Exploring Toronto Airbnb Price Throughout COVID-19*

Henry Shiffer

31 August 2021

Abstract

This paper explores how the relationship of Airbnb listing price and listing quality has changed over the course of the COVID-19 pandemic. The goal of this paper is to examine if there are certain attributes of an Airbnb listing's quality that have led to higher prices per-night or lower prices per-night in the COVID-19 impacted Toronto Airbnb market. This paper also explores how per-night Airbnb listing prices in certain Toronto neighbourhoods have varied throughout the months of February 2020 up to and including April 2021. Model results find that an Airbnb listing's bedroom amount and bathroom amount have the strongest positive effect in determining an Airbnb's nightly price throughout the months of February 2020 up to and including April 2021. Model results also find that the time duration over the course of the COVID-19 pandemic (February 2020 - April 2021) and Toronto Lockdowns and stay-at-home orders, did not provide a significant effect in determining an Airbnb listing's price per-night. This paper concludes that higher prices were associated with attributes that represent a listing of higher quality. However, it was seen that if a host is a Superhost and the amount of reviews a listing has per month led to slightly lower prices per-night. This can likely be attributed to a host understanding the market that surrounds their listing, and even if the listing is of high quality, the host makes the listing affordable to attract business.

Contents

*Code and data are available at: https://github.com/henryshiffer/Airbnb-Data-Exploration-and-Modelling

1 Introduction

The COVID-19 pandemic and resulting stay-at-home orders and social distancing measures increased the importance of housing and personal space, while simultaneously changing the social fabric of society. While all employment and economic sectors were forced into some change, the small business economy and the tourism industry saw very dramatic changes happening the most quickly. Since the start of the COVID-19 outbreak, 70% of Canadian small businesses had seen a drop in revenues by 30% or more, and it is anticipated there will be a slower recovery for small businesses once the pandemic is over (Mo et al., 2020). For tourism, the restrictions on travel in Canada have led international arrivals in Canada to fall 54.2% in March 2020 when compared to February 2020 (Liu, 2020), and hotel occupancy rates were below 20% by the beginning of April 2020 (Liu, 2020).

With the fluctuating COVID-19 restrictions on travel, and with the risks that traveling during COVID-19 impose, it is no question that Airbnb whose market relies heavily on tourism has taken a significant hit during the COVID-19 pandemic. Local lockdowns resulted in a 57.8% fall in global Airbnb booking activities (Hu & Lee, 2020). According to CNBC in May 2020, Airbnb had laid off a quarter of its workforce (Bosa & Rodriguez, 2020) and with no end to the pandemic in sight, there is no way to predict if the organization will ever recover its pre-pandemic success. This loss, while significant to the Airbnb organization, can be felt much further down by Airbnb hosts. Many hosts, whose primary income comes from rental of the units they provide, are stuck with bills and mortgages to pay, sometimes for many units.

In March 2020, Ontario Premier Doug Ford declared a provincial state of emergency which prompted the closing of many non-essential businesses and facilities across Ontario. In late July 2020, Toronto and Peel region were allowed to enter stage 3 of the Ontario reopening plan. In October 2020, Peel region, Toronto and Ottawa were rolled back to modified stage 2 for 28 days. In November 2020, Doug Ford announced that the province would move towards a 5-tier system to determine restrictions. Mid-November Toronto was announced to move into the Red Tier "Control" to provide more restrictions, and finally on November 20th, Ford announced Toronto and Peel region would be moved into lockdown effective November 23rd. This prompted all non-essential businesses to be closed along with all dine-in restaurants, gyms, and personal care services. In January 2021, Doug Ford declared a second declaration of emergency, and stay-at-home orders were declared for the province. In March 2021, Toronto and Peel region were able to exit stay-at-home orders and Toronto and Peel region entered Grey Tier "Lockdown". In April 2021, Ford announced the province would move into White Tier "Shutdown", which would impose the same measures put in place during the original province wide shutdown in March 2020.

Ontario and the city of Toronto have placed an increased control on the Airbnb and short-term rental market during the COVID-19 outbreak. In April 2020, Ontario had banned all short-term rentals (a rental period of less than 28 days) unless the unit is being rented to someone in need of housing during the "Emergency Order" (Eisenberg & Shrives, 2020). In September 2020, Toronto laid out a series of short-term rental reforms. These reforms require short-term rental operators to need to register with the city to allow short-term tenants (Stevens, 2020). The reforms also state that these landlords may only be allowed to rent out their primary residence for up to 180 nights per year.

Using Toronto Airbnb data from February 2020 up to and including April 2021, this paper will explore how price has varied with relation to an Airbnb listing's quality over the course of the pandemic and Toronto's pandemic response. This paper will primarily focus on Airbnb listings around the University of Toronto area. These neighbourhoods being: Kensington-Chinatown, Annex, University, Church-Yonge Corridor, and Bay Street Corridor. This paper defines a criterion to measure an Airbnb listing's quality, and models to represent these criteria.

2 Data

This paper uses the programming language `R` (Core Team, 2021) with `tidyverse` (Kuhn et al., 2020). `ggmap` (Kahle & Wickham, 2013) was used for the spatial diagrams. `lme4` (Bates et al., 2015) was used for fitting the linear mixed effects models. `ggpubr` (Kassambara, 2020) was used for plotting.

2.1 Source of data

The website Inside Airbnb provides scraped listing data from the Airbnb website for major cities around the world. For each month since December 2015, the Airbnb website listings are scraped for each major city and are posted in a .csv file that pertains to that particular month and year. The Inside Airbnb Project was put together by Murray Cox and John Morris. For this analysis, listing data .csv files from Toronto were downloaded from this site.

2.2 Months of interest

This paper is interested in the months surrounding the initial COVID-19 lockdowns up to a more matured COVID-19 situation in Toronto. The data was selected from the months February 2020 up to and including April 2021. These particular months were selected to allow a view of how COVID-19's progression and remission trends so far have impacted Toronto's Airbnb listings.

2.3 Initial data analysis

A data set with all months of interest is very large. To understand the data better, it will make sense to first pick a single month to examine more closely before there is analysis on the full data. The month February 2021 was selected for initial analysis.

2.3.1 Airbnb price

February 2021 Toronto Airbnb price per-night can be plotted on the X-axis with the number of properties that have that price per-night on the Y-axis.

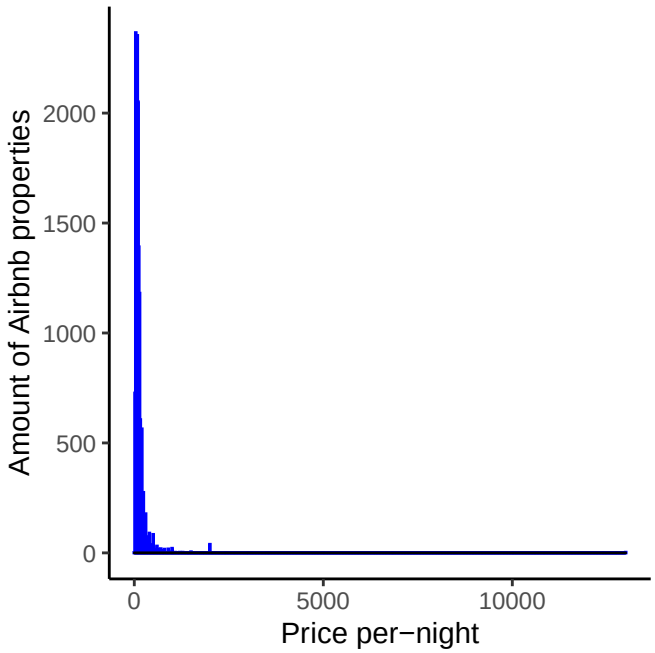

Figure 1: February 2021 Toronto price per-night distribution

It can be observed from Figure 1 that the distribution of prices is strongly right skewed. We can examine if this skew still exists zooming in closer at the prices per-night.

We can now plot the distribution of prices per-night greater than $500.

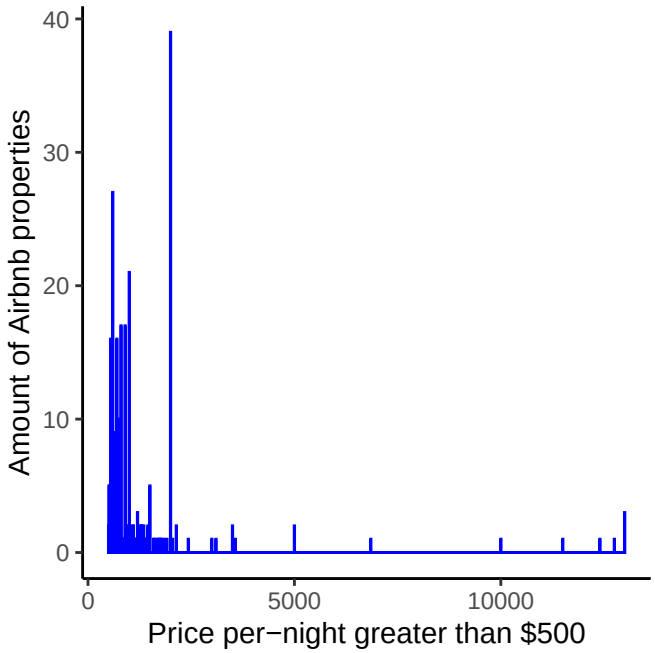

Figure 2: February 2021 Toronto price per-night $>$ 500 distribution

The right skew still exists when examining Figure 2. As we have a much better view of the outliers, we can see that there exists properties that are greater than $10,000 a night.

We can can get a better view of these higher end prices by examining the amount of properties greater than $10,000 a night.

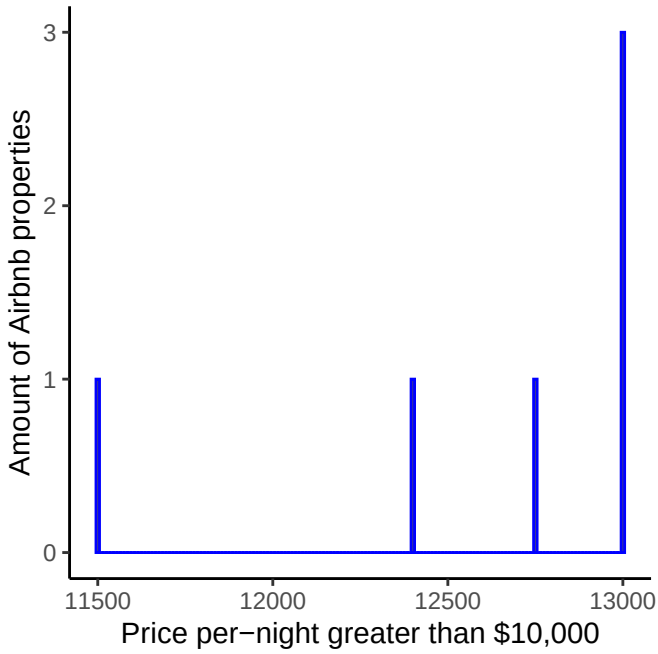

Figure 3: February 2021 Toronto price per-night $>$ 10,000 distribution

It appears from Figure 3 that the highest price in February 2021 is $13,000 a night, and there are three properties that have this price.

To see how prices in February 2021 compared to the pre-pandemic months of February 2019 and February 2020, we can plot these prices on a map of Toronto. A log scale of the prices will be used as it will be easier to see the difference between low, medium, and high prices.

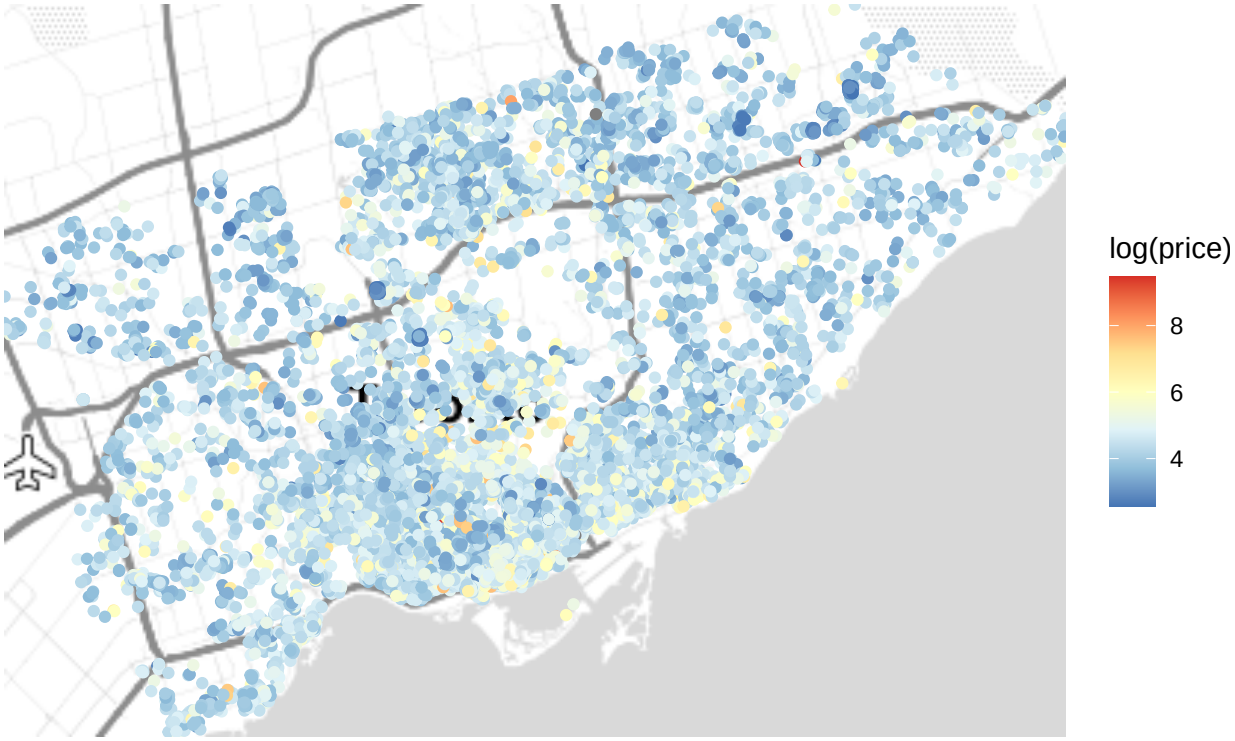

Figure 4: February 2021 listing price per-night

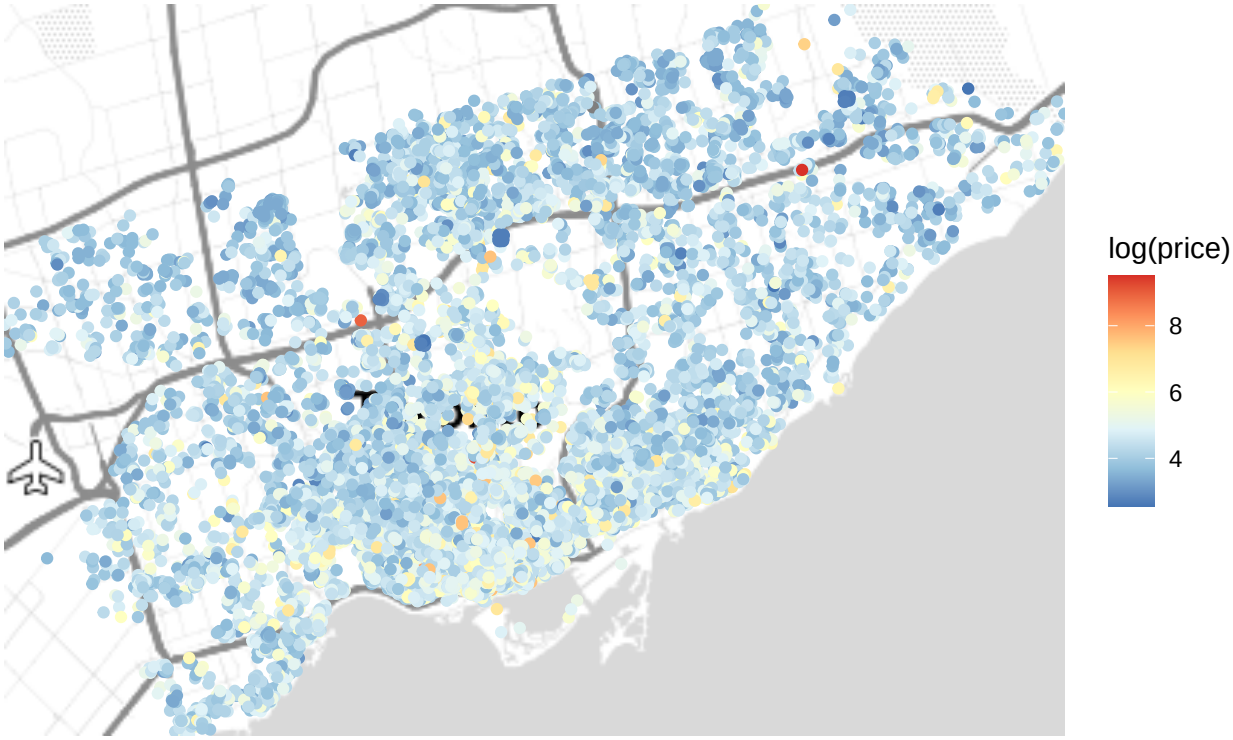

Figure 5: February 2020 listing price per-night

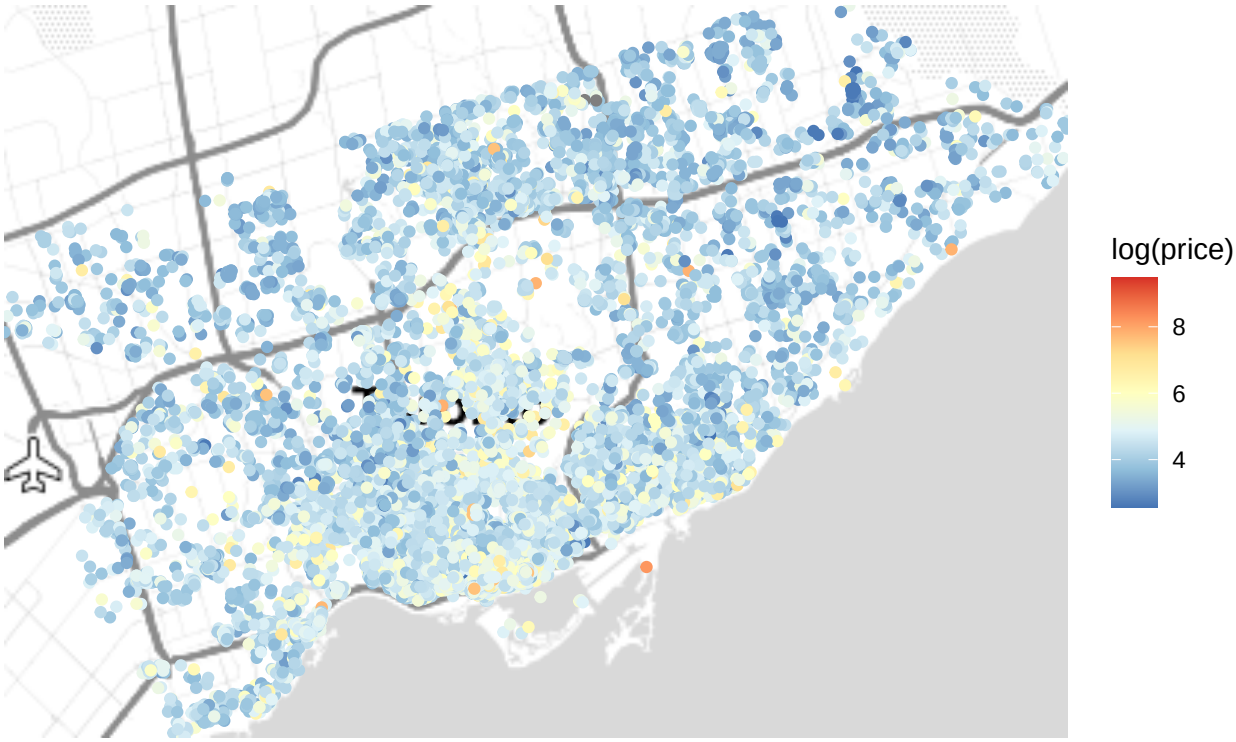

Figure 6: February 2019 listing price per-night

2.3.2 Airbnb quality

2.3.2.1 Defining quality This paper would like to examine the relationship between Airbnb price and Airbnb quality. However, Airbnb quality is subjective, and a formal definition of listing quality must be defined for quality to be measured. Listing quality is defined by this project as bedroom and bathroom amount, the rating from review scores, the neighbourhood the property is in, and if a host is a Superhost or not. This paper hypothesizes that listings that have a higher quality, will have a higher per-night price as a result.

Based on this definition:

- An Airbnb listing with a low amount of bedrooms would have lower quality than a listing with a high amount of bedrooms.
- An Airbnb listing with a low amount of bathrooms would have a lower quality than a listing with a high amount of bathrooms.
- An Airbnb lising with low review scores would have a lower quality than a listing with high review scores.

According to Airbnb's website, a Superhost is a "experienced host who provides a shining example for other hosts and extraordinary experiences for their guests" (Airbnb, 2021). Airbnb also states that a Superhost's activity is checked four times a year, to ensure that "the program highlights the people who are most dedicated to providing outstanding hospitality" (Airbnb, 2021). According to Airbnb's description, it is reasonable to believe that a listing that is ran by a Superhost would be of a higher quality than a listing ran by a host who is not a Superhost.

From here we can select variables that measure these attributes and examine their relationship with price, and the trends that exist within them.

Superhost spatial view

From the Airbnb website, a Superhost is defined as a host who has met all of the following requirements:

- "Completed at least 10 trips OR completed 3 reservations that total at least 100 nights"
- "Maintained a 90% response rate or higher"
- "Maintained a 1% percent cancellation rate (1 cancellation per 100 reservations)"
- "Maintained a 4.8 overall rating (this rating looks at the past 365 days of reviews, based on the date the guest left a review, not the date the guest checked out)"

If a listing is ran by a host that is a Superhost, the listing has a better chance of being better maintained and of higher quality to stay in.

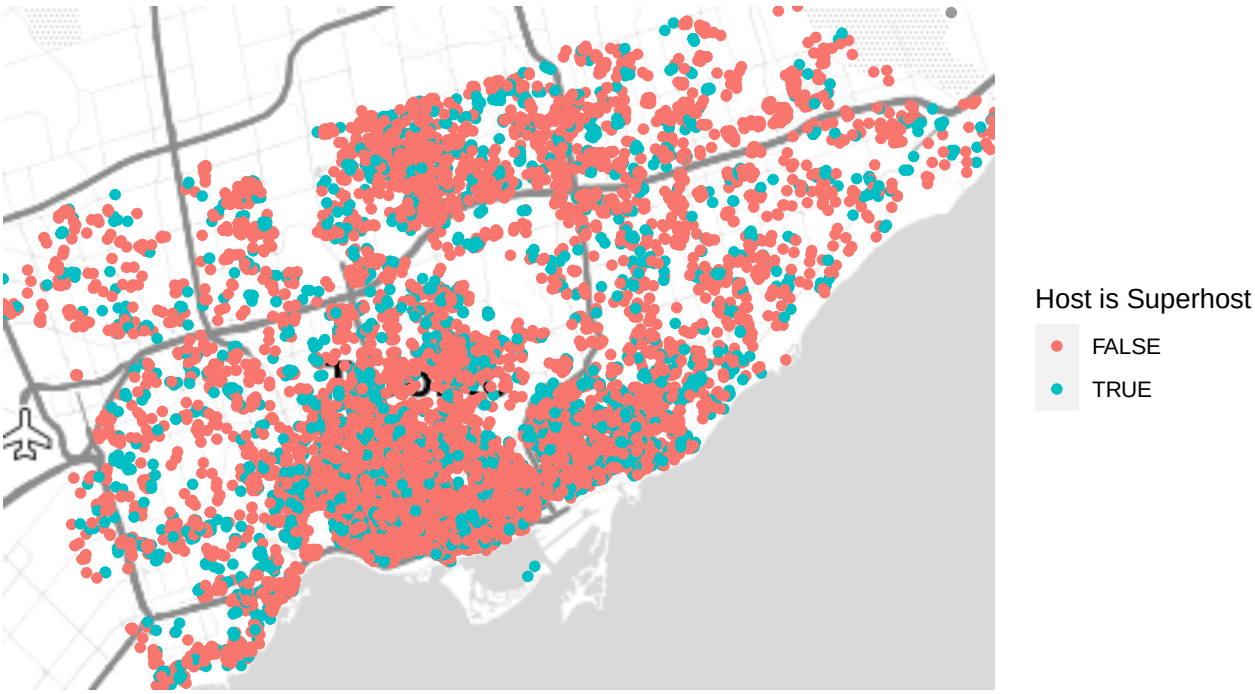

Figure 7: February 2021 Superhost listings

It can be seen from Figure 7 that there does exist a fairly high amount of Superhost listings in February 2021. However, it appears that a majority of listings are listings that do not have Superhost status.

Bedroom amount spatial view

Bedroom amounts can indicate how much space a property has, as well as how many people can comfortably live there.

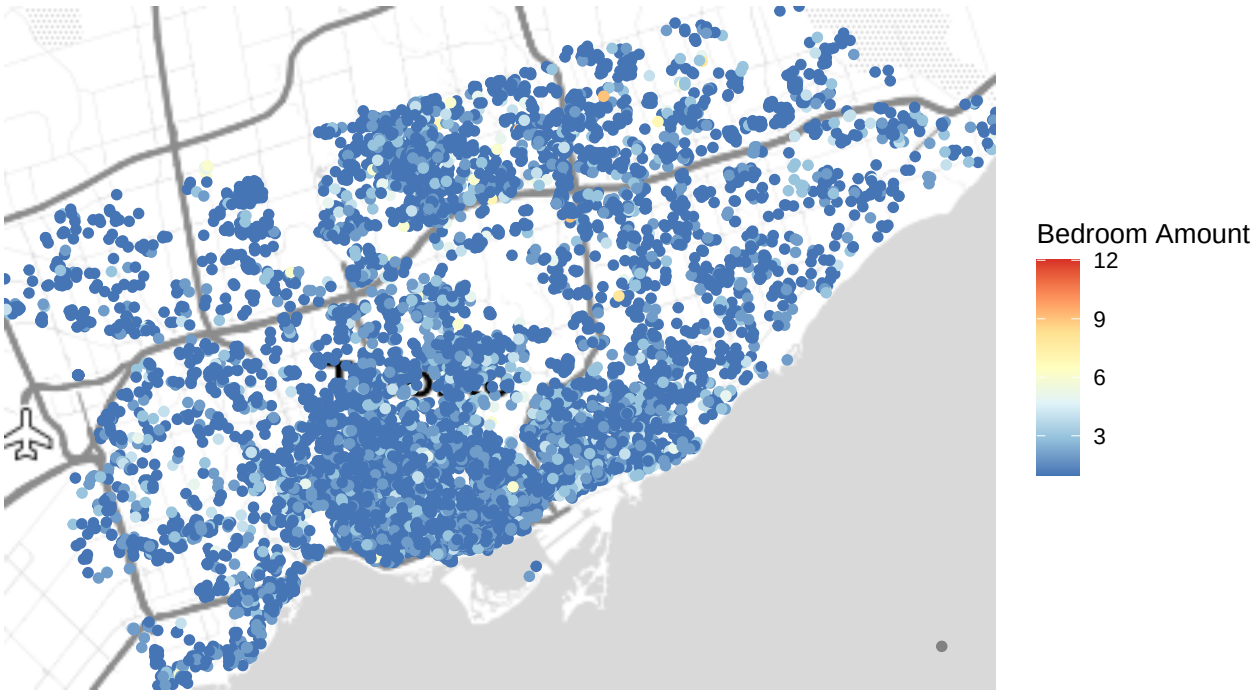

Figure 8: February 2021 listing bedroom amounts

Figure 8 shows an overwhelming number of listings that have less than ~5 bedrooms. A high amount of bedrooms is difficult to find in Toronto's housing market, and these listings reflect that.

Cleanliness score spatial view

Regardless of the pandemic, cleanliness is a good indicator of Airbnb listing quality. If a listing has received low review scores for cleanliness, it is unlikely anyone will feel comfortable staying in that property.

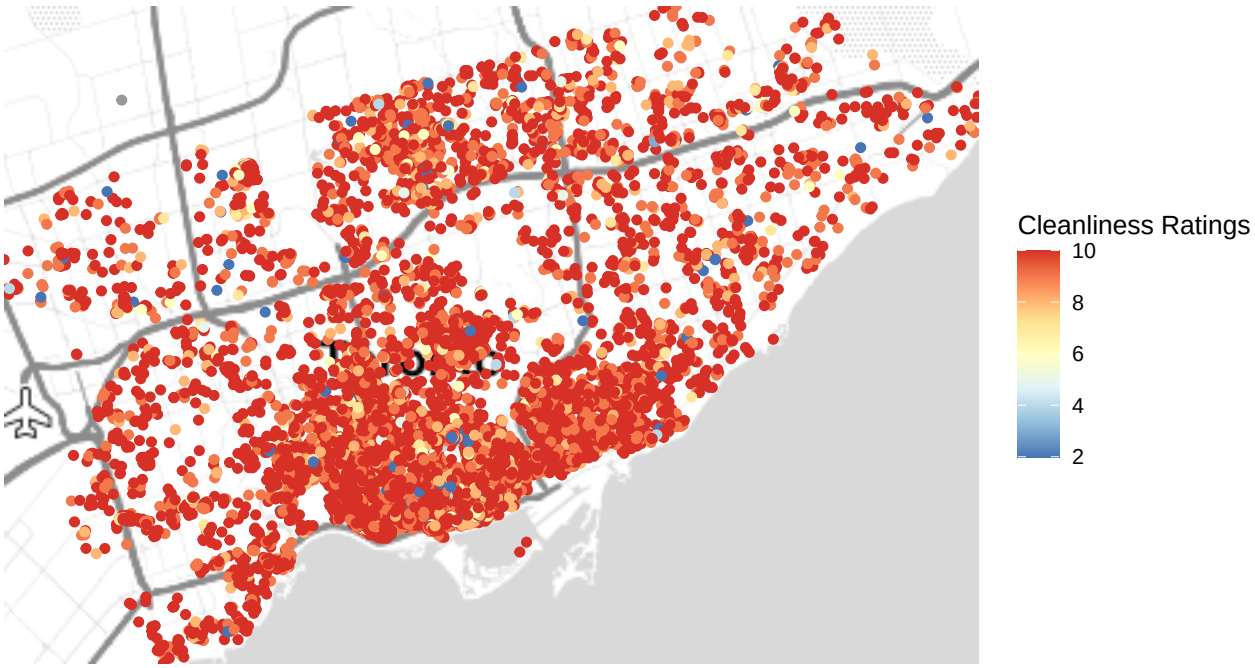

Figure 9: February 2021 listing cleanliness scores

Figure 9 shows there are many listings that have review scores for cleanliness that are 6 and below. These listings are likely not very well maintained and are possibly used just as a short-term rental property without a host living there.

Listing review score spatial view

A listing's review score indicates how comfortable that listing's experience was for a guest.

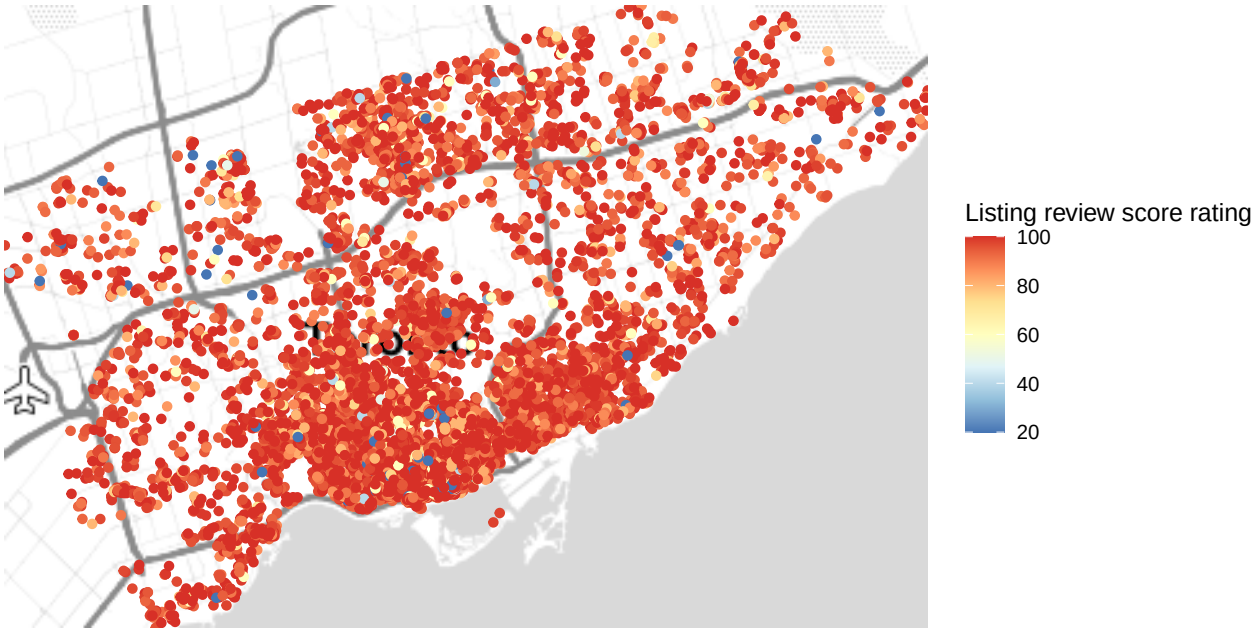

Figure 10: February 2021 listing review scores

Figure 10 shows that review scores appear to be on par with cleanliness ratings. This makes sense as a low cleanliness score would likely be associated with an overall negative stay.

2.4 Secondary data analysis

This section of the paper is focused on the full data set of the months February 2020 - April 2021. To slim the data set down and narrow focus for the paper, the neighbourhoods of interest were chosen to be neighbourhoods surrounding the University of Toronto. The neighbourhoods selected for analysis are Kensington-Chinatown, Annex, University, Church-Yonge Corridor, and Bay Street Corridor.

2.4.1 Views of price

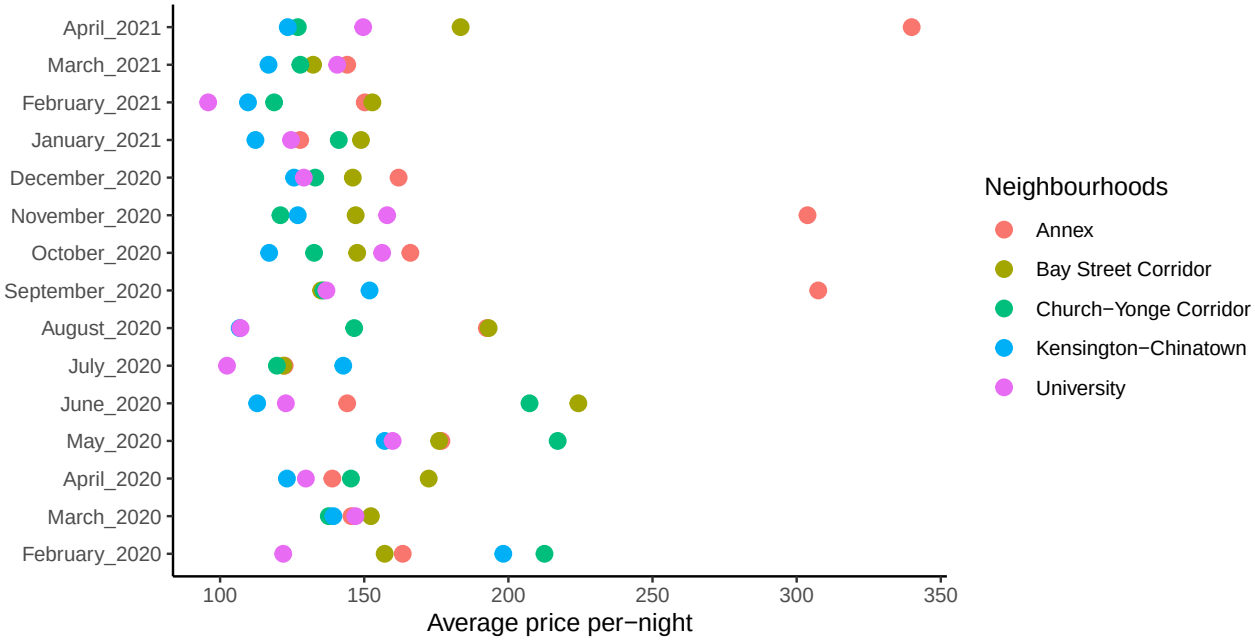

Figure 11: Average per-night price variation across Toronto neighbourhoods

When examining Figure 11, it appears Annex average prices varied greatly between each month, with fluctuations between less than $200 a night and greater than $300 a night. Bay Street Corridor prices appear to increase in the spring and summer months but stay relatively flat across the fall and winter. Church-Yonge prices appear to have an initial dip around March and April 2020, with a large increase in May and June 2020, before falling relatively flat across remaining months. Kensington-Chinatown prices appear to dip in March 2020, and fluctuate around $150 a night and below. University prices appear to dip in the summer months and the mid winter months.

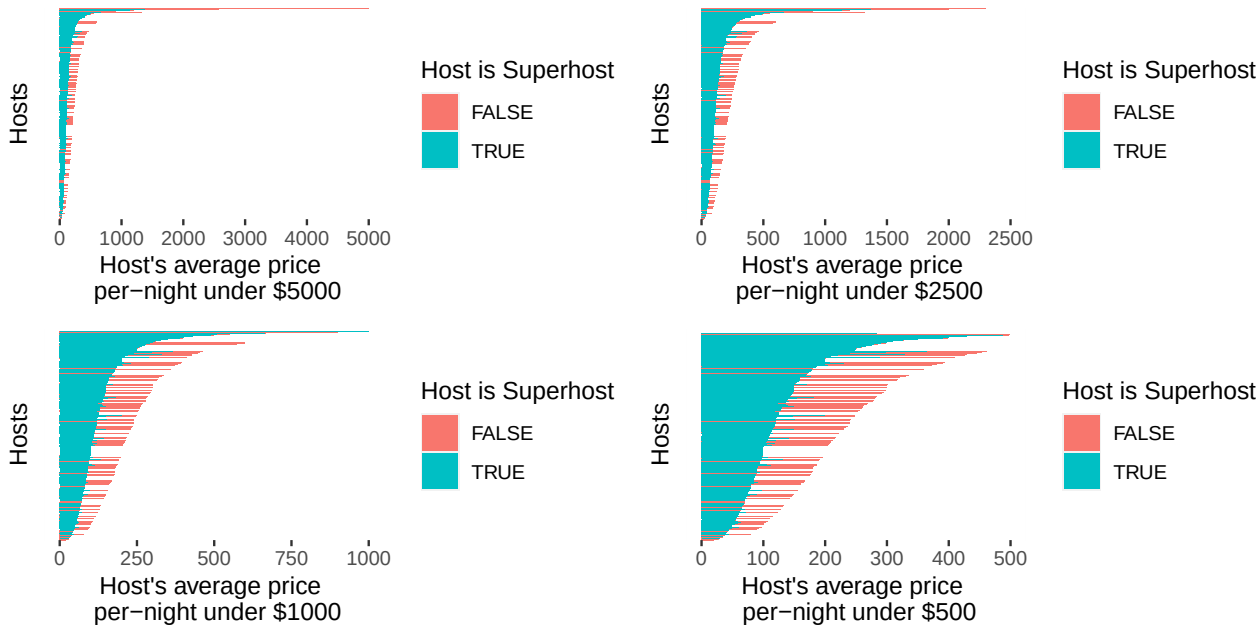

Figure 12: Average per-night price variation among Airbnb Superhosts vs Non-Superhosts

Figure 12 shows there appears to be average price variation between hosts whether the host is a Superhost or not a Superhost. It appears that many of the hosts that either have extremely high prices per night or extremely low prices per night are not Superhosts. This variation between host prices is likely as a result of the different property types that are listed under each host, as well as the different areas the listings are in.

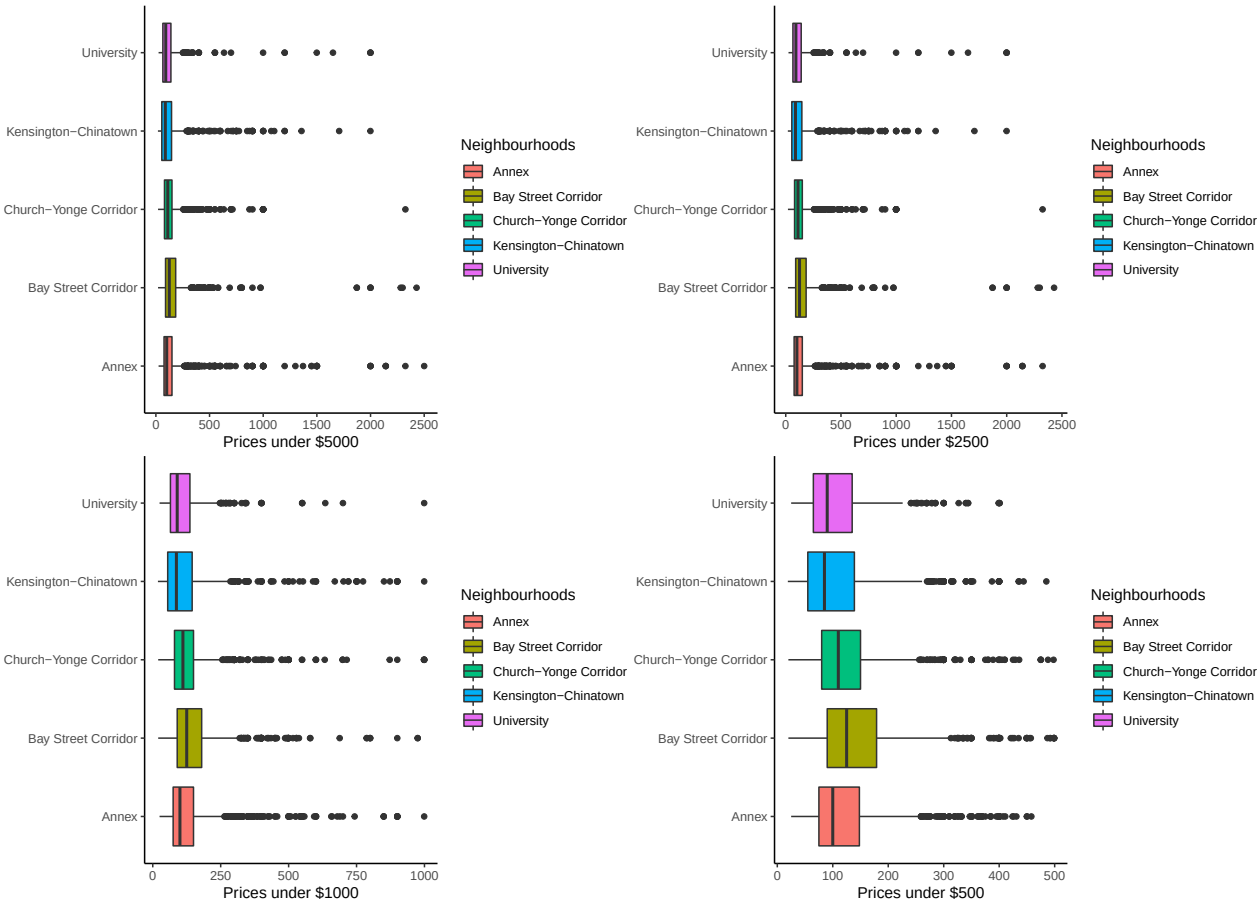

Figure 13: Airbnb price per-night boxplots by Toronto neighbourhood

It can be seen from Figure 13 that prices vary between neighbourhood. It appears Bay Street Corridor has higher prices compared to the other neighbourhoods, while Kensington-Chinatown has lower prices.

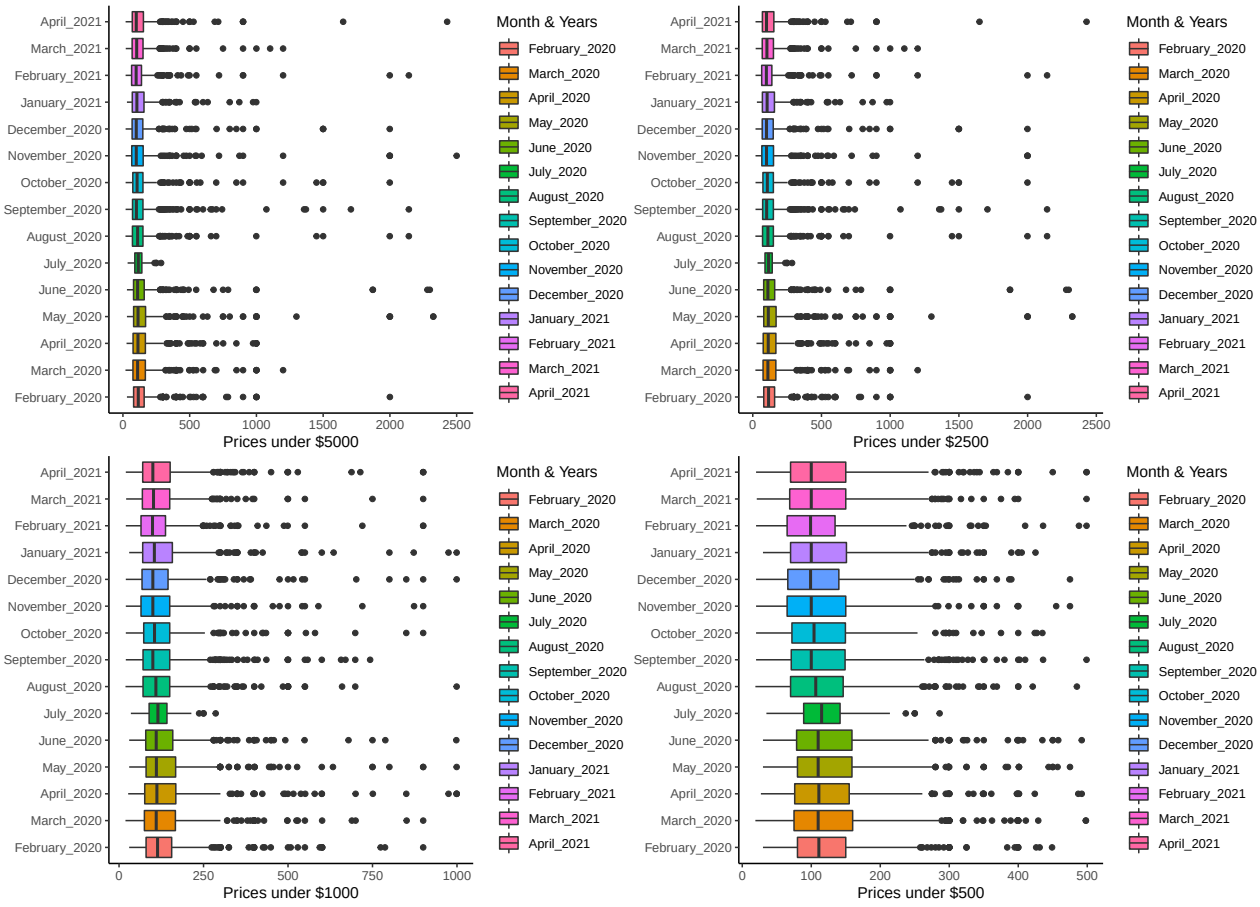

Figure 14: Airbnb price per-night boxplots by month

From Figure 14 it appears there is price variation across the different months and years. However, the price variation across the different months does not appear to be as large as the variation across the different Toronto neighbourhoods.

2.4.2 Variable selection

By this paper's definition of listing quality, these variables were selected from the data as being the most consistent with Airbnb quality:

- `bedrooms` (positive numeric variable)
- `bathrooms` (positive numeric variable)
- `neighbourhood_cleansed` (categorical variable with levels: Kensington-Chinatown, Annex, University, Church-Yonge Corridor, and Bay Street Corridor)
- `review_scores_cleanliness` (positive numeric variable between 0 and 10)
- `review_scores_rating` (positive numeric variable between 0 and 10)
- `reviews_per_month` (positive numeric variable between 0 and 10)
- `host_response_rate` (positive numeric variable between 0 and 100)
- `minimum_nights` (positive numeric variable)

- `maximum_nights` (positive numeric variable)
- `host_is_superhost` (binary variable with levels: TRUE and FALSE)

Both backwards selection and forwards selection stepwise regression with AIC criteria is ran to select the most relevant variables to predict price from this list.

2.4.2.1 Backwards selection stepwise results

Step	Variable Removed	R-Square	Adj. R-Square	C(p)	AIC	RMSE
1	review_scores_rating	0.0349	0.0268	-5.9853	46052.7716	367.7388
2	minimum_nights	0.0348	0.0271	-7.8873	46050.8705	367.6856
3	maximum_nights	0.0347	0.0273	-9.4504	46049.3113	367.6523
4	month_year	0.030	0.0269	3.7738	46036.6323	367.7249

2.4.2.2 Forward selection stepwise results

Step	Variable Entered	R-Square	Adj. R-Square	C(p)	AIC	RMSE
1	bathrooms	0.0371	0.0370	1205.025	107277.8152	397.1688
2	bedrooms	0.0428	0.0426	1695.2279	99424.9512	411.7627
3	neighbourhood_cleansed	0.0465	0.0457	1665.1344	99407.2845	411.0958
4	host_response_rate	0.0261	0.0243	394.7360	57472.7556	385.9262
5	reviews_per_month	0.0291	0.0267	-30.6483	46514.7980	365.8379

The results from both stepwise regressions match. Both stepwise regressions selected the following variables for the model:

- `bedrooms`
- `bathrooms`
- `neighbourhood_cleansed`
- `month_year`
- `reviews_per_month`
- `host_is_superhost`
- `host_response_rate`
- `review_scores_cleanliness`

3 Models

We will fit two linear mixed effects model to the data. Linear mixed effects models are chosen because we know that observations are correlated in different cluster levels within the data and ignoring these correlations may present issues.

The first cluster would be neighbourhoods. The neighbourhoods of interest for this paper are the neighbourhoods surrounding the University of Toronto. These neighbourhoods being Kensington-Chinatown, Annex, University, Church-Yonge Corridor and Bay Street Corridor. Listings within each neighbourhood are going to be more similar to each-other than they are to listings in other neighbourhoods. Part of this is due to Toronto real-estate development and city planning. For example, if an area consists of primarily office buildings, the listings in that area will be more similar to each other than they would be to an area that consists of primarily single family housing and schools.

The second cluster would be hosts. There is great variation between different hosts that have their listings on Airbnb. Some hosts have multiple listings throughout Toronto, other hosts have only one listing. As there are sometimes multiple observations listed for a single host, the models must account for this correlation within the listings a host has posted and the variation between different hosts on Airbnb.

3.1 Model 1

$$\begin{aligned}\text{price}_{ijk} = \beta_0 + \beta_1\text{Bedrooms}_{ij} + \beta_2\text{Bathrooms}_{ij} + \beta_3\text{ReviewsPerMonth}_{ij} + \beta_3\text{HostIsSuperhost}_{ij} \\ +\beta_4\text{HostResponseRate}_{ij} + \beta_5\text{ReviewScoresCleanliness}_{ij} + \beta_6\text{MonthYear}_{ij} \\ +\text{Neighbourhood}_i + \text{HostID}_{ij}\end{aligned} \tag{1}$$

Where:

- Neighbourhood_i refers to the random effect for the i^{th} neighbourhood.
- HostID_{ij} refers to the random effect for the j^{th} host in the i^{th} neighborhood.
- price_{ijk} refers to the price measured on a per-night scale of the k^{th} listing from the j^{th} host in the i^{th} neighborhood.

3.2 Model 2

A linear mixed effects model will be included with indicators pertaining to Toronto's stay-at-home orders and declarations of states of emergency.

The first indicator $\text{StayAtHomeOrder1}_{ij}$ represents Toronto entering the first state of emergency from March 2020 until July 2020.

The second indicator $\text{StayAtHomeOrder2}_{ij}$ represents Toronto entering the second declaration of emergency and stay-at-home orders in January 2021 which lasted until March 2021.

The third indicator $\text{StayAtHomeOrder3}_{ij}$ represents when in April 2021 Ford announced the province would move into White Tier "Shutdown", which would impose the same measures put in place during the original province wide shutdown in March 2020.

These indicators are included in place of the time variable MonthYear_{ij}.

$$\begin{aligned}\text{price}_{ijk} = \beta_0 + \beta_1\text{Bedrooms}_{ij} + \beta_2\text{Bathrooms}_{ij} + \beta_3\text{ReviewsPerMonth}_{ij} + \beta_3\text{HostIsSuperhost}_{ij} \\ +\beta_4\text{HostResponseRate}_{ij} + \beta_5\text{ReviewScoresCleanliness}_{ij} + \\ \beta_6\text{StayAtHomeOrder1}_{ij} + \beta_7\text{StayAtHomeOrder2}_{ij} + \\ \beta_8\text{StayAtHomeOrder3}_{ij} + \text{Neighbourhood}_i + \text{HostID}_{ij}\end{aligned} \tag{2}$$

Where:

- Neighbourhood_i refers to the random effect for the i^{th} neighbourhood.
- HostID_{ij} refers to the random effect for the j^{th} host in the i^{th} neighborhood.
- price_{ijk} refers to the price measured on a per-night scale of the k^{th} listing from the j^{th} host in the i^{th} neighborhood.

4 Results

4.1 Model 1 results

Fixed Effects	Estimate	Standard Error	T-Value	95% CI	P-value
bedrooms	51.47	11.1533	4.615	(29.61, 73.33)	4.18e-06

Fixed Effects	Estimate	Standard Error	T-Value	95% CI	P-value
bathrooms	57.94	16.5923	3.492	(25.42, 90.46)	4.91e-04
reviews_per_month	-7.40	3.9189	-1.889	(-15.08, 0.28)	0.059
host_is_superhostTRUE	-6.60	15.9865	-0.413	(-37.93, 24.74)	0.680
host_response_rate	-0.66	0.4065	-1.612	(-1.45, 0.14)	0.107
review_scores_cleanliness	5.24	7.4710	0.701	(-9.41, 19.88)	0.483
month_year [April_2021]	-16.24	34.9639	-0.464	(-84.77, 52.29)	0.642
month_year [August_2020]	-5.75	33.2231	-0.173	(-70.86, 59.37)	0.863
month_year [December_2020]	-32.61	33.9291	-0.961	(-99.11, 33.89)	0.336
month_year [February_2020]	-34.49	47.1185	-0.732	(-126.84, 57.86)	0.464
month_year [February_2021]	-34.39	35.3697	-0.972	(-103.71, 34.94)	0.331
month_year [January_2021]	-49.17	34.7210	-1.416	(-117.22, 18.88)	0.157
month_year [July_2020]	-20.49	108.7913	-0.188	(-233.71, 192.74)	0.851
month_year [June_2020]	36.45	29.6054	1.231	(-21.58, 94.47)	0.218
month_year [March_2020]	-58.54	47.1834	-1.241	(-151.02, 33.94)	0.215
month_year [March_2021]	-46.06	37.1784	-1.239	(-118.93, 26.81)	0.215
month_year [May_2020]	33.82	27.2523	1.241	(-19.60, 87.23)	0.215
month_year [November_2020]	-20.05	32.5108	-0.617	(-83.77, 43.67)	0.537
month_year [October_2020]	-30.38	33.2309	-0.914	(-95.51, 34.75)	0.361
month_year [September_2020]	1.71	31.5745	0.054	(-60.17, 63.60)	0.957

Total number of observations	Random Effects	Variance	Number of Groups
3141	HostID	16200.78	1054
	Neighbourhood	86.81	5
	Residual	114084.70	

Reading the model results, it can be shown that there are 3141 observations passed through the model, with 1054 groups for the host random effect and 5 groups for the neighbourhood random effect. There is much higher variation between hosts than there is between neighbourhoods. This can by the variation of 16200.78 for the host random effect, when compared to the variance of 86.81 for the neighbourhood random effect.

The model results show that of the variables selected to measure Airbnb quality, bedroom and bathroom amounts were statistically significant at the 5% significance level in predicting price. Bedrooms is the most statistically significant variable with t-value 4.615. The bedrooms variable has $\beta = 51.47$ with 95% CI (29.61, 73.33) and p-value 4.18e-06. The bathrooms variable is the second most statistically significant variable in the model with t-value of 3.492. Bathrooms has $\beta = 57.94$ with 95% CI (25.42, 90.46) and p-value 4.91e-04. The model results also show that out of all the other variables, reviews per month is the third most relevant variable in the model for the prediction of price with a t-value of -1.889. Reviews per month has $\beta =$ -7.40 with 95% CI (-15.08, 0.28) and p-value 0.059. While reviews per month is not significant on the 5% significant level, reviews per month is significant on the 10% significant level.

Based on how the variables currently enter the model, it does not appear that if a host is a superhost, a host's response rate, and the listing's review scores for cleanliness have an effect on an Airbnb listing's price. All three of these variables have p-values that are higher than the 5% significance level. However, out of these three variables, host response rate has the highest t-value of -1.612. Further, it can be shown that by the way that they currently enter the model, there does not appear to be much of an effect from the different months as each month and year is statistically insignificant at the 5% significance level. Future work could examine this in more detail to further explore the effect of seasonality in Airbnb prices.

4.2 Model 2 results

Fixed Effects	Estimates	Standard Error	T-Value	95% CI	P-value
bedrooms	49.83	11.1133	4.484	(28.05, 71.61)	7.75e-06
bathrooms	59.81	16.5440	3.615	(27.39, 92.24)	3.08e-04
reviews_per_month	-7.64	3.8735	-1.971	(-15.23, -0.04)	0.049
host_is_superhostTRUE	-7.84	15.8672	-0.494	(-38.94, 23.26)	0.621
host_response_rate	-0.19	0.1580	-1.213	(-0.50, 0.12)	0.225
review_scores_cleanliness	4.44	7.4328	0.598	(-10.12, 19.01)	0.550
first_state_of_emergency	18.66	14.3879	1.297	(-9.54, 46.86)	0.195
second_state_of_emergency	-32.52	20.1444	-1.614	(-72.00, 6.96)	0.106
third_state_of_emergency	-8.11	30.5720	-0.265	(-68.03, 51.81)	0.791

Total number of observation	Random Effects	Variance	Number of Groups
3141	HostID	16118.98	1054
	Neighbourhood	105.84	5
	Residual	114051.39	

Similar to model 1, in model 2 there is much higher variation between hosts than there is between neighbourhoods. This can by the variation of 16118.9 for the host random effect, when compared to the 105.84 neighbourhood random effect.

The model 2 results show that of the variables selected to measure Airbnb quality, bedroom amount, bathroom amount, and reviews per month were statistically significant at the 5% significance level in predicting price.

Bedroom amount is the most statistically significant variable with t-value 4.848. Bedrooms has $\beta = 49.83$ with 95% CI (28.05, 71.61) and p-value 7.75e-06. Bathroom amount is the second most statistically significant variable in the model with t-value of 3.615. Bathrooms has $\beta = 59.81$ with 95% CI (27.39, 92.24) and p-value 3.08e-04.Reviews per month is the third most statistically significant variable in the model with t-value -1.971. Reviews per month has β = -1.971 with 95% CI (-15.23, -0.04) and p-value 0.049.

Similar to model 1, based on how the variables currently enter the model, it does not appear that if a host is a superhost, a host's response rate, and the listing's review scores for cleanliness have an effect on an Airbnb listing's price. All three of these variables have p-values that are higher than the 5% significance level. However, out of these three variables, host response rate has the highest t-value of -1.213.

Similar to how it was with the different months in model 1, each Toronto state of emergency is statistically insignificant at the 5% significance level. Similar to seasonality, future work can explore the Toronto lockdown effect in more detail with a different modelling strategy.

5 Discussion

5.1 Listing price's relation with listing quality

This paper finds that out of the respective variables to measure Airbnb listing quality, bedroom and bathroom amount have the greatest effect on price. Looking at the results of both models, we can see that the for every increase of 1 in a listing's bedroom amount or bathroom amount, the price of the listing increases by an average of around $49.83 to $59.81 per-night. Both model results also show that a higher bathroom amount in a unit leads to the highest increase in an Airbnb listing's price compared to the other predictors. These results are on par with the paper's definition of Airbnb quality. It is believed that a higher number of bedrooms and bathrooms would indicate a listing of higher quality, and a listing of higher quality would have a higher price per-night.

This paper also finds that for every increase in reviews per month by 1, there is a decrease in Airbnb listing price per-night by an average of around $-7.64 to $-7.40 per-night. These results suggest that a lower price

may indicate a higher amount of reviews per month as there are likely more bookings occurring at that listing due to the lower price. These results indicate that if an Airbnb listing has a higher amount of reviews per month, their increased traffic has likely resulted in a slight decrease in price.

Both of the paper's model results also indicate that if a host is a Superhost, there is a decrease in listing price per-night by an average of around $-7.84 to $-6.60 compared to if a host was not a Superhost. This can likely be attributed to Superhosts understanding the market value of their property, and attempting to make their price more attractive to Airbnb customers. Also, Superhosts for Airbnb have had to achieve a certain amount of reserved stays to qualify for Superhost status. According to the Airbnb website, to be eligible to become a Superhost, you must have completed 3 reservations that total at least 100 nights (Airbnb, 2021). This may suggest the slightly lower average price per-night if the host would like to retain their Superhost status.

The paper finds that for an increase in cleanliness review scores by 1, on average there is an increase in price per-night of around $4.44 to $5.24. This result seems to be on par with the paper's definition of quality, and that a higher quality listing would have a higher review score for cleanliness and in turn a higher price. Based on this data, this result suggests that Airbnb listings that place more value in cleanliness may have slightly higher prices per-night.

Overall, the paper's results seem to be on par with the paper's discussion of quality. Higher prices were associated with attributes that represent a listing of higher quality. However, it was seen that if a host is a Superhost and the amount of reviews of a listing per month led to slightly lower prices. This can likely be attributed to a host understanding the market that surrounds their listing, and even if the listing is of high quality, the host makes the listing affordable to attract business. However, the results did not indicate month and year being significant in predicting Airbnb listing price. This could be a result of the paper failing to capture the seasonality with the generalized linear mixed model that was fit.

5.2 What is next for Toronto's Airbnb market?

COVID-19 has changed the Airbnb market in Toronto from what it once was. According to the Canadian Press, throughout COVID-19, there has been an influx of former Airbnb units that are being moved to being long-term rentals or being sold (Balakrishnan, 2020). According to a report of July 2020 data from Toronto's Regional Real Estate board, there was a spike in sale listings within Toronto condos buildings that "allow, previously allowed or are known to be popular on Airbnb" (Balakrishnan, 2020). It is difficult to say what this means for Toronto's short-term rental afforablility in the long run. CTV news has found that in April 2021 average rent prices have begun to rise again in Toronto from the overall rent price decline throughout the pandemic (Katawazi, 2021). For the rest of the pandemic period, an Airbnb host may feel the need to maintain a higher listing price per-night to compensate for the infrequency of guests and the money the unit could of made on the rental market. With the easing of Ontario COVID-19 restrictions, and the eventual opening of the Canadian border for international travel, there is a chance that more Airbnb listings may appear on the market at affordable prices to appeal to travelers. However, if many Airbnb hosts feel that it is more profitable to list as a long term rental or to even sell, and if Ontario continues short-term rental restrictions, it is likely that the pandemic impacted Toronto Airbnb market will maintain its difference from its pre-pandemic success.

5.3 Weaknesses

There was a high amount of the listings in the full data-set (February 2020 - April 2021) after the neighbourhoods of interest were selected that were incomplete cases. These observations had to be discarded from the data-set used for modelling due to missing values for the variables of interest in the model. There was 7,824 listings in the full data-set, and 4,682 listings had to be omitted before being passed into both models due to these missing values. This meant that both models only were fit on 3141 listings. While still a large amount of listings, it needs to be noted that the models were fit on only 40% of the total amount of listings that were in the areas of Kensington-Chinatown, Annex, University, Church-Yonge Corridor and Bay Street Corridor listed on Airbnb anytime within the time frame of February 2020 up to and including April 2020.

July 2020 data did not contain bathroom amount. Bathroom amount for July 2020 had to be constructed

by searching for July 2020 unique listing id in other months, and using that unique listing id to map those bedroom amounts onto the July 2020 data. July 2020 listings also did not contain nightly price, but instead contained weekly price. Nightly price for July 2020 had to be constructed by averaging weekly price by 7. After omitting missing values for the data to be fed into the model, July 2020 had an extremely low amount of listings. July 2020 had only 11 listings in the data after the various cleaning and omitting of missing values, which is extremely low compared to 246 listings in June 2020 and 175 listings in August 2020. As July 2020 had such a low amount of data being passed through the model, model results for July 2020 should be met with criticism.

With how the linear mixed effect models were fit, the models failed to capture a seasonality and a lockdown effect. No month during the period of February 2020 - April 2021 was significant in the prediction of listing price, and no lockdown indicator was significant as well. The Toronto Airbnb rental market took a major hit during COVID-19, and rental prices dropped throughout Toronto. It would make sense to see some effect of time on Airbnb listing price as tourism was halted during these months and lockdowns in Toronto were implemented to stop the spread of COVID-19.

Since linear mixed effect models were used, we are able to assess regression assumptions the same way we would if modelling an ordinary multiple linear regression model. When assessing model diagnostics in the Appendix, we can see some violations of the linearity assumption and the normally distributed errors assumption, however the homoscedasticity assumption appears to be met. In Figure 15 and Figure 18, the model residuals vs the predictor is plotted, there appears to be a pattern in both plots which indicates that the linearity assumption may be violated in model 1 and 2. When assessing homoscedasticity in the Figure 16 and Figure 19, the values are mostly randomly scattered about 0 which indicates the homoscedasticity assumption is likely met in both models. When assessing normally distributed errors in the Figure 17 and Figure 20, there are heavy tails which indicates some deviation from normality.

Finally, the analysis did not contain enough pre-pandemic data to really assess how COVID-19 had impacted the Toronto Airbnb market. For a more accurate view of COVID-19's impact on listing price and quality, there should be data taken from 2017-2019 and compared to 2020.

5.4 Next Steps

For a next step, it would be helpful to fit a model of data taken from the year of 2019 to provide a more thorough comparison to model 1. It also would be helpful to further explore the effect of seasonality and the lockdowns with another model that is more sensitive to time. Lastly, a model fit on each neighborhood of interest would be helpful in providing a more accurate view of how that neighbourhood's price shifted with relation to quality over the course of the COVID-19 pandemic.

Appendix

A Model Diagnostics:

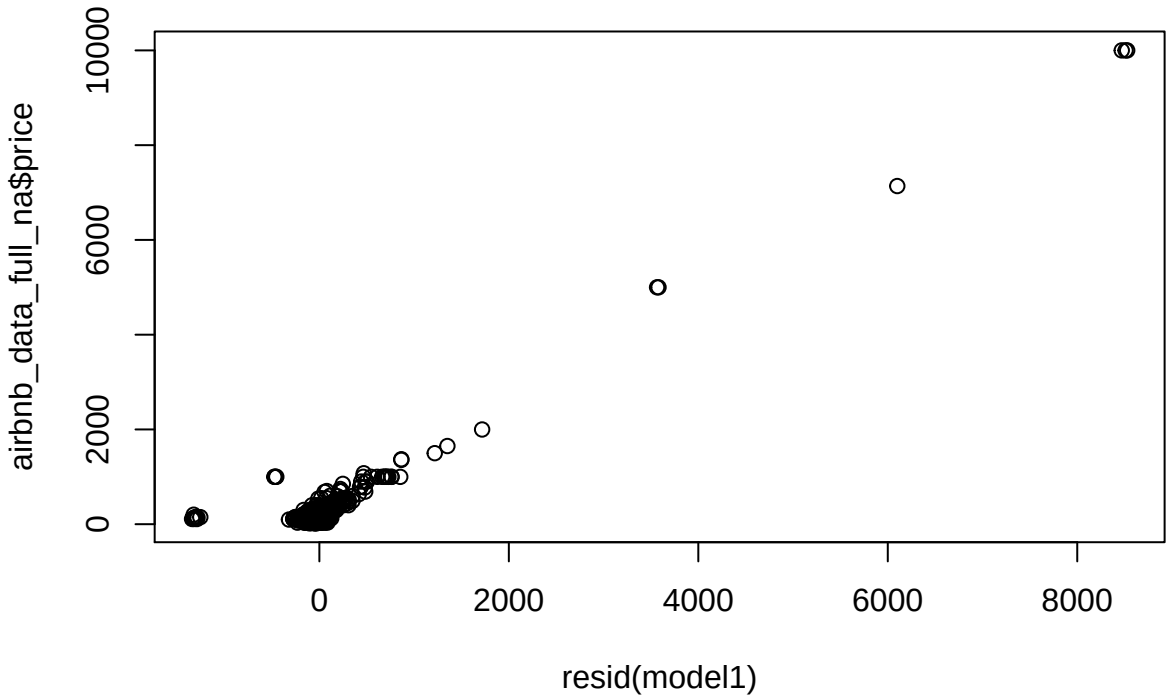

Figure 15: Model 1 residual Plot

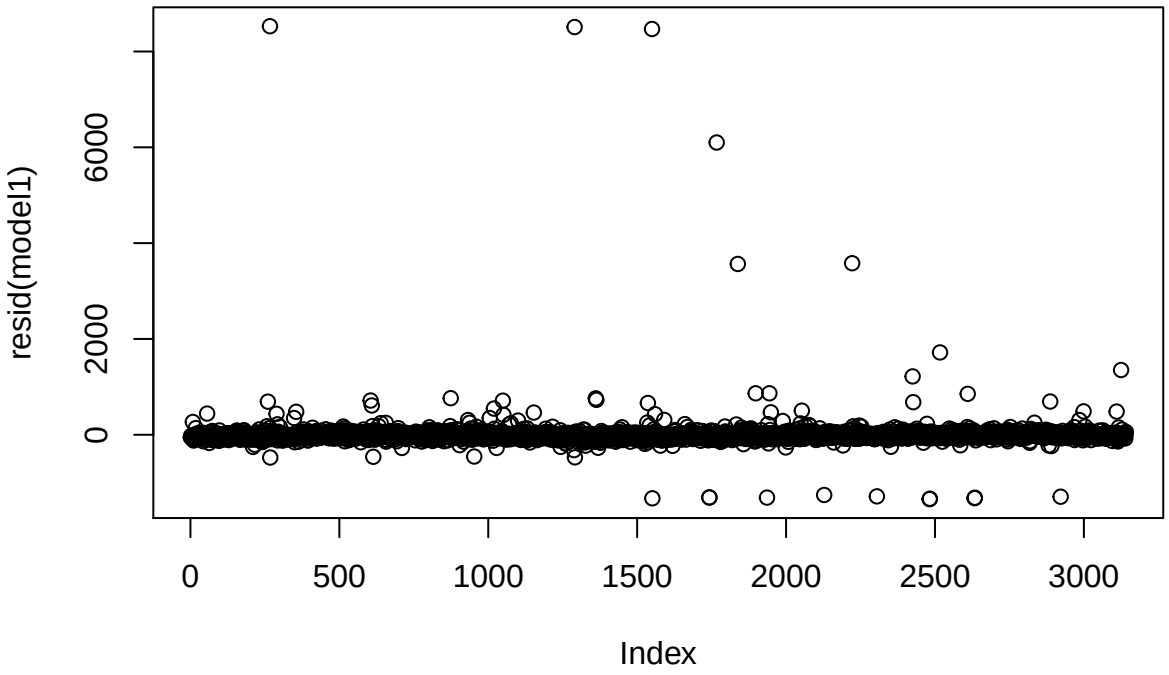

Figure 16: Model 1 fitted vs residual plot

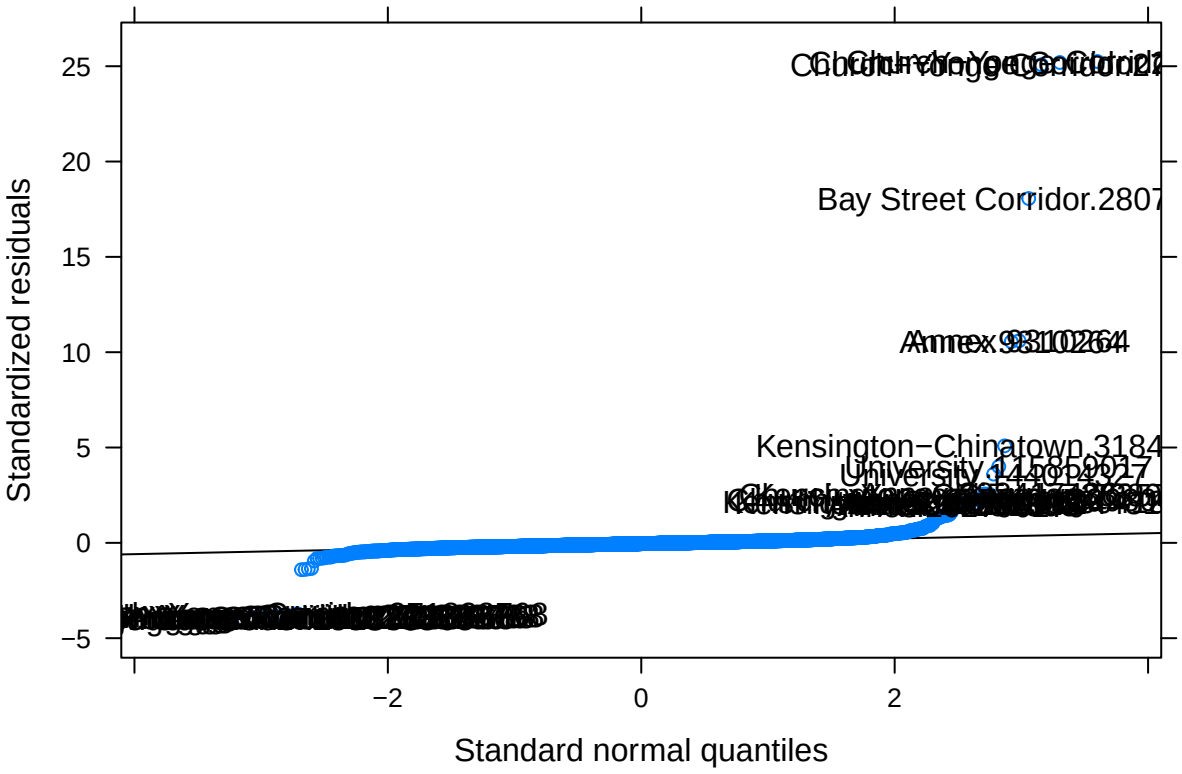

Figure 17: Model 1 QQplot

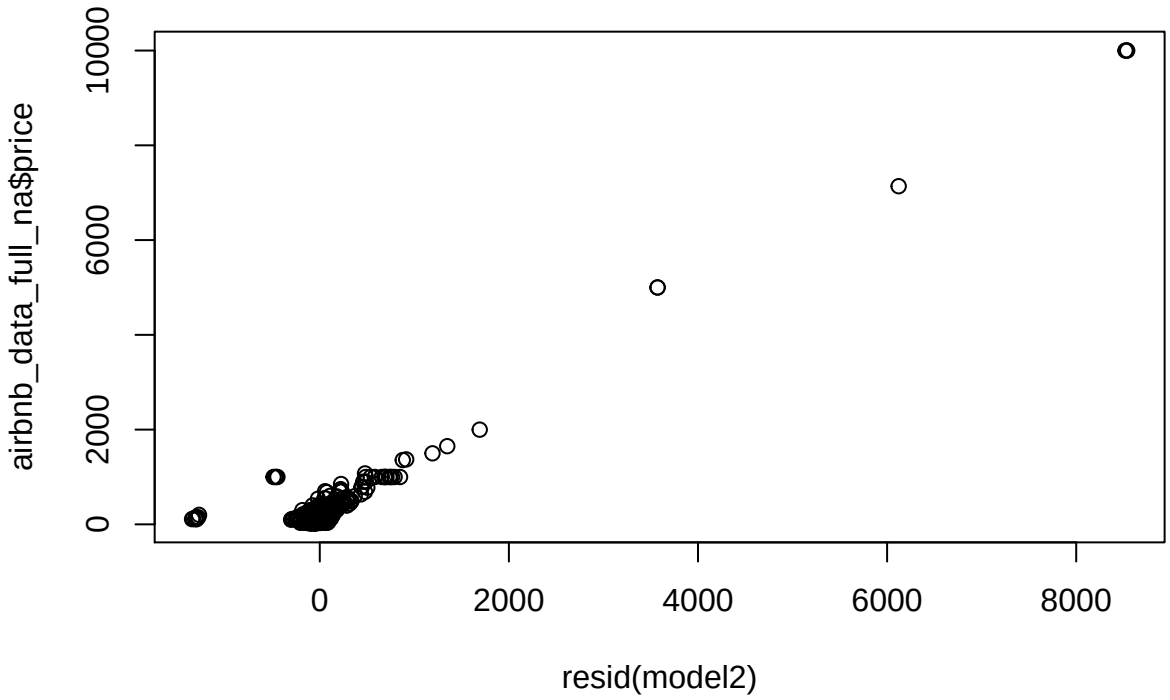

Figure 18: Model 2 residual plot

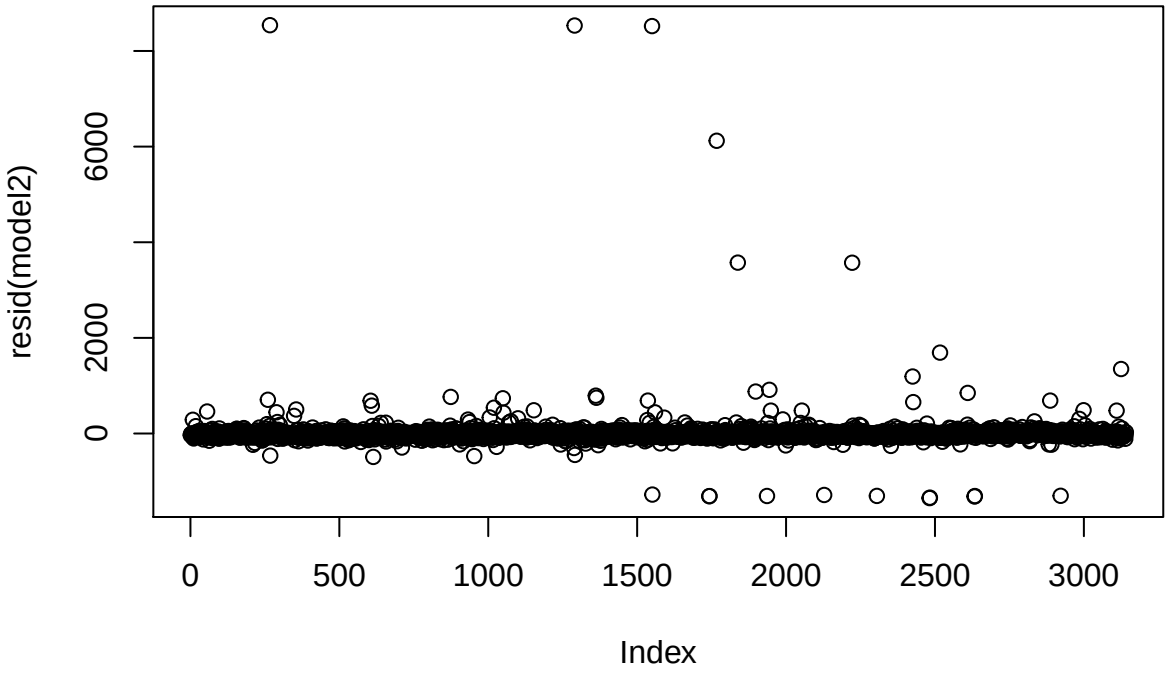

Figure 19: Model 2 fitted vs residual plot

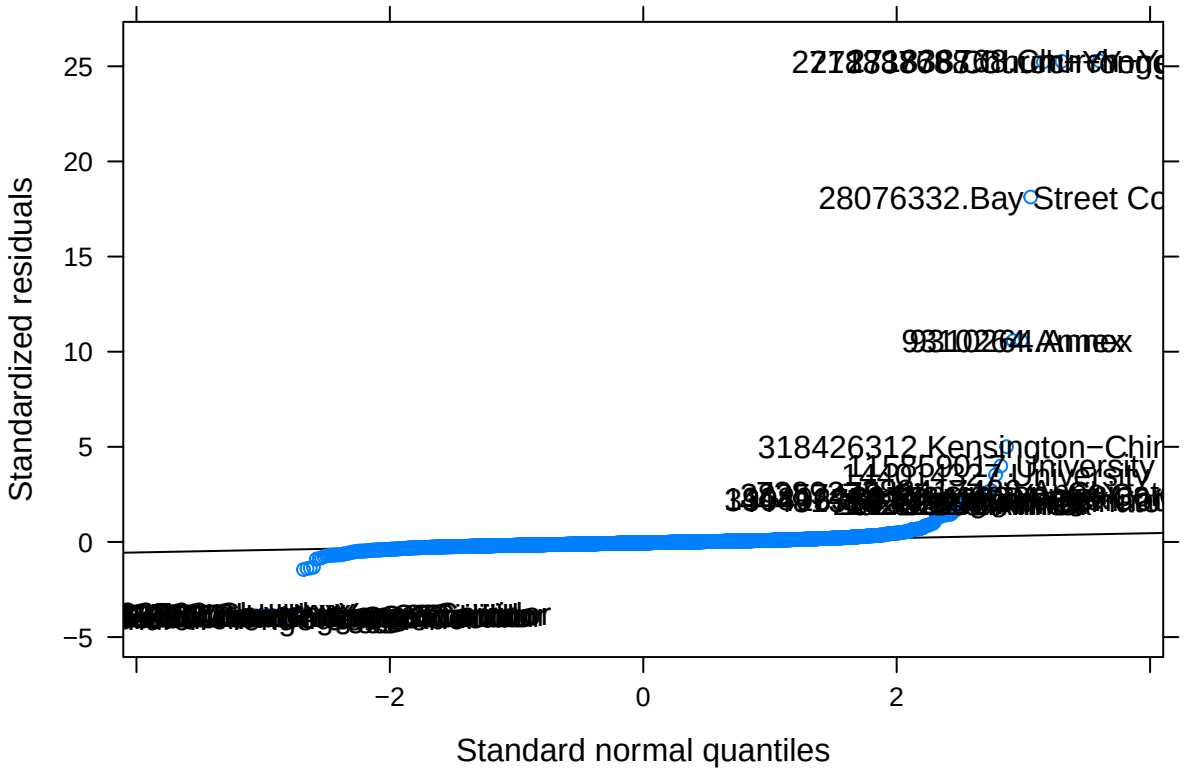

Figure 20: Model 2 QQplot

B References

Balakrishnan, A. (2020, July). Former Airbnb units in Toronto moving to long-term rental market. CP24. https://www.cp24.com/news/former-airbnb-units-in-toronto-moving-to-long-term-rental-market-1.5036301?cache=ahlalmzxbysj%3FclipId%3D104056%3FautoPlay%3Dtrue%3Fot%3DAjaxLayout%3Fot%3DAjaxLayout%3FautoPlay%3Dtrue%3FcontactForm%3Dtrue%3FclipId%3D1921747.

Bates D., Maechler M., Bolker B., Walker S. (2015). Fitting Linear Mixed-Effects Models Using lme4. Journal of Statistical Software, 67(1), 1-48. doi:10.18637/jss.v067.i01.

Deirdre Bosa, S. R. (2020, May 5). Airbnb to lay off nearly 1,900 people, 25% of company. CNBC. https://www.cnbc.com/2020/05/05/airbnb-to-lay-off-nearly-1900-people-25percent-of-company.html. Eisenberg, R. (2020, June 5). The long-term impact of covid-19 on short-term rentals. WeirFoulds LLP. https://www.weirfoulds.com/the-long-term-impact-of-covid-19-on-short-term-rentals.

Helliwell, J. F., Schellenberg, G., & Fonberg, J. (2020). Life Satisfaction in Canada Before and During the COVID-19 Pandemic. Social Analysis and Modelling Division Statistics Canada, 457.

Hu, M and Lee, A., Airbnb, COVID-19 Risk and Lockdowns: Local and Global Evidence (2020). Available at SSRN: https://ssrn.com/abstract=3589141 or http://dx.doi.org/10.2139/ssrn.3589141

Kassambara A. (2020). ggpubr: 'ggplot2' Based Publication Ready Plots. R package version 0.4.0. https://CRAN.R-project.org/package=ggpubr

Katawazi. (2021, April 14). Toronto rent prices begin to go back up After plummeting for more than a year. CTV News Toronto. https://toronto.ctvnews.ca/toronto-rent-prices-begin-to-go-back-up-after-plummeting-for-more-than-a-year-1.5387285.

Kahle D., Wickham H., ggmap: Spatial Visualization with ggplot2. The R Journal, 5(1), 144-161. URL http://journal.r-project.org/archive/2013-1/kahle-wickham.pdf

Liu, H. (2020). The Economic Impact of Travel Restrictions on the Canadian Economy due to the COVID-19 Pandemic. Economic Insights.

Mo, G., Cukier, W., Atputharajah, A., Boase, M. I., & Hon, H. (2020). Differential impacts During COVID-19 in Canada: A look at diverse individuals and their businesses. Canadian Public Policy, 46(S3). https://doi.org/10.3138/cpp.2020-072

R Core Team (2021). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

Stevens, P. (2020, August 27). Toronto introduces series of short-term Rental reforms from 10 September. Short Term Rentals. https://shorttermrentalz.com/news/toronto-rental-reforms-september/.

What is a Superhost? - Airbnb help centre. Airbnb. (2021). https://www.airbnb.ca/help/article/828/what-is-a-superhost.

Wickham et al., (2019). Welcome to the tidyverse. Journal of Open Source Software, 4(43), 1686, https://doi.org/10.21105/joss.01686