The Effect of COVID-19 on the Airbnb Market ー Comparing November 2019 to December 2020

Radhika Rita Kulkarni

Statistics Department, Student ID: 917212169, Email: rrkulkarni@ucdavis.edu

This article addresses the effect of the coronavirus pandemic on the temporary housing market of Airbnb in New York City from 2019 to 2020, with the context of the gig economy and gentrification in the city via postal code extrapolation. Our analysis identified that price per night after the pandemic decreased compared to before, which lines up with the economic recession as a result of stay-at-home orders and shutdowns, and that the minimum number of nights to stay had fewer offerings for a small number of nights compared to pre-pandemic, but more over 30 minimum night stays due to new Airbnb policy enacted in December 2019. Additionally, the pandemic has caused many inequities to be exacerbated, and in the Airbnb market, we found that hosts generally profited from the pandemic regardless of where their listings were located, rich or poor postal codes. We also found that the number of available days for listings in poor neighborhoods and rich neighborhoods were significantly different from before to after the pandemic. These findings are useful in further investigation into the continuation of the current pandemic in the near future and can aid Airbnb in extrapolating this relatively small-scale study to larger locales.

I.   INTRODUCTION

The following paper addresses the question of how the COVID-19 pandemic has changed the Airbnb market in New York City in the space of one year, from the end of November 2019 to beginning of December 2020. The COVID-19 pandemic became a worldwide phenomenon in March 2020, but was identified in December 2020, and its effects on public health and safety resulted in a complete overhaul of hygienic practices and social life. Most people in the U.S. and other countries were governmentally mandated to abide by shelter-in-place restrictions and social distancing, which led to virtual learning and work, the continuation of the work week into the weekend, and stay-at-home vacations rather than urban tourism and globe-trotting adventures. It is logical then, to assume that, as a result of these restrictions, the number of people globally who would travel to New York City and rent an Airbnb for a certain number of days would be reduced from pre-pandemic levels. Does this mean that hosts, or people who list their properties for rent on Airbnb, would charge different prices due to the economic recession? Would they offer fewer listings or change the minimum number of nights to stay? And, during a pandemic, is it possible that Airbnb rentals could have become more gentrified such that only the rich or medically fortunate would have the luxury of renting a vacation residence in an economically and socially unstable period?

The long-term impact of these questions is undeniably crucial to how Airbnb will operate in the future. This study is a good microcosm of how the COVID-19 pandemic has impacted a relatively small market of a large urban area, which can be extrapolated to larger markets all over the world.

II.   BACKGROUND

Since the start of the widespread recognition of the Covid-19 pandemic around March 2020, much research and discussion has been centered on its impact. We review literature on tourism, hotels, and gig-work in the Covid era, as each of these topics are somehow related to Airbnb and its business model.

It was not unexpected by researchers that a global pandemic would reduce tourism and upend the industry of hotels and hotel-style housing. As international flights were rapidly cancelled and tourist attractions shut down, tourism came to a standstill. Hotels were predicted in March 2020 to "be the first property type to be affected by the coronavirus due to reduced tourism and travel, and a slowdown in economic activity," with the virus expected to "exacerbate hotel cash flow declines and rising expenses from wages and real estate taxes" (Jiang et al. 2020). Since, like hotels, Airbnb offers a form of temporary housing, also used often by tourists, we asked if the company would be similarly affected in one of the most famous tourist destinations in the world: New York, New York.

Like on-demand transportation or delivery service companies like Uber, Lyft, or Doordash; Airbnb fulfils its brand of supplying vacation housing through gig workers: temporary, un-benefited independent contractors who often use their own property to provide the service. One study from August 2020 on the effect of Covid-19 on the gig economy concluded that "there is no long-term relationship between Covid-19 incidence and Gig economy," (Umar et al., 2020) even stating that "platform economy has been affected positively by Covid-19 in the short run". This implies, in contradiction to many other studies on the topic, that the numbers of listings may not change before and after the pandemic.

42   A different study found that "the global panic associated with COVID-19 may have enduring

43   consequences on travel" that will require extreme measures to counteract, which somewhat

44   contradicts the conclusion of the aforementioned study by assessing the impact of Covid on travel

45   to be negative in the short and long term.

46   So, by employing gig workers as well as being driven by travel and tourism, it appears that

47   Airbnb should be, respectively, positively and negatively affected by the pandemic. Through our

48   own analysis of Airbnb's listing data in NYC, which we will take to be a microcosm of its service in

49   cities around the world, we hope to understand the true impact of the Coronavirus on Airbnb—

50   whether positive or negative.

51   **III.   DESCRIPTIVE ANALYSIS**

52   **A. Data Cleaning**

53   After acquiring the listing data supplied by Airbnb for 2019 and 2020, we restrict and restructure

54   the data for reasonable analysis. First, we perform minor cleaning by dropping those features present

55   in one dataset but not the other, removing listings marked as hidden by Airbnb, and formatting

56   numeric variables into numbers.

57   To the end of having the most up-to date data possible, we restrict to only listings that have

58   been reviewed in the year on which the sample was taken: 2019 and 2020, respectively, for the 2019

59   and 2020 datasets. This way, listings that have not been reviewed for several years, and that might be

60   inactive, are removed. As the data contains several outliers, skewed high in terms of price, we also

61   remove any listings that have nightly prices greater than $1000.

62   Using Python's Geopy package, we perform reverse geolocation on the latitude and longitude

63   values provided for each listing to extract the postcode and the borough of New York City in which

64 they are located. We then scrape the richest postcodes from the Internal Revenue Service through

65 zipdatamaps.com and label the listings as rich or not rich.

66 **B. Data Cleaning**

67 After cleaning, our variables are limited to:

68 - id, a unique, numeric identifier for each listing

69 - name, a string containing the name and description of the listing

70 - host_id, a unique numeric identifier for each host. This is not unique to listings, as

71 multiple properties can be attributed to a single host

72 - host_name, the name of the host

73 - neighbourhood, the neighbourhood in which the listing is located

74 - latitude, the latitude element of the listing's geographic coordinates

75 - longitude, the longitude element

76 - room_type, type of room: either 'Entire home/apt', 'Private room', or 'Shared room'.

77 The Post-Pandemic dataset has the additional category of 'Hotel room'

78 - price, price per night limited to less than $1000

79 - minimum_nights, minimum number of nights the host requires an occupant to rent a

80 listing

81 - number_of_reviews, total number of reviews for each listing

82 - last_review, most recent date a listing was reviewed, restricted to the year in question

83 - reviews_per_month, average reviews per month for a listing

84 - calculated_host_listings_count, number of total listings in the dataset for the host of the

85 current listing

86   • availability_365, number of days in the year during which the listing is available to rent. After cleaning, most zero availability listings are occupied rather than inactive

88   • postcode, postal code, extracted via reverse geolocation from latitude and longitude

89   • rich, indicator of whether or not the postcode in which the listing is located falls into a set of 'rich' postal codes as defined by the zipdatamap.com through the IRS

91   • borough, the borough of New York City in which the listings is located, or the suburb if the listing is not located in NYC. Also extracted via reverse geolocation

93   **IV.   EXPLORATORY DATA ANALYSIS**

94   As our question of interest regards how the Airbnb landscape looks before and after the pandemic, we generate various plots of the different features using pre- and post- pandemic data in order to see if there is a visual difference between the years.

97   We are additionally interested in the difference in these differences when wealth is concerned: whether the listings are in a wealthy neighborhood or not; as well as the means of the host, which we approximate by the number of properties they have posted.

100   **A. Listing Locations and Availability Before and After the Pandemic**

101   *1. Nonzero availability lost and new listings by wealth*
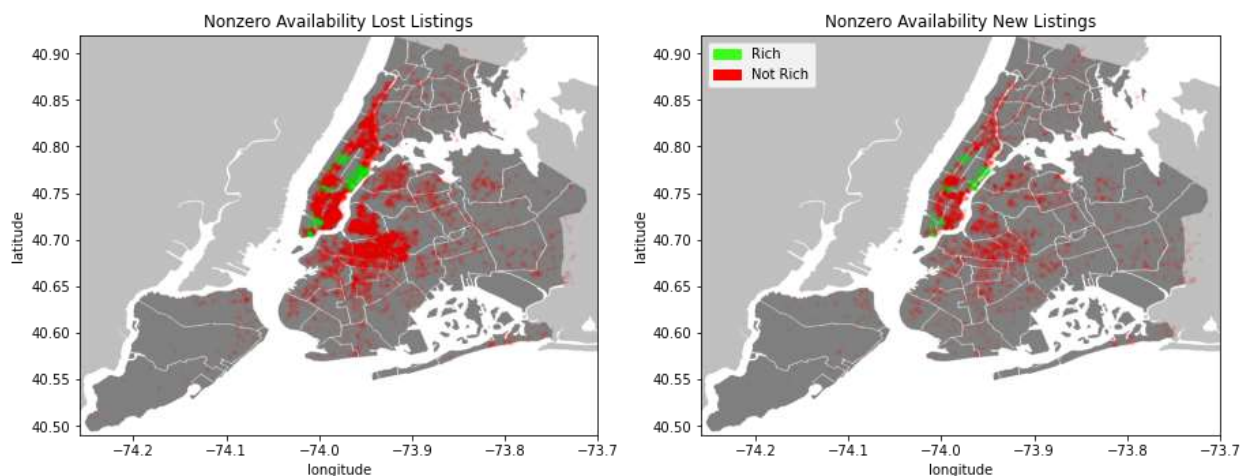

102

103   FIG. 1. A map of lost (present only in 2019, before the pandemic) and New (only in 2020,

104   during the pandemic) nonzero availability listings, colored by whether they are located in a wealthy

105   postal code.

106   We use the unique IDs of the listings to identify which are lost and which are new. Nonzero

107   availability means that the listings are available to rent for at least 1 day.

108   The new listings are visibly sparser in the plot for both rich and not rich postcodes, indicating

109   that more listings were removed than added during the pandemic for both categories.

110   **2. Nonzero availability total listings by wealth**

111   TABLE I. Summary statistics of mean nonzero availability in days pre and Post-Pandemic, by

112   listings in rich versus not rich postal codes.

	Rich Postal Code	Not Rich Postal Code
Pre-Pandemic	182.6621	165.5433
Post-Pandemic	226.9327	211.0596

113

114   Our summary statistics show that, both before and after the pandemic, the average non-zero

115   availability in rich postal codes is greater than that in non-rich postal codes. And, contrary to what

116   might have been expected, the average number of non-zero listings increases both after the

117   pandemic. This may be indicative of more listings being posted but going unoccupied.
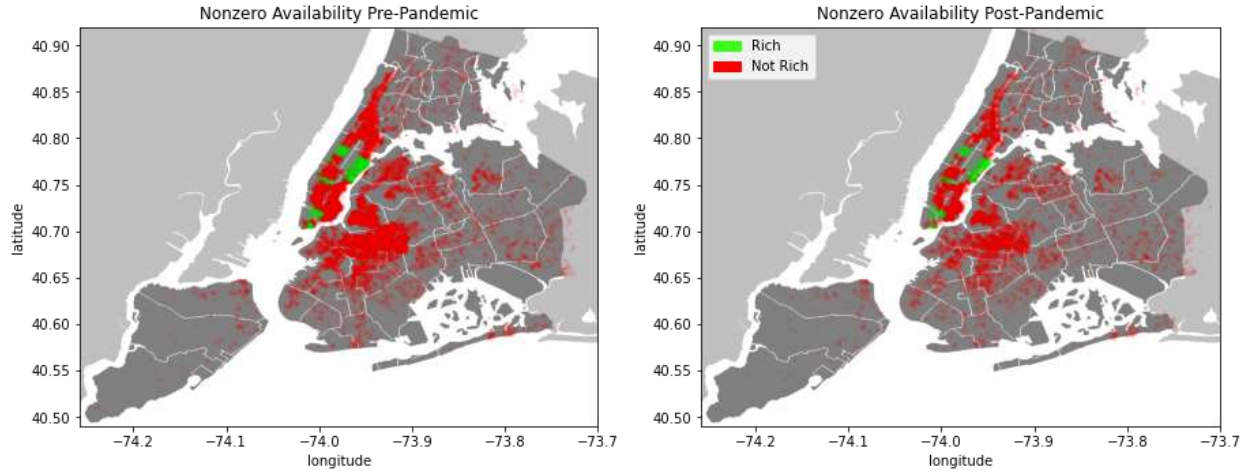
FIG. 2. A map of total nonzero availability before and during the pandemic. As in the previous plot, the post listings are much fewer in all locations, with the starkest difference visible in Manhattan and Brooklyn.
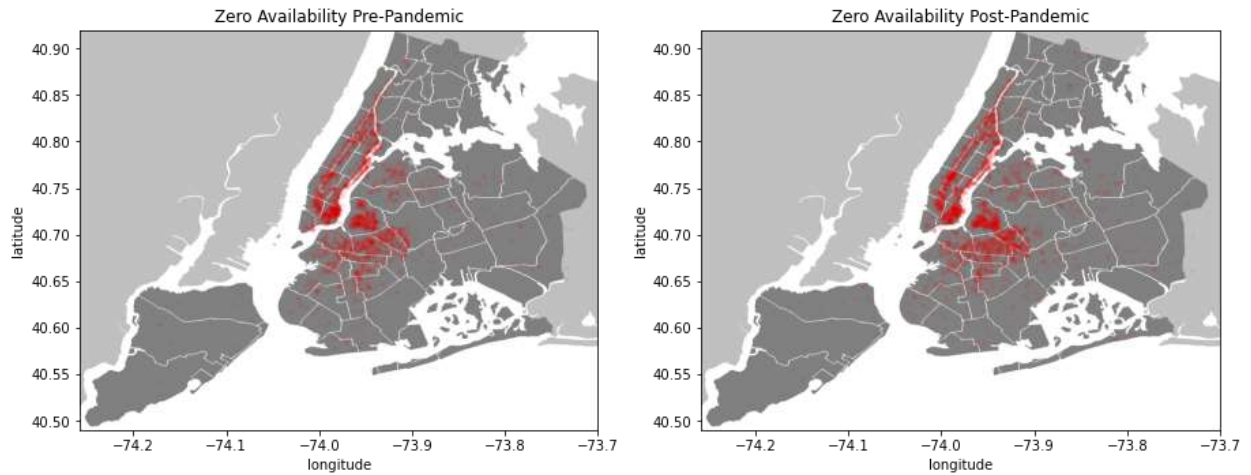
3. Zero Availability

FIG. 3. We now plot the zero availability listings, which have availability set to zero because they are either occupied or disabled. Since we are only considering listings that had been reviewed in the last year at the time they were sampled, we suppose that the majority of these are occupied rather than inactive.

B. Host Listing Counts Before and After the Pandemic

1. Host listings counts by number of listings

130   TABLE II. Summary statistics of mean days of nonzero availability pre- and post-pandemic, by

131   listings in rich versus not rich postal codes.

	< 5 Listings Per Host	Between 5 & 20 Listings Per Host	> 20 Listings Per Host
Pre-Pandemic	1.2822	1.4164	59.6944
Post-Pandemic	1.3267	1.5182	55.8864

132

133   This table shows that the mean listings per host increases for hosts with fewer than 20 listings

134   but decreases for hosts with more than 20 listings. This difference is not only mitigated, but starkly

135   reversed when one considers the fact that hosts with more than 20 listings to their ID have over 50

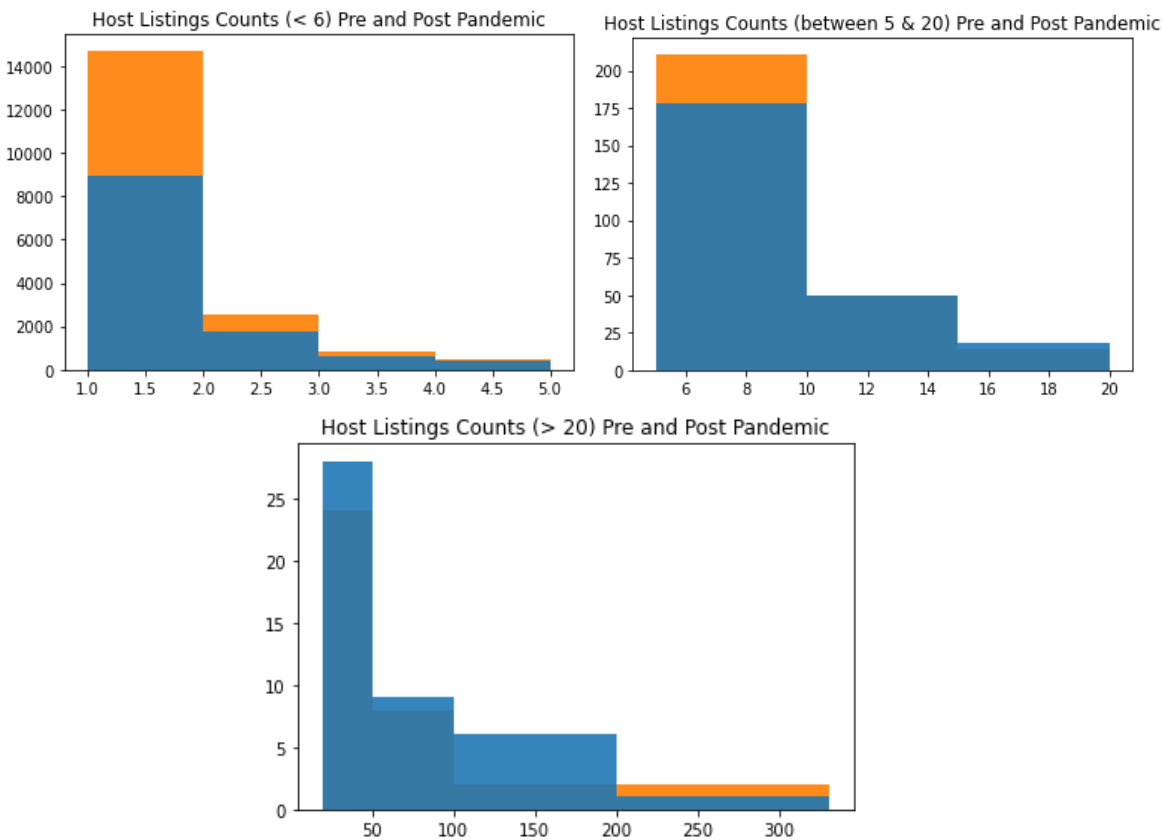136   listings on average, where hosts with fewer than 20 have at most two on average.

137

138

139 FIG. 4. This figure set of nonzero host listings counts gives more nuance to the contents of Table

140 II, which implies that the mean listings for hosts with less than 20 listings increased overall, while

141 the opposite is true for hosts with greater than 20 listings. Here, we see that, for host with less than

142 10 listings or greater than 200 listings, the orange pre-pandemic values are greater than the blue

143 post-pandemic values.

144 ### 2. *Host listings counts by wealth of postal code*

145 TABLE III. Mean listings per host, before and after the pandemic, separated by listings in rich

146 versus not rich postal codes.

	Not Rich Postal Code	Rich Postal Code
Pre-Pandemic	1.6553	1.5192
Post-Pandemic	2.2323	1.6864

147

148   When separating by postal code rather than host listings count, the difference between the

149 categories of listings is far less apparent. Still, rich postal codes have more average listings per host;

150 and again, the average listings per host increases after the pandemic. This time, the value for rich

151 postal codes increases considerably more than that for non-rich postal codes.
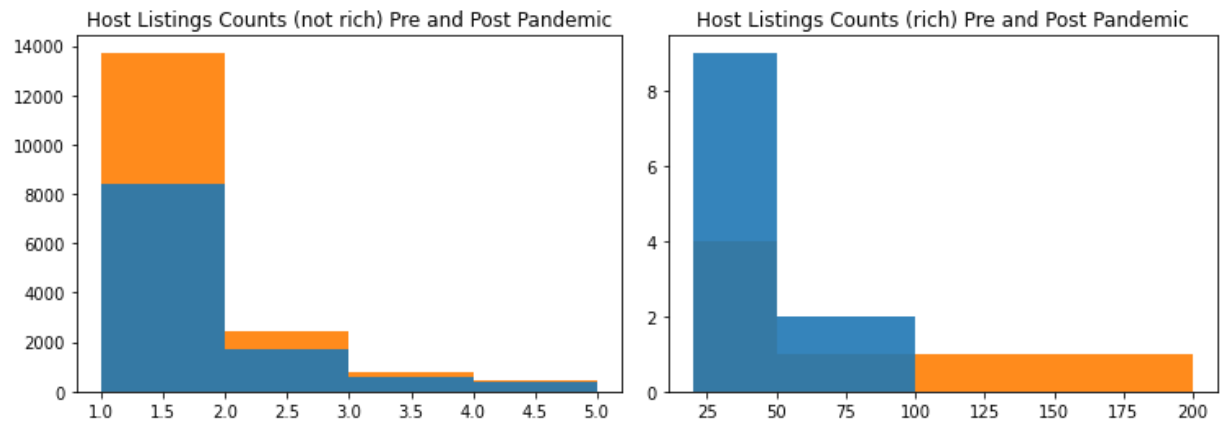

152

153   FIG. 5. This figure of host listing counts pre- and post- pandemic, by rich and not-rich postal code,

154   adds context to the contents of Table III. Before-pandemic values are again shown in orange. We

155   see that, for non-rich postal codes, the pre-pandemic values are consistently higher than post-

156   pandemic ones, indicating that the mean listings per host in these areas decreased. For rich postal

157   codes, while the post-pandemic value is similarly lower for hosts with over 100 listings, it is sharply

158   higher between 25 and 100 listings. We also see that hosts in rich neighborhoods have at least 25

159   listings each, while in poor neighborhoods they do not have more than 5.

160   **C. Other Notable Data Visualizations: Price Per Night, Minimum Nights, & Room**

161   **Types Before and After the Pandemic**

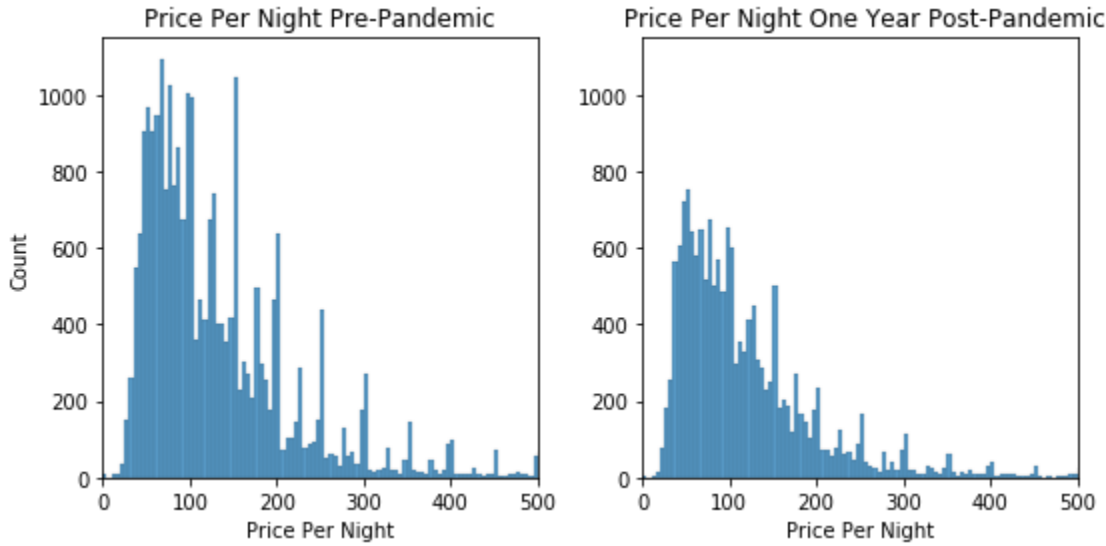162   *1. Price per night pre- and post-pandemic*

163

164   FIG. 6. This side-by-side plot of the distributions of price per night, instituted by host, before and

165   after the pandemic (one year from Nov. 2019 to Dec. 2020), visibly demonstrates the decrease in

166   prices per night one year into the pandemic. The distribution shapes are nearly the same, but the

167   highest price drops from over $1000 to just under $800, a dramatic drop of 200 dollars per night,

168   which would otherwise be unlikely to occur under normal circumstances.

169   *2. Minimum nights frequency, by room type: joinplot and heatmap*

Minimum Nights Before and After Pandemic

170
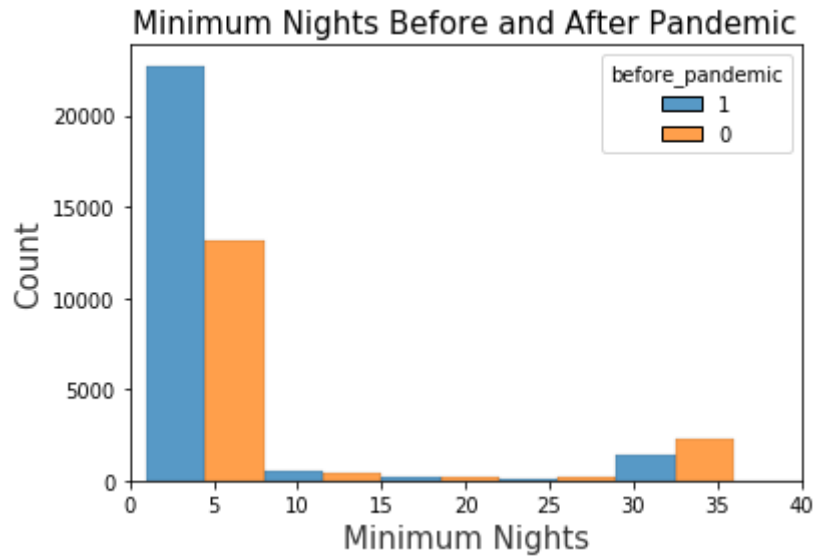
FIG. 7. The figure of minimum nights before and after the pandemic (1 indicates that

before_pandemic is true, and 0 indicates that post-pandemic is true) reveals that a small number (0-

8) of minimum nights, set in advance by hosts, is far more frequently listed before than after the

pandemic. The count in the pre-pandemic decreases by nearly 10,000 listings for the small number

of minimum nights. On the other hand, the number of minimum nights on the larger end of the

scale grew by a small amount for minimum nights greater than or equal to 30. There could be several

factors as to why this would be the case, such as less tourism due to the pandemic and thus fewer

hosts who offer listings with low number of minimum nights, and who instead offer minimum

nights over 30, or due to Airbnb policy changes, stay-at-home mandates, and more.

180
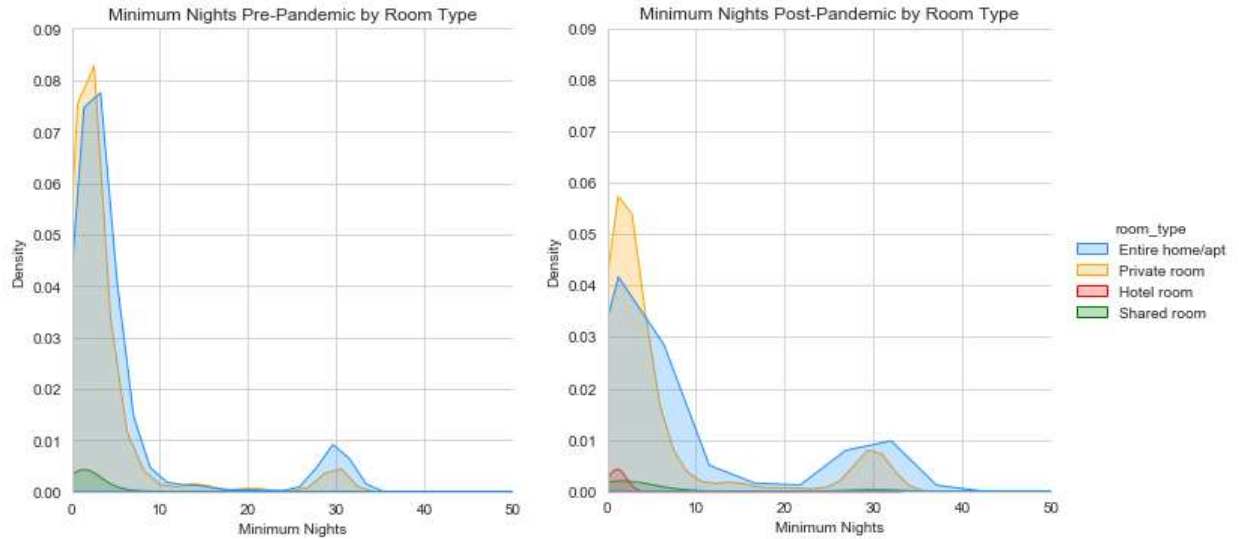
FIG. 8. This figure of a comparison between the minimum number of nights instituted by hosts,

grouped by room type, shows a similar reduction in the number of minimum nights as in FIG. 7.

For the pre-pandemic plot (left), the data extracted from Nov 2019 did not include the room type of

hotel, while for post-pandemic, it was likely significant enough to be included by Airbnb. Minimum

nights for entire homes or apartments decreased in density after the pandemic by around 0.02, and

the gap between the density of entire home or apartment and private room increased after the

pandemic. The green distribution of shared rooms flattened out and decreased, which makes sense

post-pandemic, as it is risky to share a room with someone who might have the virus. While more

hosts offered hotel rooms for a few nights, the red distribution is still not very large.
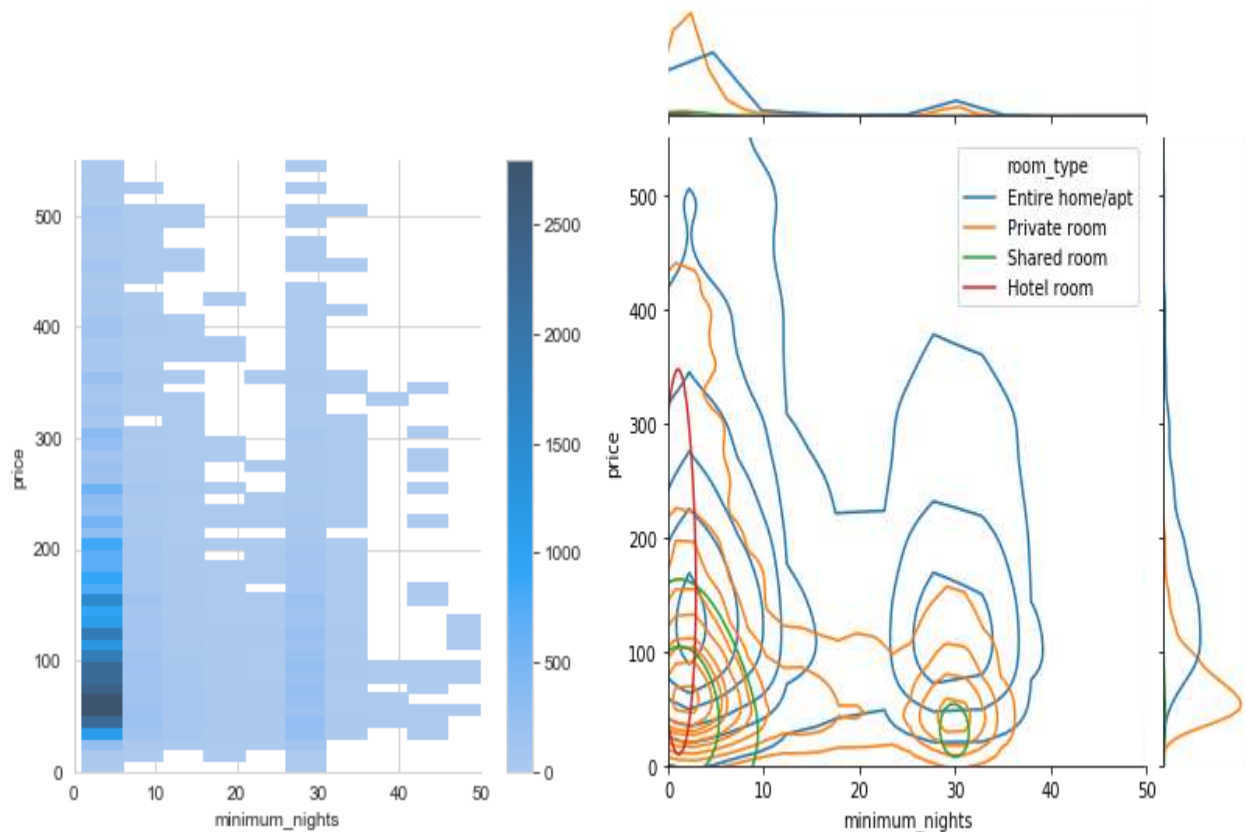
190

Fig 9. The two plots above show the relationship between minimum number of nights and price per

night and includes listings from both before and after the pandemic: the left, emphasizing frequency,

and the right, grouping by room type. The plot above and to the left is a heatmap of minimum

nights against price and reveals that the most common bookings over the last two years are in the

price range of approximately 300 to 200 dollars per night, with a minimum night stay of 5 nights.

The joinplot above, right has the marginal distribution of the minimum number of nights on top of

the graph, and the marginal distribution of price on the right side of the graph. The main plot in the

center has oval and circular shaped figures that grow larger as the price increases. These oval and

circular shapes indicate clusters of points that grow larger, more varied, and more stretched as price

increases and the number of points in the area gets sparser. For example, private room (orange) is

clustered at minimum nights near 0 and 30, with tighter ovals where there are more data values and

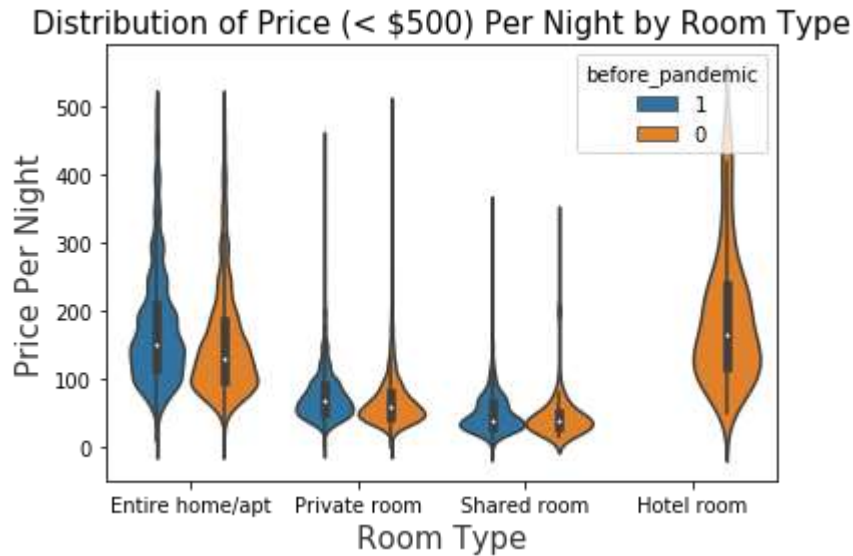less ovals and more white space between the orange curves as price grows.

Distribution of Price (< $500) Per Night by Room Type

203

204 FIG. 10. This grouped violin plot of price less than 500 dollars per night for the room types before

205 and after the pandemic (1, or blue color, indicates pre-pandemic prices; 2, or orange, indicates post-

206 pandemic prices) shows the trend that the post-pandemic violins are smoother than their blue

207 counterparts, and are slightly wider towards the end. The smoothness means that there are fewer

208 minor clusters of the listings in a certain price range, and the wideness of the violin indicates that

209 more listings are in the corresponding price range. This means that post-pandemic, the listings are a

210 bit more evenly spread out than pre-pandemic, and that the price ranges set by hosts are more

211 universal in New York City, likely due to the economic downturn: people being less likely to spend

212 more than a reasonable amount per night. Private rooms have a slightly higher price per night than

213 shared rooms; and hotel rooms, which are only available post-pandemic, have very high variability.

214

215 **V.   INFERENTIAL ANALYSIS**

216 **A. Before and After Pandemic: Mean Listings per Host**

217 *1.   Before and after pandemic: mean listings per host*

218 We would like to determine whether mean listings per host in rich neighborhoods were greater

219 after the pandemic than before.

220 *a. Hypothesis.*

221
$$H_0: F_1(x) = F_2(x)$$

222
$$H_A: F_1(x) \geq F_2(x),$$

223 where $F_1(X)$ is the distribution for group 1: rich neighborhood mean listings per host pre-

224 pandemic, and $F_2(x)$ is the distribution for group 2: rich neighborhood mean listings per host post-

225 pandemic.

226 *b.   Assumptions.* Let us test for assumptions of a parametric two sample test. From the below

227 distribution (FIG. 11.) of the rich neighborhood calculated host listings (with no duplicate host ids),

228 the data is not Normal in shape, is asymmetric, and has heavy tails. From the QQ-Plot below we see

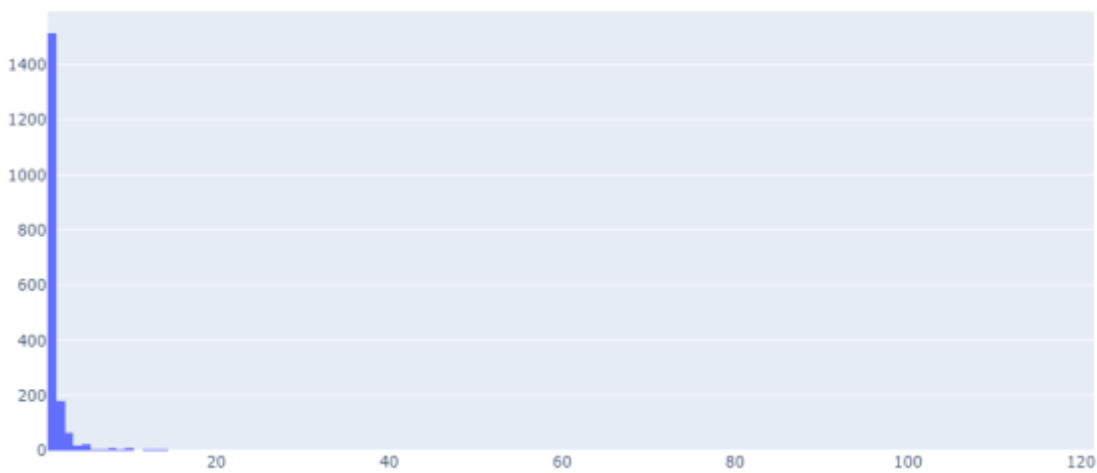229 a very heavy tail which includes several outliers that appear influential.

230

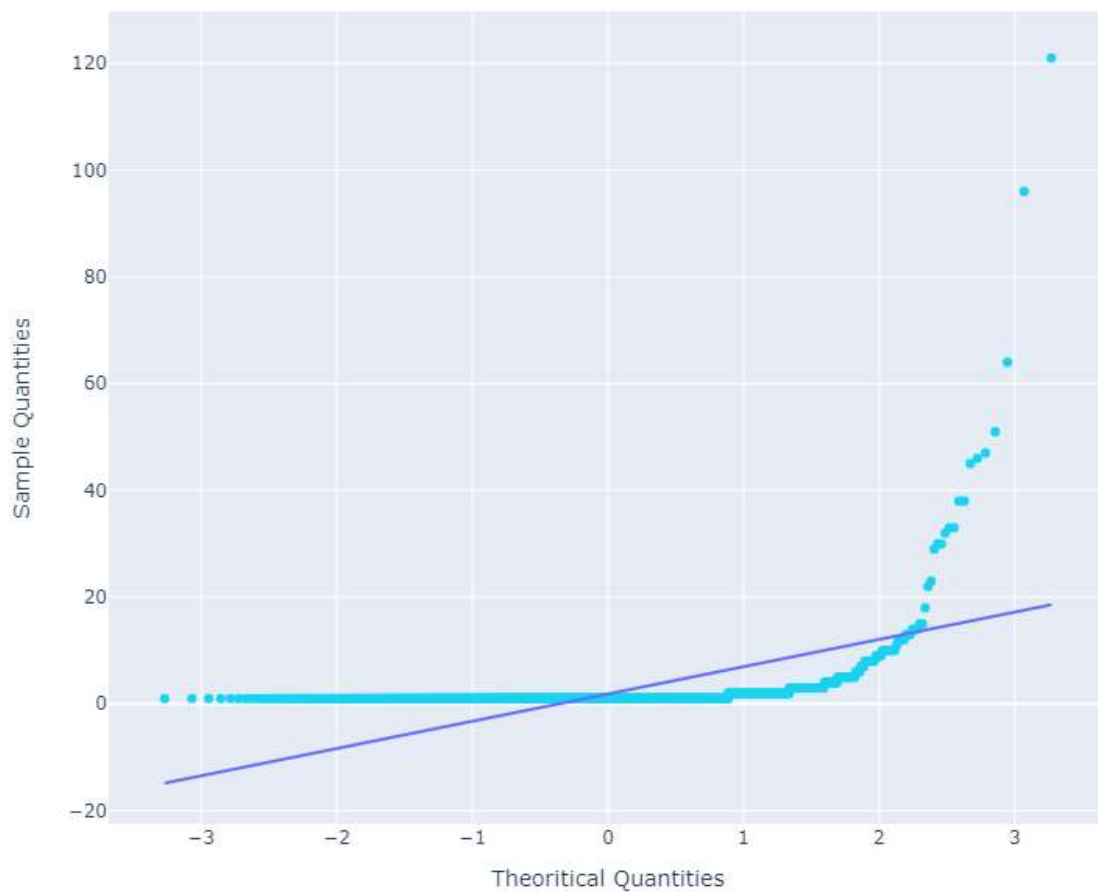231 FIG. 11.

Quantile-Quantile Plot: Rich Neighborhoods

232

233   FIG. 12.

234   Let us now test for Normality using Shapiro-Wilks Test.

Test Statistic	p-value	Comments	
0.14855217933654785	0.0	Sample does not look Gaussian (reject H0)	

235

236   From the skewed distribution of the rich neighborhood hosts data, the QQ-plot, and the

237   Shapiro-Wilks test, we conclude that our sample is not Normally distributed. Therefore, the

238   assumptions for a parametric two-sample test are violated.

239

240   *c.   Wilcoxon Rank Sum Test.* Let us now use the methodology of the nonparametric two-sample

241   test. Because our distribution exhibited heavy tails and asymmetry (skewness) as well as clear

242   outliers, we shall perform the Wilcoxon Rank Sum test for the mean calculated host listings for rich

243   neighborhoods. We use this test because ranks tend to make inferences more robust to outliers, and

244   from nonparametric methodology, the mean is a better measure of center with the distribution

245   shape we have. The Wilcoxon Rank Sum Test tends to have higher power when the distribution is

246   skewed, and outliers are present than the Permutation Test.

247   For the Wilcoxon rank sum test, we have the assumptions of independence of the samples and

248   equal variance. Because the samples were sampled randomly from the population of Airbnb listings

249   in New York City and that each sample was taken in a different year and filtered by whether the

250   listings had been reviewed in that year alone, we can assume independence. The boxplot below of

251   the samples show that both the pre- and post-pandemic data have a similar spread, so we can
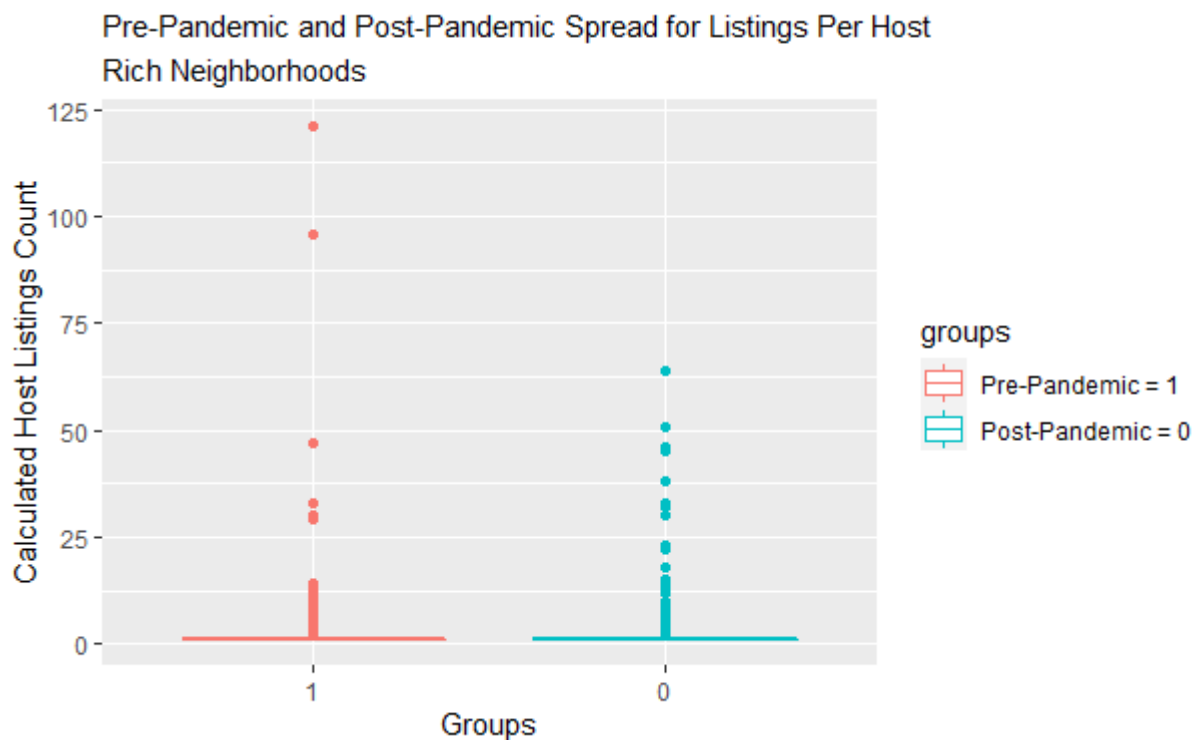
252   assume equal variance.

253

254 FIG. 13.

255 From our Wilcoxon Rank Sum test, we get a test statistic of $Z = -4.2322$ and p-value of 1.157e-

256 05. If, in reality, the distribution of mean calculated host listings for rich neighborhoods before and

257 after the pandemic were the same, we would observe our test statistic or more extreme with

258 probability of 1.157e-05.

259 **2. Poor neighborhoods**

260 We would like to determine whether poor neighborhood mean listings per host were greater

261 after the pandemic than before.

262 *a. Hypotheses.*

263 $$H_0: F_1(x) = F_2(x)$$

264 $$H_A: F_1(x) \geq F_2(x),$$

265 where $F_1(X)$ is the distribution for group 1: poor neighborhood mean listings per host pre-

266 pandemic, and $F_2(x)$ is the distribution for group 2: poor neighborhood mean listings per host post-

267 pandemic.

268 *b. Assumptions.* Let us again test for assumptions of a parametric two sample test. From the

269 below distribution (FIG. 14.) of the poor neighborhood calculated host listings (with no duplicate

270 host ids), the data is not Normal in shape, is asymmetric, and has heavy tails. From the QQ-Plot

271 below we see a very heavy tail which includes several outliers that appear influential.
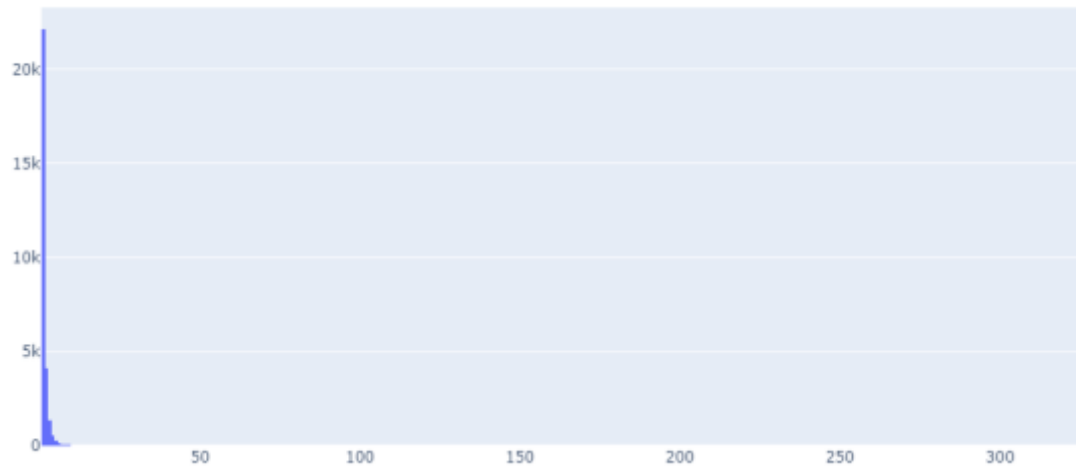
272

273   FIG. 14.

Quantile-Quantile Plot: Poor Neighborhoods
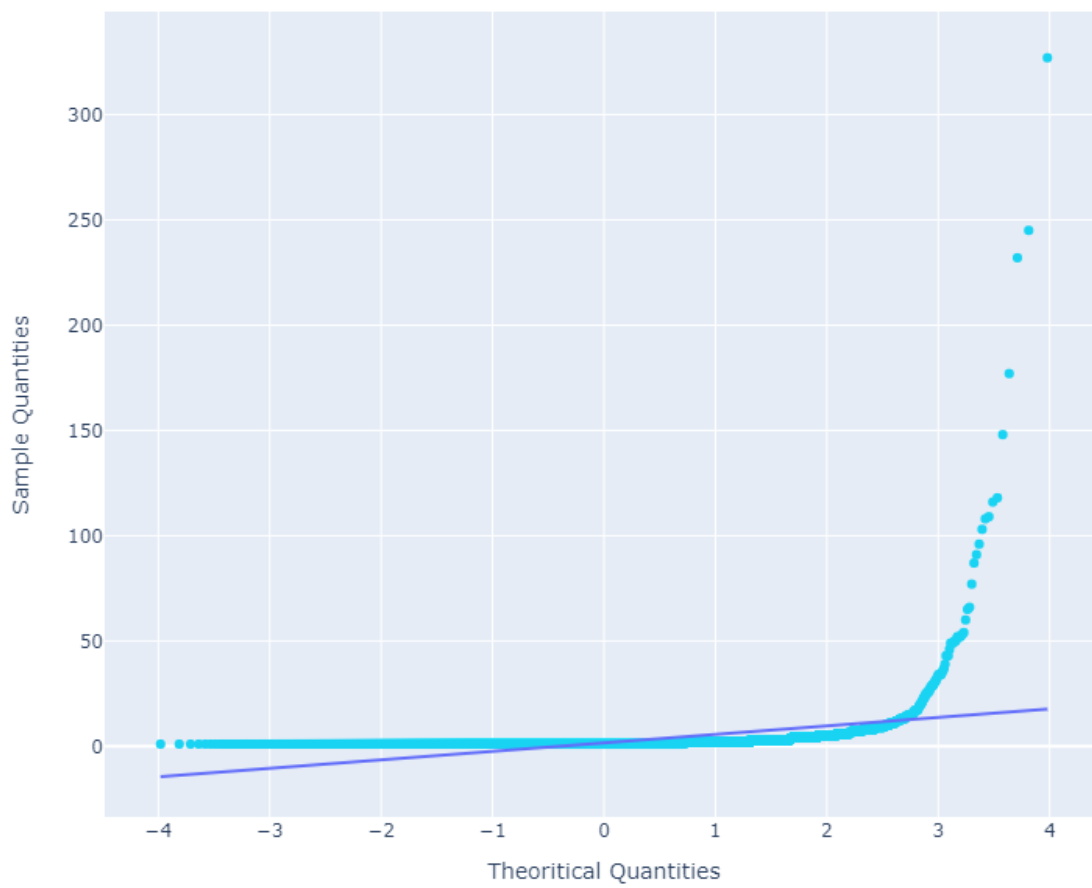
274

275   FIG. 15.

276   Let us now test for Normality using Shapiro-Wilks Test.

Test Statistic	p-value	Comments	
0.08792716264724731	0.0	Sample does not look Gaussian (reject H0)	

277

278   From the skewed distribution of the poor neighborhood hosts data, the QQ-plot, and the

279   Shapiro-Wilkes test we conclude that our sample is not Normally distributed. Therefore, the

280   assumptions for a parametric two-sample test are violated.

281   *c.*   *Wilcoxon Rank Sum Test.* Let us now use once again the methodology of the nonparametric

282   two-sample test. For similar reasons as for the Rich data, we shall perform the Wilcoxon Rank Sum

283   test for the mean calculated host listings for poor neighborhoods. The data is again independent,

284   and our boxplot below shows that pre- and post-pandemic have a similar spread, so we can assume
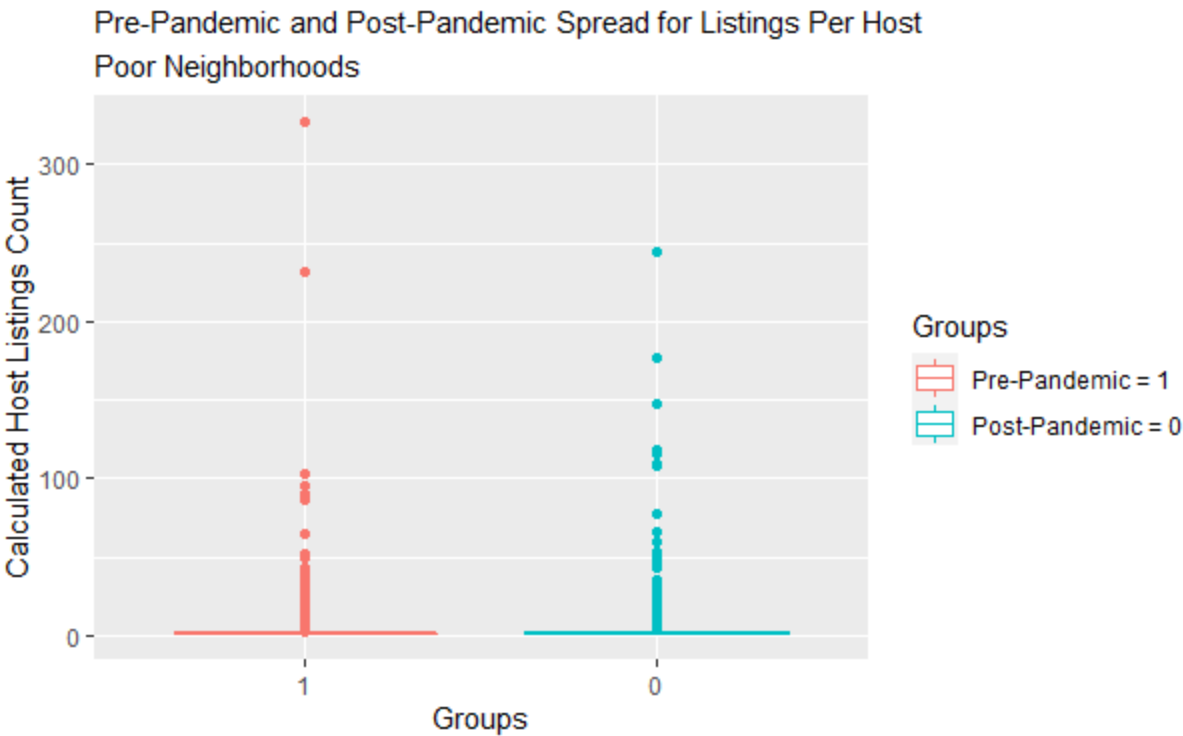
285   equal variance.

286

287   FIG. 16.

288   From our Wilcoxon Rank Sum test, we get a test statistic of Z = -6.565 and a p-value of 2.602e-

289   11. If, in reality, the distribution of mean calculated host listings for poor neighborhoods before and

290   after the pandemic were the same, we would observe our test statistic or more extreme with

291   probability 2.602e-11.

292   Because the p-value = $2.602e - 11 \leq \alpha = 0.05$, we reject the H_0. Therefore, at the 5%

293   level, we conclude that the mean listings per host in poor neighborhoods post pandemic is greater

294   than the mean listings per host in poor neighborhoods pre pandemic.

295   **B.   Before and After Pandemic: Mean Days of Nonzero Availability**

296   *1.   Rich neighborhoods*

297   We would like to determine whether mean days of nonzero availability in rich neighborhoods

298   have changed after the pandemic compared to before. We think this claim may be true because after

299   the pandemic hit in December 2019, Airbnb increased the minimum number of nights people

300   should stay at the same place, in order to "[crack] down on the availability of illegal short-term

301   rentals" (Shared Economy Tax, 2019) and data sharing laws, so hosts would have to offer their

302   listings for a longer period of time. It's possible, however, that hosts did not heed this policy because

303   they could only offer listings for a smaller number of days, because of fear of COVID-19

304   transmission in their homes, or because they were conducting illegal activities in short-term rental

305   Airbnbs. There could be many factors that we cannot account for. Thus, we test for a significant

306   difference in the mean days of nonzero availability for rich neighborhoods.

307   *a. Hypotheses.*

308                                 $$H_0: F_1(x) = F_2(x)$$

309                            $$H_A: F_1(x) \geq F_2(x), F_1(x) \leq F_2(x),$$

310   where $F_1(X)$ is the distribution for group 1: rich neighborhood mean days of nonzero

311   availability pre-pandemic, and $F_2(x)$ is the distribution for group 2: rich neighborhood mean days of

312   nonzero availability post-pandemic.

313   *b.*   *Assumptions of Normality.* From the below distributions of the rich neighborhood nonzero

314   availability (number of days available greater than 0), the distributions for both pre- and post-

315   pandemic, respectfully, are not Normal in shape. The first distribution, for pre-pandemic rich

316   neighborhoods, is skewed right and has two peaks. The second distribution, for post-pandemic rich

317   neighborhoods, is more skewed to the left, and has a heavy tail. The distributions do not look the

318   same.

319

320
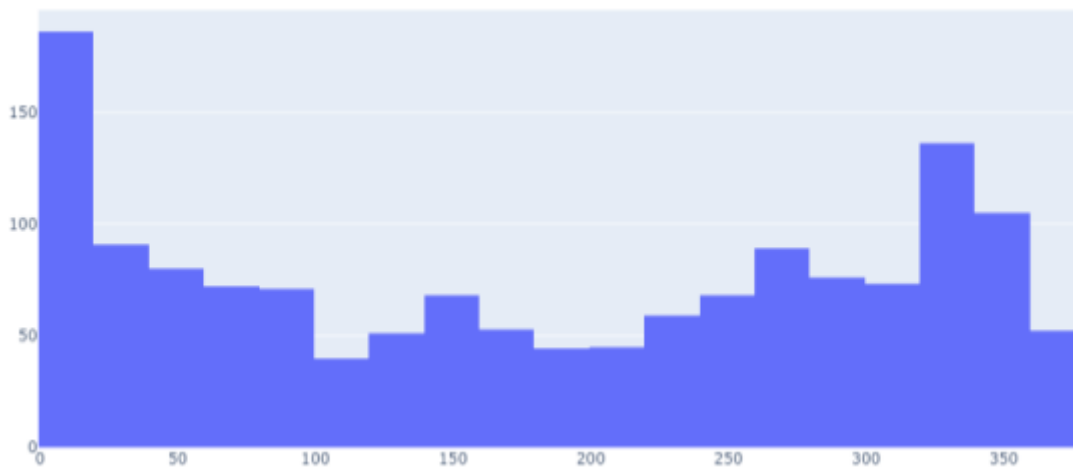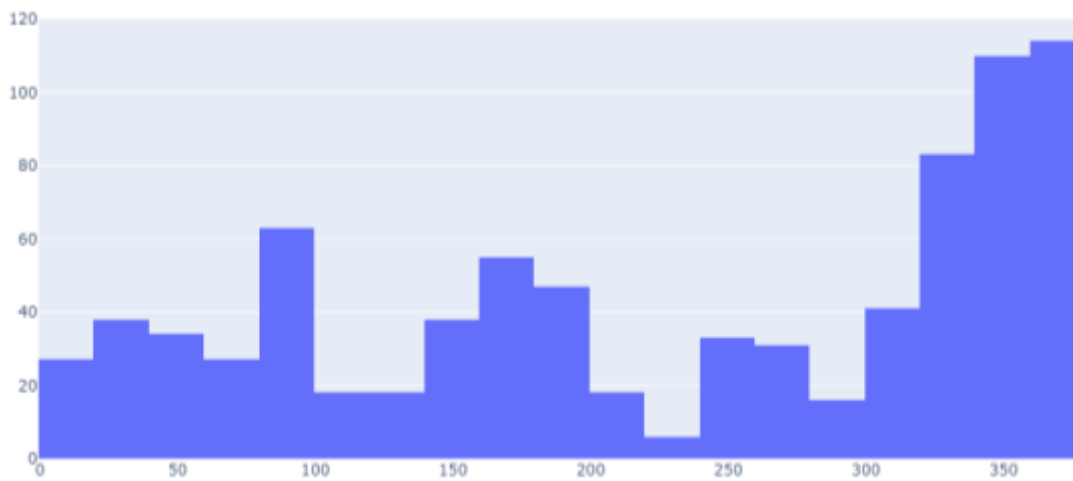
321   *c.* *Kolmogorov Smirnov Test.* Because these two distributions appear to be different, we will test to

322 see if the two samples are from the same distribution. We will use the Kolmogorov-Smirnov Test to

323 determine if there is a difference in mean days of nonzero availability in rich neighborhoods before

324 and after the pandemic. Under the assumption that the samples are from the same distribution, the

325 samples have been independently sampled in different years, with only listings with reviews in the

326 same year as sampled included in each sample; we have that the assumptions for the Kolmogorov-

327 Smirnov Test have been met.

328   From our K-S test, we get a test statistic of 0.21494 and p-value of 1.2212453270876722e-15. If

329 the distribution of mean days of nonzero availability for rich neighborhoods before and after the

330 pandemic were the same, we would observe our test statistic or more extreme with probability of

331 almost 0.

332   Because our p-value= $1.2212453270876722e-15 \leq \alpha = 0.05,$ we reject the H_0. At the

333 5% level, we conclude that there is a significant difference in mean days of nonzero availability in

334 rich neighborhoods before and after the pandemic.

335 **2. Poor Neighborhoods**

336   We would like to determine whether poor neighborhood mean days of nonzero availability have

337 changed after the pandemic compared to before the pandemic. For the same reasons as for rich

338 neighborhoods, namely the new policy by the city of New York and Airbnb that requires longer

339 stays by occupants of a listing to prevent illegal activity in short term rentals, but also other factors

340 that we cannot control for, we test for a significant difference rather than increase or decrease in the

341 mean days of nonzero availability for poor neighborhoods.

342   *a. Hypotheses*

343 $$H_0: F_1(x) = F_2(x)$$

344 $$H_A: F_1(x) \geq F_2(x), F_1(x) \leq F_2(x),$$

345   where $F_1(x)$ is the distribution for group 1: poor neighborhood mean days of nonzero

346   availability pre-pandemic, and $F_2(x)$ is the distribution for group 2: poor neighborhood mean days

347   of nonzero availability post-pandemic.

348   *b.   Assumptions of Normality*

349   From the below distributions of the poor neighborhood nonzero availability (number of days

350   available greater than 0), the data for both pre- and post-pandemic, respectively, are not Normal in

351   shape. The first distribution, for pre-pandemic poor neighborhoods, is skewed right, and has a heavy

352   tail with a relatively large peak. The second distribution, for post-pandemic poor neighborhoods, is

353   skewed strongly to the left, and has a heavy tail with two much smaller peaks. The assumption of

354   Normality has not been met, so we cannot use the parametric two sample test. The distributions do

355   not look the same.

356

357
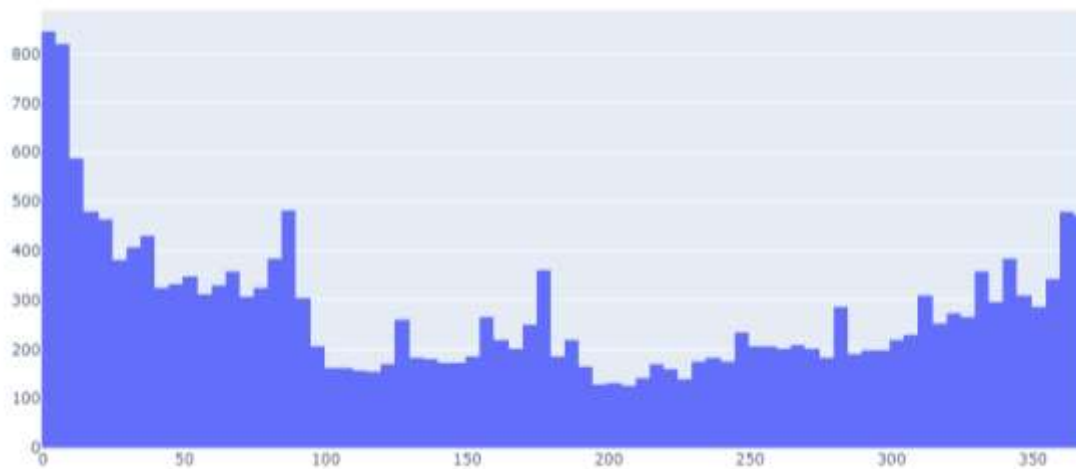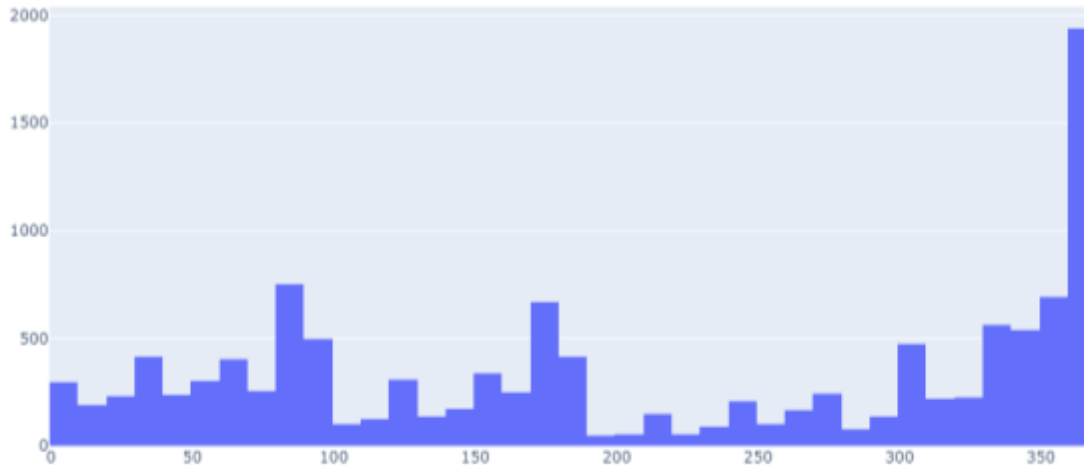
358   *c.*   *Kolmogorov Smirnov Test.* Because these two distributions for poor neighborhoods also appear

359 to be different like the rich neighborhoods, we will test to see if the two samples are from the same

360 distribution. We again use the Kolmogorov-Smirnov Test, this time to find if there is a difference in

361 mean days of nonzero availability in poor neighborhoods before and after the pandemic. Under the

362 same assumptions as the previous test, we have that the assumptions for the Kolmogorov-Smirnov

363 Test have been met.

364   From our K-S test, we get a test statistic of 0.17175 and p-value of 1.1866766660863536e-194. If

365 the distribution of mean days of nonzero availability before and after the pandemic for poor

366 neighborhoods were the same, we would observe our test statistic or more extreme with probability

367 of almost 0.

368   Because our p-value= $1.1866766660863536e - 194 \leq \alpha = 0.05$, we reject the H_0. At the

369 5% level, we conclude that there is a significant difference in mean days of nonzero availability in

370 poor neighborhoods before and after the pandemic.

371 **VI.**   **CONCLUSION**

372   The goal of this paper was to determine some of the myriad ways in which the COVID-19

373 pandemic has significantly affected the Airbnb market in New York City by comparing data from

November 2019 to December 2020, a timespan of approximately one year. During this relatively short of time, one of the major events in our lifetimes occurred, upending the economy, public health and safety, social activities, and every part of people's daily lives. One would be hard-pressed to find even a single person who has not been impacted by the pandemic, so for the largest city in the US and arguably the most famous city in the world, the results have been transformative.

In a small sector of the temporary housing market is the Airbnb marketplace, which has suffered the loss of tourists to New York City, real estate price devaluation, and hosts moving to warmer and less infected locales like Miami or the Midwest. Thus, the main customers who would rent Airbnbs in NYC have been greatly depreciated, and any influx from those who would ordinarily rent an Airbnb from out-of-town, or country would hesitate to visit such a famous location in the circumstances of the pandemic.

Earlier, we questioned if this would mean that hosts would charge different prices due to the economic recession, or if they would offer less listings or change the minimum number of nights to stay. Through data visualizations and summary statistics we discovered that in fact, the price per night hosts charged fell dramatically after the pandemic, which is likely due to the economic downturn and hosts offering cheaper prices to get back tourists and those who are unconcerned about the pandemic. Additionally, the number of lost listings was higher than the ones gained after the pandemic, and for hosts who had between 20 and 200 listings, the average listings per host increased.

Another interesting find that we developed from data visualizations and confirmed through hypothesis testing was that after the pandemic, both poor and rich neighborhoods had average mean listings per host that were higher than before, and that the average non-zero availability in rich postal codes is greater than that in non-rich postal codes. It seems likely as a result that only hosts who could survive medical (through expensive treatments, healthcare, and access to good doctors) and

398   financial hardships could continue to profit from Airbnb rentals, which would become more

399   gentrified.

400   We deduced that a possible reason for why both rich and poor neighborhoods had growth in the

401   mean listings per host, which means that hosts listed more properties rather than less one year after

402   the start of the pandemic, is that despite having listings in either or both poor and rich

403   neighborhoods, the hosts themselves are almost certainly not poor. To be able to offer even one

404   listing post pandemic may be expensive enough to maintain, let alone more than one. Many hosts

405   may even have listings in both poor and rich neighborhoods and may have profited from low real

406   estate prices in New York City due to people moving out of the city after the pandemic hit. This

407   also supports the referenced study on the effects of the pandemic on the gig economy, which

408   concluded that the pandemic would cause businesses that run on the gig or platform economy to

409   flourish.

410   Our four hypothesis tests, using the nonparametric techniques of Wilcoxon Rank Sum and

411   Kolmogorov-Smirnov Test, sought to determine whether rich neighborhoods had significant

412   differences in mean listings per host and mean number of nonzero available days pre and post

413   pandemic, and whether poor neighborhoods had significant differences in mean listings per host

414   and mean number of nonzero available days pre and post pandemic. We found that the mean

415   listings per host post-pandemic for both rich and poor neighborhoods were significantly larger than

416   before the pandemic, and this is supported by our summary statistics for these categories. We also

417   found that there was a significant difference for both rich and poor neighborhoods pre and post

418   pandemic, but we would need more hypothesis testing and a rigorous method to account for the

419   various factors associated with the new policy by Airbnb and the fallout from the coronavirus

420   pandemic.

421   The results of our study have far reaching implications on the impact of the pandemic for

422   Airbnb rentals in New York City as well as other gig economies, as it is likely that the pandemic will

423   continue to influence our daily lives, two, five, and even ten years into the future. Understanding

424   how hosts, who determine how customers will interact with the market in rich and poor

425   neighborhoods, will set prices, minimum nights to stay, number of days available out of 365, and

426   other factors, in conjunction with city and Airbnb policy will allow us to gain a better sense of what

427   may occur in the future under similar circumstances. Having this microcosm of a study would allow

428   for extrapolation into larger markets than New York, such as Tokyo, Delhi, or Shanghai Airbnb or

429   other gig marketplaces.

430

431

432   **REFERENCES**

433   Dgomonov. (2019, August 12). "New York City Airbnb Open Data," Kaggle. Retrieved April 05,

434   2021, from https://www.kaggle.com/dgomonov/new-york-city-airbnb-open-data

435   Hoisington, A. (2020, March 17). "5 insights about how the COVID-19 pandemic will affect

436   hotels," Questex. Retrieved May 01, 2021, from

437   https://www.hotelmanagement.net/own/roundup-5-insights-about-how-covid-19-

438   pandemic-will-affect-hotels

439   Garrido-Moreno, A., Garcia-Morales, V., Martin-Rojas, R. (2021, April 1). "Going beyond the curve:

440   Strategic measures to recover hotel activity in times of COVID-19," ScienceDirect.

441   Retrieved May 01, 2021, from

442   https://www.sciencedirect.com/science/article/pii/S0278431921000712#sec0080

443   Jeon, S. and Ostrovsky, Y. (2020, May 20). "The impact of COVID-19 on the gig economy: Short-

444   and long-term concerns," Statistics Canada. Retrieved May 01, 2021, from https://epe.lac-

445   bac.gc.ca/100/201/301/weekly_acquisitions_list-ef/2020/20-

446   21/publications.gc.ca/collections/collection_2020/statcan/45-28/CS45-28-1-2020-19-

447   eng.pdf

448   Jiang, Y. and Wen, J. (2020, June 23). "Effects of COVID-19 on hotel marketing and management: a

449   perspective article," International Journal of Contemporary Hospitality Management.

450   Retrieved May 01, 2021, from

451   https://www.emerald.com/insight/content/doi/10.1108/IJCHM-03-2020-

452   0237/full/html#sec005

453   (2018). "Map of 25 Wealthiest Zip Codes in New York Based on Average Adjusted Gross Income,"

454   IRS. Retrieved April 05, 2020, from

455   https://www.zipdatamaps.com/economics/income/agi/state/wealthiest-zipcodes-in-new-

456   york

457   (2019). "How to Navigate the NYC Airbnb Law," Shared Economy Tax. Retrieved May 03, 2021,

458   from https://sharedeconomycpa.com/blog/nyc-airbnb-law/

459   Kelleher, S. R. (2021). "New Airbnb Hosts Have Earned $1 Billion Since The Pandemic Began,"

460   Forbes Magazine. Retrieved May 03, 2021, from

461   https://www.forbes.com/sites/suzannerowankelleher/2021/02/18/new-airbnb-hosts-have-

462   earned-1-billion-since-the-pandemic-began/?sh=507b49565bd4

463