat = "identity", position = "dodge")
```



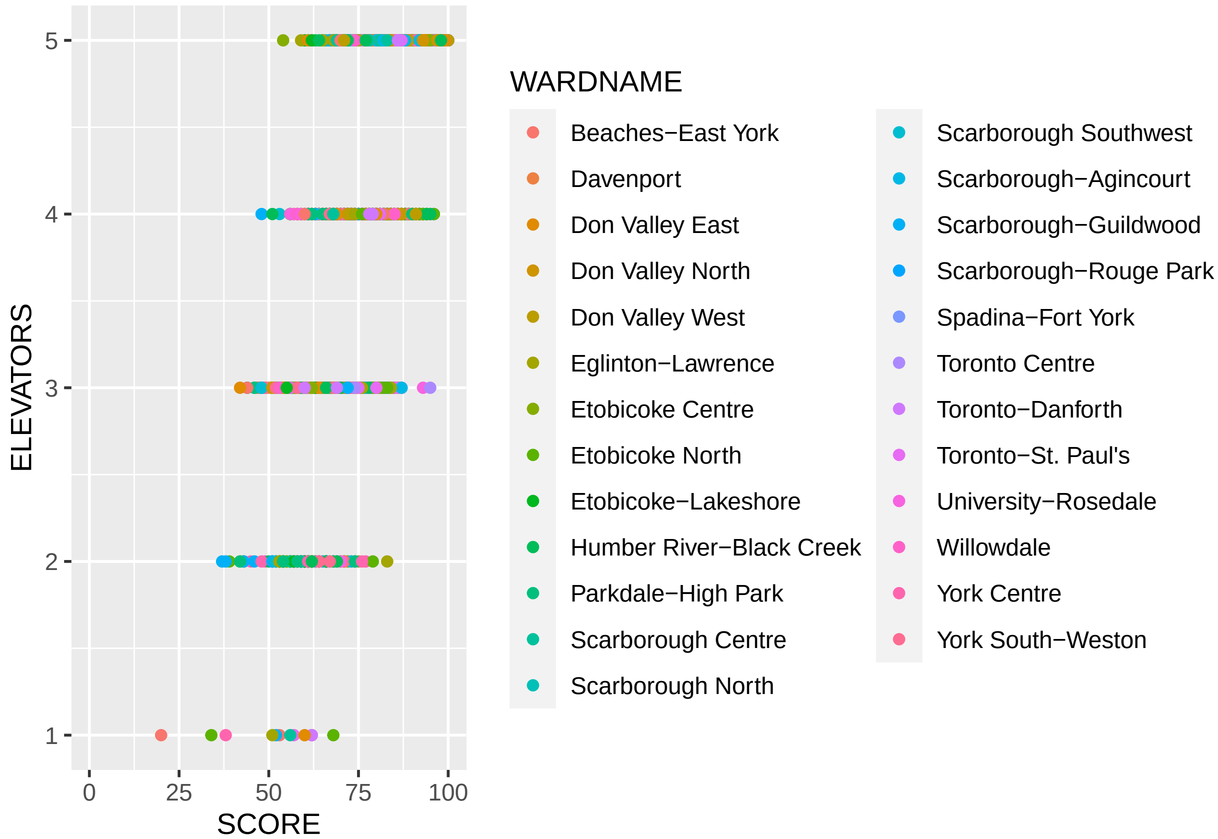

```r
#change fill
#Figure 3
data %>%
  ggplot(mapping = aes(x = SCORE, y = STAIRWELLS, color = WARDNAME)) +
  geom_point()
```

```
## Warning: Removed 1 rows containing missing values (geom_point).
```



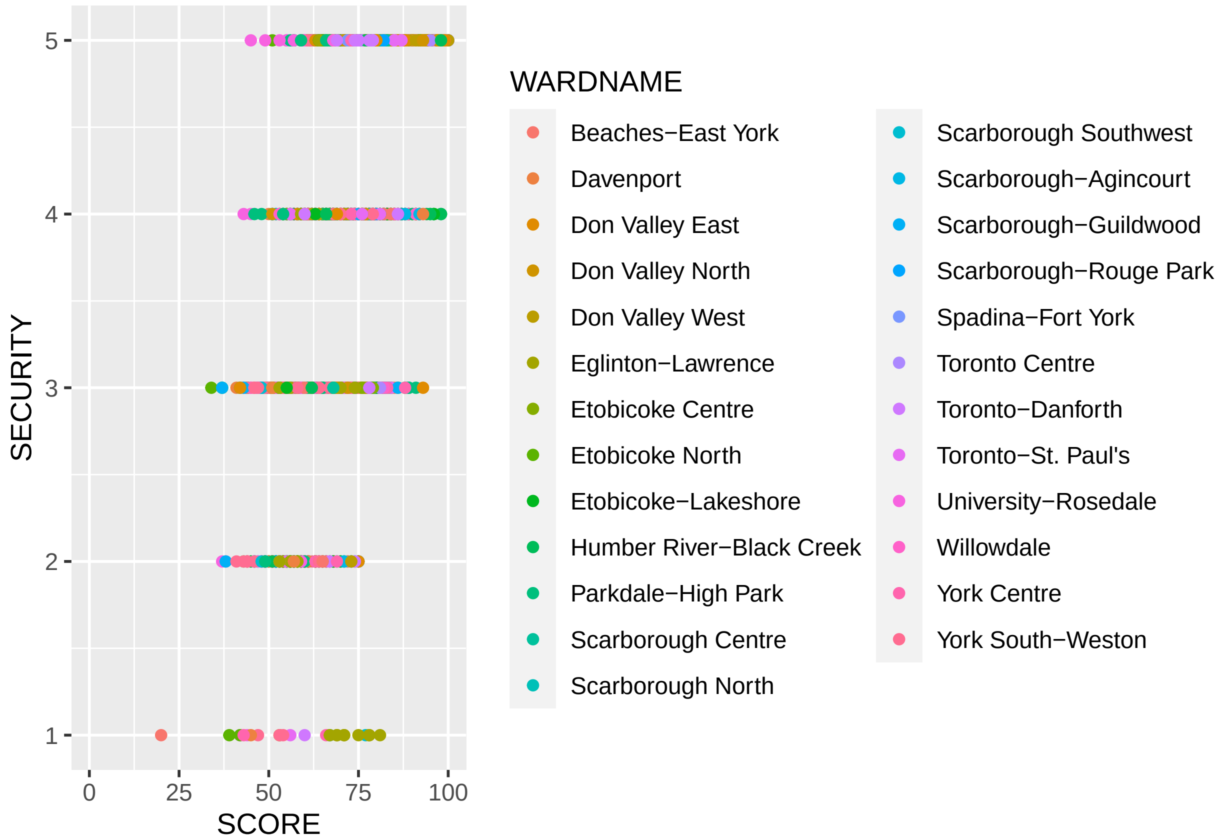

```
data %>%
  ggplot(mapping = aes(x = SCORE, y = ELEVATORS, color = WARDNAME)) +
  geom_point()
```

```
## Warning: Removed 3657 rows containing missing values (geom_point).
```



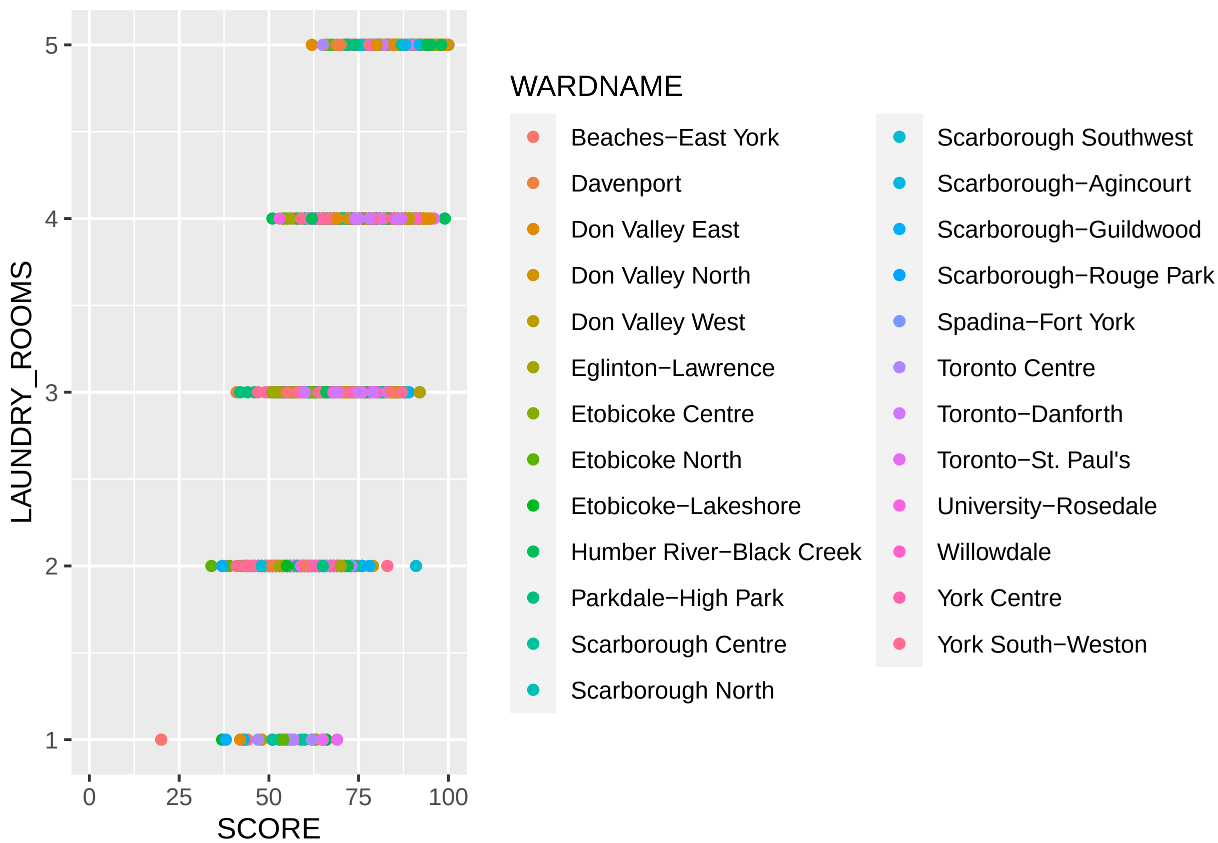

```
data %>%
  ggplot(mapping = aes(x = SCORE, y = SECURITY, color = WARDNAME)) +
  geom_point()
```

```
## Warning: Removed 2 rows containing missing values (geom_point).
```




```
data %>%
  ggplot(mapping = aes(x = SCORE, y = LAUNDRY_ROOMS, color = WARDNAME)) +
  geom_point()
```

```
## Warning: Removed 491 rows containing missing values (geom_point).
```



4 Results

5 Discussion

5.1 Initial Graphs

The RentSafeTO bylaws say that a buildings next evaluation is determined based on their score. If the score is less than 50, then a complete buliding inspection and audit is conducted. This means that the building condition is not fit for living. The main aim of the program should be to see make sure each buiding provides a liveable space that is both safe for current residents and future ones. Figure 1 tells us that most of the residences fall in the 66-85 range, which means that their nexct evaluation comes 2 years later, This should be considered a good things as there a re very few apartments with scores lower than 50.

The table tells us that while some building did get a perfect score, some building also got zero points. While this is good to know, it is not helpful for our analysis. We do get to know that the average score is 72.4 which holds in line with our conclusion from the bargraph of categories.

Figure 2 shows us that while all wards are relatively close in scores, there are a few that on average better buildings than others. Don Valley North, Scarborough North, Scarborough Agincourt are the top three wards in terms of apartment quality. Davenport, Humber River-Black Creek and Parkdale- High Park are the wards with the three lowest wards. This gives us some insight as to how the residences are being maintained around Toronto, and can even help someone narrow down their search if they are looking for a place.

5.2 Scatterplots

Our dataset consisted on multiple columns. Most of them told us the number of facilities in each apartment buliding. One would assume these features help a bulidings score, and these scatterplots give us some insight as to if this is true or not. Since there are a lot of facilites to account for, I decided to focus on the ones I find most important according to me, These plots do show an indication that having more is better. An increase in the number of stairwells, security, laundry rooms and elevators tend to give a higher evaluation scor, as indicated by Figure 3.

While their are other factors included in the dataset, such as number of lobbies and garbage areas, these factors do not tend to play a crtical role in evaluations and these are also not things most people consider while looking at prospective residences.

5.3 Growth

An important aspect of introducing and keeping large amounts on data on a program such as RentSafeTO in a populous and fast growing city like Toronto is to observe growth and see if it has been helpful or not. I calculated the averege scores over all the buildings and grouped them by the year of evaluation. The first year the prgram was conducted (2017), the average score was only 65. But since these were strict laws and building owners followed them, we saw a large incerease in the average score over the next 2 years, going up all the way to 79 in 2019. A few points have dropped in the last 2 years, but I think this can be attributed to the COVID-19 pandemic.

```
#Table 2
year_mean_scores <- mut_data %>%
  group_by(YEAR_EVALUATED) %>%
  summarise(value = mean(SCORE))
year_mean_scores
```

```
## # A tibble: 6 x 2
##   YEAR_EVALUATED value
##            <dbl> <dbl>
## 1           2017  65.1
## 2           2018  72.9
## 3           2019  79.3
## 4           2020  77.5
## 5           2021  76.4
## 6             NA  68
```

5.4 Weaknesses and next steps

A weakness that is evident can be seen in Table 2 above, which is the missing values. I could have removed the row from the table but it should be noted that this dataset contained a lot of missing cells in multiple rows and columns. A lot of columns are also not that helpful when you look at the way they are presented. The OTHER_FACILITIES column gives us the number of extra facilities certain luxury apartments might have. howver, since there is no description is provided, we have no idea of knowning what these features are.

I belive a good step to make as this program continues is to maybe modify the cateogry values that are currently set. As we observed, the average score was 72 and the baseline is at 50. If this baseline is increased, the average score will go up and the quality of buildings all arund the city will increase. Adding some comments for each inspection will also help keep a track of any persisting problems that the buliding has not fixed over many evaluations.

Appendix

A Additional details

B References

https://www.toronto.ca/community-people/housing-shelter/rental-housing-tenant-information/rental-housing-standards/apartment-building-standards/rentsafeto-for-tenants/

https://www.toronto.ca/community-people/housing-shelter/rental-housing-tenant-information/rental-housing-standards/apartment-building-standards/rentsafeto-for-building-owners/rentsafeto-building-evaluations-and-audits/