

# Apartment building in Toronto in 2020 - The correlation between the high quality buildings and the number of stories, and the type of security and stairwells

Xinjing Guo - 1005086620

## Research question:

In this project, my interest area is the apartment building evaluation in Toronto in 2020. According to Lepkova (2016), he points out that construction industry takes an important role in the development of the city. And Dennis (2016) also shows that in Toronto, the high-rise buildings replace the early department. It is because Toronto, where has a large number of people, is the largest city in Canada, more apartments are built. It means the apartment building is significant for people to live in. Thus, nowadays, compared to prices, the quality of the apartment is more essential for people. This is because the most important purpose of the building is to resist the external wind chill and protect yourself. Thus, researching apartment building evaluation is important and residents can have a better overview of the apartment buildings.

## Background:

The security and stairwells are the most important factors for the high apartment buildings as a lot of papers think they are the conditions of the building. According to Ho (2008), he claims that safe building means minimizing people's injury and death. The stairwells are significant under the emergency. When the building has a fire or emergency, the stairwell is the only way to evacuate people. Thus, the condition of the stairwells affects apartments quality. Waller (1978) claims that "the burglary figure for Toronto was 840 per 100,000 population" and Rollwagen (2016) points out that the apartment buildings have more probability have the crime since there has a more common room and many households on each floor. Thus, the security of the department is significant to protect residents' safety and property. Moreover, the higher story of an apartment building should have higher requirements on safety so that people live in high-rise apartments will feel safe. In this project, the research question is whether the total high-quality buildings correlate with the number of stories and the type of security and stairwells in Toronto in 2020. Since based on the research, the security and stairwells of the building are significant of the department buildings, and the number of storied takes an important role in the apartment buildings. Then, based on the type of security and stairwells, and the number of stories, we can know how they relate to the total score.

## Data source:

I select the dataset from the open data Toronto: <https://open.toronto.ca/dataset/apartment-building-evaluation/>

The dataset is about the score of apartment buildings with more than 3 stories or has more than 10 units. These standards scores are based on the RentSafeTO to ensure residents' safety. If the total score is less or equal to 50, the buildings would be inspected in all common areas. This dataset includes 40 variables and 20 variables are the scores of the apartment conditions such as security, lobby rooms, elevators. The score is from 1 to 5 which 1 means the worst and 5 means the best. Also, one of the variables is the overall score of the building. And one of the numerical variables is the number of stories in a building.

In order to research whether the high score buildings have the correlation with the type of security and stairwell (high or low), and the number of the stories of buildings, using the table and plots can observe easily.

## Exploratory data analysis:

Based on the dataset, the more recent year is, the more valuable the result has. So, I want to study the most recent year of data, I select the data only in 2020. The units of the buildings should be greater than 0 and less or equal to 600. It is because as the apartment buildings are not higher than 40 levels in the dataset, the units of a building greater than 600 are impossible. Then, I delete the outliers of units in the dataset. Moreover, the stories must be greater than 0, or the data would be meaningless which the data has 0 stories that should not be included. Furthermore, in order to observe the different scores of security and stairwell better, I divide 1 to 3 into a group and name it “low”, and others named “high”.

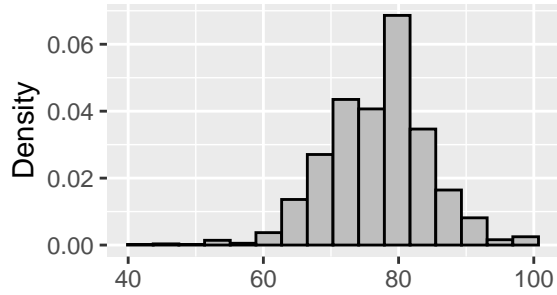


Fig.1 score of building evaluation

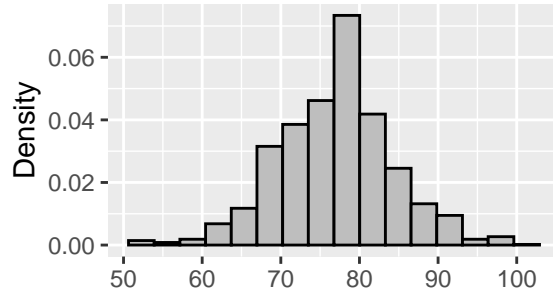


Fig.2 score of building evaluation above 50

### Discussion of what is observed in the EDA:

Figure 1 shows the scores of apartment buildings in Toronto in 2020. This plot is a slight left-skewed histogram since from range 40 to 100, most buildings got scores from 70 to 88 which is quite higher. There were some variables that were skewed which means there are some buildings that have low scores. Also, as I am only interested in high-quality apartment buildings, and as the total score is below 50, buildings would be inspected in all common areas. Thus, I only choose the data in which the total score is above 50.

Then, I got figure 2. This histogram is the buildings that have scored above 50 which are high-quality apartment buildings. This histogram is approximate to the normal distribution which most buildings get a score around 76 from the range 50 to 100. the mode is about 76 and the density is above 0.7 which is quite large.

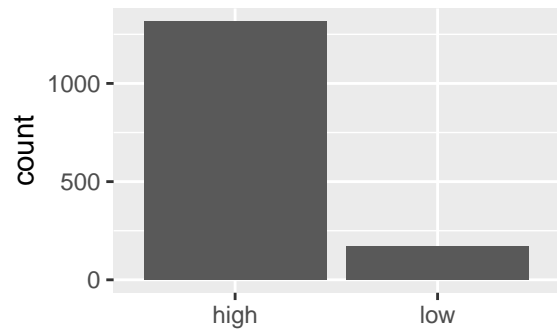


Fig.3 Type of security

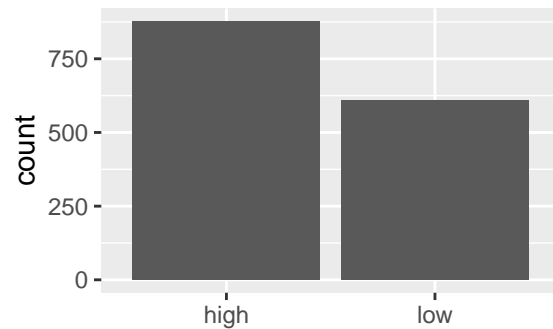


Fig.4 Type of stairwells

Figure two is the bar plot which is about the type of security which is from high score building in Toronto in 2020. From this plot, over 1250 apartment buildings have high security and only about 200 buildings from Toronto Community Housing have low security. Thus, it means that most buildings have a very high-security system in apartment buildings in Toronto in 2020.

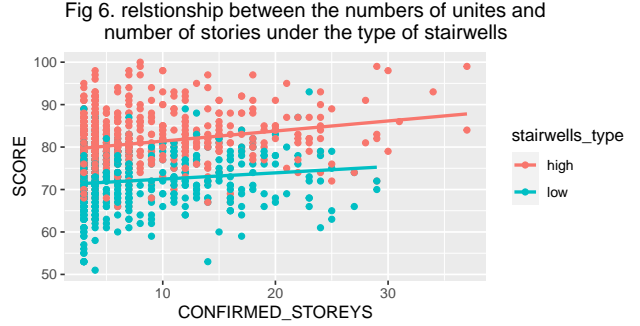
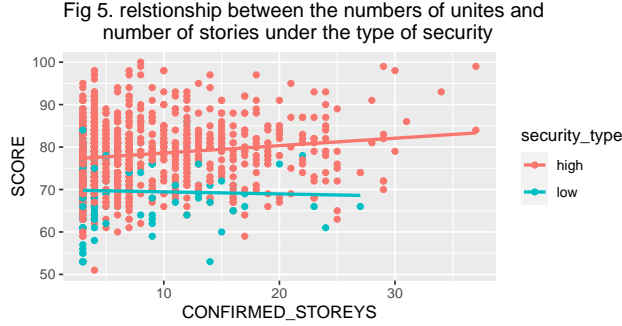
Figure three is the bar plot which is about the type of stairwells. If the score of stairwells is from 1 to 3, it would return low, else return high. Then, it shows that for the condition of stairwells, about 875 buildings are high and about 625 buildings are low. It shows that there is a very large of buildings that had the problem with the condition of stairwells.

Since the gap between high and low security and stairwells is large, it is meaningful to research them.

Table 1: Summaries for the number of storeies in a building in 2020 in Toronto

min	median	max	mean	sd
3	4	37	6.851178	5.645533

Table 1 is the summary of the number of stories of high-quality apartment buildings in Toronto in 2020. The minimum levels is 3 and the maximum number of stories is 37 for a building. Then, the range of the histogram is 36. Moreover, since the mean is 6 and the median is 4, it means that most apartment buildings are around 4 to 6 which is quite low. The standard deviation of is 5 which is a little large. Then, it means that the range of the histogram of the number of stories is large.



Since the study, whether the high-quality buildings has the correlation with the type of security and stairwells, using scatter plots, can observe their relationships.

Figure 5 is the scatter plot about the relationship between the number of units and the number of stories under the type of security in Toronto in 2020. From the plot, it shows that for the high condition of securities, as the stories increase, the score also increases. For the low condition of security, as the stories increase, the score decrease. Then, the direction under the low security is negative.

Figure 6 is about the relationship between the number of stories and the number of scores under the type of stairwells. Under the different types of stairwells, the slopes are both positive. It exhibits if the building has a high condition on stairwells, the score would increase as the stories increase. For the low condition of the stairwells, as the stories of buildings increase, the score is also increased.

## Linear model:

Based on this study, using a multiple linear regression model is appropriate. The response variable is the score of the apartment buildings have and the predictors are the type of security and stairwell, and the number of stories of the building. Since the response variable Y is the total score. It is dependent and based on figure 2, it is a normal distribution, it conforms to the assumption of the linear regression model. Moreover, based on figure 5 and figure 6 shows the relationship between the total score of buildings and the number of storied based on the type of security and stairwells. It exhibits that they have a correlation with each other. Thus, it fits the assumption of the linear regression model. Also, predictor X, they are an independent numerical variable and categorical which is follow the assumption. Thus, using the linear regression model is an appropriate statistical tool to address my research question. Some things I plan to keep an eye on when working on my data analysis is researching what is the specific correlation between the total score and the number of stories, and the type of security and stairwells. So, I will use the hypothesis test and confidence interval to study in the future study.

## References

- Baptiste Auguie (2017). `gridExtra`: Miscellaneous Functions for “Grid” Graphics. R package version 2.3. <https://CRAN.R-project.org/package=gridExtra>
- Dennis, R. (2016). Apartment housing then and now: vignettes from Toronto and Montreal (Le logement en appartement alors et aujourd’hui: vignettes de Toronto et de Montréal). *British Journal of Canadian Studies*, 29(2), 249–272. <https://doi.org/10.3828/bjcs.2016.12>
- JJ Allaire and Yihui Xie and Jonathan McPherson and Javier Luraschi and Kevin Ushey and Aron Atkins and Hadley Wickham and Joe Cheng and Winston Chang and Richard Iannone (2020). `rmarkdown`: Dynamic Documents for R. R package version 2.6. URL <https://rmarkdown.rstudio.com>.
- Hao Zhu (2021). `kableExtra`: Construct Complex Table with ‘kable’ and Pipe Syntax. R package version 1.3.4. <https://CRAN.R-project.org/package=kableExtra>
- Ho, D. C.-W., Chau, K.-W., King-Chung Cheung, A., Yau, Y., Wong, S.-K., Leung, H.-F., Siu-Yu Lau, S., & Wong, W.-S. (2008). A survey of the health and safety conditions of apartment buildings in Hong Kong. *Building and Environment*, 43(5), 764–775. <https://doi.org/10.1016/j.buildenv.2007.01.035>
- Lepkova, N., Butkiene, E., & Belej, M. (2016). Study of Customer Satisfaction with Living Conditions in New Apartment Buildings. *Real Estate Management and Valuation*, 24(3), 52–70. <https://doi.org/10.1515/remav-2016-0021>
- Rollwagen, H. (2016). The Relationship Between Dwelling Type and Fear of Crime. *Environment and Behavior*, 48(2), 365–387. <https://doi.org/10.1177/0013916514540459>
- Waller, I., & Okihiro, N. (2019). Burglary Rates in Toronto and Other Areas. In *Burglary* (pp. 15–22). University of Toronto Press. <https://doi.org/10.3138/9781487574208-004>
- Wickham et al., (2019). Welcome to the tidyverse. *Journal of Open Source Software*, 4(43), 1686, <https://doi.org/10.21105/joss.01686>
- Yihui Xie and J.J. Allaire and Garrett Grolmund (2018). *R Markdown: The Definitive Guide*. Chapman and Hall/CRC. ISBN 9781138359338. URL <https://bookdown.org/yihui/rmarkdown>.
- Yihui Xie and Christophe Dervieux and Emily Riederer (2020). *R Markdown Cookbook*. Chapman and Hall/CRC. ISBN 9780367563837. URL <https://bookdown.org/yihui/rmarkdown-cookbook>.