Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

Elsevier hereby grants permission to make all its COVID-19-related research that is available on the COVID-19 resource centre - including this research content - immediately available in PubMed Central and other publicly funded repositories, such as the WHO COVID database with rights for unrestricted research re-use and analyses in any form or by any means with acknowledgement of the original source. These permissions are granted for free by Elsevier for as long as the COVID-19 resource centre remains active.
Impact of the COVID-19 pandemic: Insights from vacation rentals in twelve mega cities

Shuai Liang \textsuperscript{a,b,c}, Hong Leng \textsuperscript{a,b,**}, Qing Yuan \textsuperscript{a,b}, Chao Yuan \textsuperscript{c,*}

\textsuperscript{a} School of Architecture, Harbin Institute of Technology, Harbin 150000, China
\textsuperscript{b} Key Laboratory of Cold Region Urban and Rural Human Settlement Environment Science and Technology, Ministry of Industry and Information Technology, Harbin 150000, China
\textsuperscript{c} School of Design and Environment, National University of Singapore, Singapore 117566, Singapore

A R T I C L E   I N F O
Keywords:
COVID-19 pandemic
Airbnb
Vacation rentals
Tourism
Risk perception
Megacities

A B S T R A C T
Coronavirus disease 2019 (COVID-19) is a challenging global problem. COVID-19 has caused shocks to various urban systems, and the tourism industry is no exception. We analyzed the impact on vacation rentals by conducting diachronic data mining on nearly 10 GB of rental data (calendar, listings, and reviews) in twelve highly internationalized megacities distributed across Asia, Europe, America, and Oceania based on the data set from the Inside Airbnb website. All twelve cities were adversely affected. The specific time of the impact is related to the pandemic’s outbreak and enforced lockdowns policies. Affected by the epidemic, reservation rates decreased, tourists preferred renting in suburbs instead of city centers, the proportion of foreign tourists in all destinations dropped sharply, tourist sentiment scores fluctuated dramatically especially among foreigners, and people focused less on tourism related activities. This study reveals the changing illustrations of vacation rentals in highly internationalized megacities under the pandemic’s influence. It offers a methodological assessment framework to monitor the hospitality sector over time and aims to serve as a reference for preparedness in similar cities worldwide.

1. Introduction

On March 11, 2020, the World Health Organization (WHO) declared the novel coronavirus (COVID-19) as an unprecedented global pandemic (World Health Organization, 2020). Asia, Europe, and the United States have successively become the epicenters of the pandemic. As of August 30, 2020, infections were confirmed in more than 190 countries, territories, and areas, totaling nearly 25 million cases with more than 800,000 deaths have been reported, and suspected cases have been growing exponentially worldwide (World Health Organization, 2020a). Similar to the severe acute respiratory syndrome (SARS) that emerged in 2003, COVID-19 is an airborne illness transmittable in humans (Chan et al., 2020; Vellingiri et al., 2020). To prevent the transmission, health communication strategies and preventive measures swept the globe (Huang et al., 2021; Thu, Ngoc, Hai & Tuan, 2020). Countries imposed restrictions on human mobility internationally. Borders were closed, cruise vessels were docked, air fleets were suspended, and hotels, restaurants, and touristic sites were shutdown. This halted global communications and exchanges, and indeed, the multidimensional and interconnected impacts at the global scale challenged current systems and led to a worldwide recession (Chakraborty & Maity, 2020 Sharifi & Khavarian-Garmsir, 2020).

The COVID-19 is actually a combination of a natural disaster, a sociopolitical crisis, and economic crisis (Zenk & Kock, 2020), constituting a blend of several disasters and crisis typologies (Ritchie & Jiang, 2019). The complexity and interconnectedness existing in the process are in line with the law of chaos theory proposed by E. N. Lorenz that suggests that a small change in one part might trigger greatly different scenarios in other parts (Gleick, 2011). Researchers have already proven the impact of COVID-19 in many areas, including pollutant emissions (Kumar et al., 2020), air quality (Latif, Dominick, Hawari, Mohtar & Othman, 2021 Wang & Li, 2021), noise (Rumpler, Venkataraman & Göransson, 2020), human mobility (Beria & Lunkar, 2021Shakibaie, de Jong, Alpkökin & Rashidi, 2021; Zhang, Qian & Hu, 2020, the global supply chain(Guan et al., 2020) and energy consumption(Jiang, Fan & Klemes, 2021; López Prol & O, 2020). As one of the most important global employers (1 in 10 jobs are directly related to tourism, UNWTO, 2020, 2020b) and the major GDP contributor for several countries,

* Corresponding authors.
E-mail addresses: s.liang@hit.edu.cn (S. Liang), hitlaura@126.com (H. Leng), hityuanqing@hit.edu.cn (Q. Yuan), akiyuan@nus.edu.sg (C. Yuan).

https://doi.org/10.1016/j.scs.2021.103121
Received 7 March 2021; Received in revised form 3 June 2021; Accepted 22 June 2021
Available online 13 July 2021
2210-6707/© 2021 Elsevier Ltd. All rights reserved.
tourism is highly vulnerable to risks and disasters internationally (Farzaneghan, Ghollipour, Feizi, Nunkoo & Andargoli, 2020) and presumably was affected in this pandemic. Tourism is used to and has become resilient in bouncing back from various crises and outbreaks (Becken & Santana-gallego, 2020; Blake & Sinclair, 2003; Novelli, Gussing Burgess, Jones & Ritchie, 2018). However, we still lack an understanding of how disasters and crises with multiple topologies affect tourism. Depicting the scenarios of tourism affected by COVID-19 could help us further understand the impact mechanism of such complex disasters and crises.

To date, we have accumulated rich experience regarding the impacts of infectious diseases on tourism. Related studies on the impacts of disease-induced epidemics, including SARS (Zeng, Carter & De Lacy, 2005), the bird flu (Rittichainuwat & Chakraborty, 2009), Ebola (Cahyanto, Wiblishauser, Pennington-Gray & Schroeder, 2016; Novelli et al., 2018), and influenza pandemics (Page, Yeoman, Munro, Connell & Walker, 2006), show comparable patterns in which they were all affected in local areas within a limited scope. However, none had implications for the global system as the COVID-19 pandemic did (Gössling, Scott & Hall, 2020). International tourist arrivals were estimated to have dropped by 78%, and 120 million direct tourism job cuts represented seven times the impact of September 11 (the largest decline in the history) (UNWTO, 2020b; UNWTO, 2020a). The coronavirus exposed the vulnerabilities of tourists’ destinations suffering from high infection rates, especially densely populated megacities that are highly exposed the vulnerabilities of tourists.

Considering all the aspects, this paper attempts to address these questions. First, does COVID – 19 have an impact on vacation rentals in highly internationalized megacities? Second, if so, when does the pandemic affect the vacation rentals? Last, in what ways does it have an impact, and to what extent? Revealing the impact of COVID-19 on vacation rentals in highly internationalized cities on a global scale can provide references for other cities globally, help us understand the potential impact of the pandemic better, and is of great significance for formulating appropriate measures in the future. And a methodological assessment framework was proposed to monitor the hospitality sector through longitudinal comparative analysis. We adopted statistical

Fig. 1. The workflow of the methodology.
analysis and textual analysis to explore the aspects that may be affected by the pandemic and the corresponding degree of impact (Q3). In addition, we checked and compared the affected time according to the policy schedule implemented by each city (Q2). Finally, we could respond to the impacts that COVID-19 imposed on tourism (Q1).

2.1. Sample selection

When selecting sample cities, we comprehensively considered the severity of the pandemic, the spatial distribution of the case cities, the tourist attractiveness, the rental market activity, and the accessibility of rental data. A total of 12 typical megacities were selected (Fig. 2). These cities were all seriously affected by the pandemic despite the outbreak times of the cities being different (using confirmed COVID-19 cases reported by the Center for Systems Science and Engineering (CSSE) at John Hopkins University) (Appendix A). The basic information of the twelve megacities can be seen in Table 1. Note that the cities we selected are widely distributed across the world and have varying populations. With data from Euromonitor International’s 2019 travel report, these sample cities all ranked in top 100 city destinations with large amount of annual international tourists arrivals. Apart from being important tourist attractions, they are also highly internationalized political, economic, and cultural megacities with global influence. Moreover, most selected cities owned active rental markets with available rental data (McCarthy, 2016). Though the numbers of listings in Beijing (China), Singapore (Singapore), and Vancouver (Canada) were relatively small, we still brought them into our analysis considering the representativeness of the seriously affected cities, the tourist attractiveness and the balance of the geographical distribution. To control the spread of the virus, countries and regions have adopted various policy measures, including border blockades and city closures. We sorted out the corresponding policies that may curb human mobility and the implementation periods by referring to the large data set released by the CoronaNet Research Project (Cheng, Barcelo & Kubinec, 2020) (Table 1) (Appendix B).

Airbnb® is the world’s largest online rental platform. It has more than three-quarters of a billion guest arrivals to date and is accessible in 62 languages across 220 countries and regions (Airbnb, 2020). Airbnb’s property and room listings have already rivaled the world’s largest hotel firms Byers, Proserpio & Zervas, 2013; Guttentag, 2015b Volgger, Taplin & Pforr, 2019. The Airbnb site is publicly available. Rental data were

Fig. 2. The distribution of the twelve studied megacities.

### Table 1

| Country                  | Continent | International tourist arrivals (million) | Population (million) | Airbnb listing number | Policy adopted | Main Policy type | Time span   |
|--------------------------|-----------|----------------------------------------|----------------------|-----------------------|---------------|-----------------|-------------|
| Barcelona (Spain)        | Europe    | 6.67                                   | 5. 494               | 23,000                | External border restrictions | 03.17–06.21 |
| Beijing (China)          | Asia      | 4.00                                   | 19. 618              | <20,000               | External border restrictions/ Lockdown | 01.23-03.22 |
| Berlin (Germany)         | Europe    | 5.93                                   | 3. 552               | <20,000               | External border restrictions/ Lockdown | 03.16–06.30; 03.23–04.26 |
| Copenhagen (Denmark)     | Europe    | 3.07                                   | 1. 321               | 20,000                | External border restrictions | 03.14–06.01 |
| London (United Kingdom)  | Europe    | 19.54                                  | 9. 046               | 47,000                | Lockdown      | 03.26–07.04    |
| New York City (US)       | America   | 13.60                                  | 18. 819              | 46,000                | Lockdown      | 03.20–05.15    |
| Paris (France)           | Europe    | 17.50                                  | 10. 901              | 78,000                | External border restrictions | 03.30–04.15 |
| Rio de Janeiro (Brazil)  | America   | 2.29                                   | 13. 293              | 33,000                | External border restrictions | 03.23–04.22–07.05 |
| Rome (Italy)             | Europe    | 9.97                                   | 4. 210               | 23,000                | Lockdown      | 03.10–06.03    |
| Singapore (Singapore)    | Asia      | 18.64                                  | 5. 792               | <20,000               | Lockdown      | 04.07–06.02    |
| Sydney (Australia)       | Oceania   | 4.16                                   | 4. 792               | 20,000                | External border restrictions | 03.16–06.15 |
| Vancouver (Canada)       | America   | 3.20                                   | 2. 531               | <20,000               | External border restrictions | 03.18–06.30 |
Table 2
Categories inside the data sets.

| Items   | Descriptions                                                                                     | Remarks                                                                                     |
|---------|-------------------------------------------------------------------------------------------------|--------------------------------------------------------------------------------------------|
| listings| Listings data show about a hundred attributes for each of the listings, including its location,   | Location information for listings are anonymized by Airbnb. In practice, this means the   |
|         | room type, supporting facilities, price and other aspects.                                      | location for a listing on the map, or in the data, will be within 0-450 feet (150 metres) of  |
|         |                                                                                                 | the actual address, which is within our required limitation.                                |
| calendar| Calendar data provide details about booking for the next year by listing. Main attributes used | If the listing was booked, then its available status is “False”, otherwise, the status is  |
|         | include ‘listing_id’ (discrete), ‘date’ (date) and ‘available’ (categorical).                    | “True”. What need to be mentioned is that, if the host withdraws their house and marks it  |
| reviews | Detailed reviews given by the guests are with 6 attributes. Key attributes include ‘date’ (date), | Comments are usually written in multiple languages within the comments of one city, e.g.,   |
|         | ‘listing_id’ (discrete), ‘reviewer id’ (discrete) and ‘comment’ (textual).                      | English, Spanish, French or Chinese.                                                       |

derived from the data set available at Inside Airbnb.com (“Inside Airbnb, n.d.”) built by Murray Cox. It is an independent, non-profit website that collates publicly accessible data collected from the Airbnb site (the site is not associated with or endorsed by Airbnb). The data set includes the availability calendar for the upcoming 365 days and the reviews for each listing. Data are verified, cleansed, analyzed and aggregated. No “private” information is used. The data set covers approximately one hundred cities around the world, including their recent rental data and archived data. For each city, the data set includes all listings in selective markets, all reviews under each listing, and each listing’s unique characteristics. We then interpret the data structure, and the three main categories inside are shown in Table 2.

A data volume of nearly 10 GB was collected for the sample cities for a three-year period (2018, 2019, and 2020). This includes listings data, calendar data, and updated review data. Because the cities lack data for individual months, we left this part of the data vacant. While the periods before and after the outbreak in 2020 are primary for the study, a comparative analysis with previous years ensures the elimination of seasonal fluctuations (Chang, Chen, Che & Chu, 2019).

2.3. Data mining

The data set mentioned in 2.2 was interpreted by data mining utilizing R. It was used to explore the changes in rental booking demand, the shifts in the reservations’ spatial pattern, and the variations in attitude feedback from tourists. We excluded the rows and columns containing null values.

Rental demand mining——Two time spans (December 2018-May/June 2019 and December 2019-May/June 2020) were selected for longitudinal comparison analysis to eliminate the influence of seasonal changes. It is assumed that people tend to book three months in advance. Therefore, we examined the booking rates of vacation rentals in the next three months. Here, we define “Reservation rate = Number of bookings (when the available status of the listing is “FALSE”) / Total number of listings.” We first tabled the column “date” in “calendar,” then transformed it into a data frame and determined the total number of the listings for every day. We then filtered the rows when the column of available is “FALSE” in “calendar,” then we determined the reservation rate for each day using the equation above and mapped it out. Review data are a good reflection of tourists’ experiences and preferences. We tentatively explored the relationship between the review numbers and the severity of the pandemic. As the populations of different countries and cities are different, the absolute values of confirmed cases could not properly reflect the pandemic’s severity. Thus, we used the number of confirmed cases per million people in the country where the city was located. We conducted statistical analysis on the relationship between the review numbers and the confirmed cases per million people during the pandemic period in 2020. We plotted the review numbers for the same period in 2019 to obtain a better understanding of the effects caused by the pandemic.

Review language mining——We extensively used comments (reviews) in our analysis. The data set contained reviews written in multiple languages. We assume that the language used in the tourists’ comments is their native language, which roughly indicates their origin countries. It should be noted that some countries’ official languages are

Fig. 3. Booking Rate of the twelve megacities in two time spans. For each city, the lower part represents the later time span from December 2019-May/June 2020, while the upper part represents the earlier time span from December 2018-May/June 2019. Every column along the time span shows the reservation rates for the remaining days of the month in which the data were collected and the next three months at the right month. Every grid in the picture represents the unavailable rate in each day. Red represents “high unavailable rate”, while blue represents “low unavailable rate”.

| Listing_id | Date | Available | Reviewer_id | Comment |
|------------|------|-----------|-------------|---------|
| 12345      | 2019-01-01 | True    | 1234      | “Great stay!” |
| 67890      | 2019-02-15 | False   | 5432      | “Needs improvement.” |
| 11122      | 2019-03-31 | True    | 2134      | “Excellent experience.” |

```c
import pandas as pd
import matplotlib.pyplot as plt

# Load the data
data = pd.read_csv('data.csv')

# Convert available status to True/False
data['Available'] = data['Available'].apply(lambda x: True if x == 'True' else False)

# Plot the booking rates
plt.figure(figsize=(10, 6))
for city in data['City'].unique():
    rate = data[data['City'] == city]['Reservation Rate'].values
    plt.plot(rate, label=city)
plt.legend()
plt.xlabel('Time')
plt.ylabel('Reservation Rate')
plt.title('Booking Rate Comparison')
plt.show()
```
Table 3
Grading table of booking number changes for each month compared to the counterparts in the previous year.

| City     | December | January | February | March | April | May    | June |
|----------|----------|---------|----------|-------|-------|--------|------|
| Rome     | 3.35%    | 1.51%   | 10.65%   | -22.50% | -42.79% | -42.44% | 0.00%|
| Barcelona| 0.19%    | 0.68%   | -0.72%   | -21.99% | -36.10% | -38.15% | 0.00%|
| NYC      | 13.90%   | -0.06%  | 4.77%    | -7.64% | -16.96% | -21.43% | -21.32%|
| Paris    | -3.29%   | -5.54%  | 0.95%    | -12.07% | -15.38% | -16.71% | 0.00%|
| Berlin   | -0.56%   | -1.07%  | -1.11%   | -8.89%  | -11.85% | -16.02% | 0.00%|
| Singapore| 0.00%    | 0.00%   | 0.00%    | 3.64%  | -3.62%  | -14.23% | 0.00%|
| London   | 1.33%    | -1.36%  | 2.86%    | -7.85% | 2.95%  | -13.57% | 0.00%|
| Vancouver| 2.51%    | -1.75%  | 7.77%    | -8.23% | -20.62% | -13.41% | 0.00%|
| Copenhagen| -1.93%| 3.84%   | 0.46%    | -3.12% | -7.32%  | -9.16%  | 0.00%|
| Sydney   | 0.87%    | -3.32%  | 2.12%    | -4.26% | -4.93%  | -3.07%  | 2.25%|
| Rio de Janeiro | 3.05% | 7.35% | 7.35% | 2.61% | -1.49% | 2.74%   | 0.00%|
| Beijing  | 5.80%    | 5.34%   | 68.41%   | 67.16% | 58.54% | -46.61% | 0.00%|

Fig. 4. The changes of review numbers against the severity of the pandemic situation. The gray column represents the daily review numbers in 2019, while the blue one 2020. The orange curve illustrates the pandemic’s trends with the daily accumulated confirmed cases per million in their countries. The yellow dotted lines mark the implemented start time of policies concerning external border restrictions or lockdowns.
the same, and some certain languages are used in a various number of countries. Therefore, we can first make judgments based on the changing trends of the proportion of domestic and foreign tourists in countries/regions where the mother tongue is not widely used, such as Spain (Barcelona-Spanish), China (Beijing-Chinese), Denmark (Copenhagen-Danish) and Brazil (Rio de Janeiro-Portugal), and then extend it to other cities for further analysis. We carried out language detection for the comments made after January 1, 2020, using R’s language detection function and obtained the language corresponding to each comment. By plotting the proportion of different languages in tourists’ comments for each day during the pandemic period, we could obtain the possible changes in the composition of tourists for each city. It is worth noting that in this case, the decline in the proportion of languages cannot accurately reflect the change in the proportion of tourists from the exact country of origin.

Spatial pattern mining——We added location information in “listings” to “calendar” by joining with the same column of the two items. As the distribution of bookings within a certain day may have greater contingency and randomness, we used the overall order distribution for a whole month in the three years (May 2018, May 2019, and May 2020) for analysis. The number of days that every listing was reserved (out of 31 days for each month) was calculated. We define “Occupancy rate of a listing = Number of the days the listing reserved (when the available status of the listing is ‘FALSE’)/ 31” here. Based on the location information, we draw them on the map to obtain the reservations’ spatial distribution. We used an optimized hotspot analysis tool in ArcGIS to illustrate the spatial clustering distribution of high and low reservation rates by calculating the Gi bin index. The Gi Bin field identifies statistically significant hot and cold spots, corrected for multiple testing and spatial dependencies using the false discovery rate (FDR) correction method. Features with a Gi Bin value of either +3 or −3 were statistically significant at the 99% confidence level (CI); +/-2 bins 95% CI; +/−1 bins 90% CI. Features with 0 were not statistically significant.

Textual data mining——Comments were written in multiple languages, and different algorithms and matching libraries were involved in emotion recognition for different languages. At present, multilanguage emotion mining technology is not mature. Therefore, to simplify the analysis process, we used R to filter out the comments written in English (the share of comments written in English for sample cities is attached in the Appendix C) and performed text filtering to remove common stop words and phrases that did not contribute significantly to the meaning of the comments. After that, we adopted the “sentiment_by ()” function in R for sentiment analysis. This function has the NRC Word-Emotion Association Lexicon resource employed by Mohammad and Turney (2013) built-in, in which words in English were assigned emotion ratings based on Plutnik’s (1980) wheel of emotions.
Words in the NRC Word-Emotion Association Lexicon were tagged by positive-negative polarity classes, from which we could obtain the sentiment scores for each comment. Theme focus mining could help us better understand the variety of things tourists focused on before and after the pandemic’s outbreak. The comments in February and April for each city were used to build the theme model. We then set up the corpus, cleaned up the text, established the text matrix, and calculated the word frequency.

3. Results

3.1. Changes of rental demand

The calendar map of the booking rate for sample cities in two time spans (December 2018-May/June 2019 and December 2019-May/June 2020) is shown in Fig. 3. We only analyzed the booking rate for the upcoming three months. The available status turns to be “False” when the listing was booked, and “True” when not been booked. This means that “the unavailable rate” could represent “the booking rate.” Compared with the first period, we can conclude four categories for the sample cities (Table 3).

- In the first category, most of the reservations in March, April, and May showed a significant downward trend compared with the same period in the previous year, e.g., Barcelona and Rome. For the second category, the reservation rates showed a slightly downward trend, e.g., New York City, Paris, Berlin, Singapore, London, Vancouver, and Copenhagen. These two kinds of categories are easily understood because these cities were heavily affected by the pandemic.  
- In the third category, the reservation rates remained basically unchanged, e.g., Sydney and Rio de Janeiro. When referring to the pandemic situation in Appendix A, these two cities were not seriously affected by the pandemic at the beginning of the outbreak.

Fig. 6. (a) The spatial patterns of the occupancy rate for each listing, (b) The spatial patterns of the occupancy rate for each listing Rental data for Singapore in May 2018 is missing.
Therefore, the travel plans of tourists to these cities were not heavily affected.

- The fourth category was Beijing, which is very special. An abnormal phenomenon was observed from the map; starting in February 2020, a large number of “super high unavailable rates” appeared, and the unavailable rate reached nearly 90%. It is speculated that a large number of landlords took back their houses during the pandemic period instead of renting them out and marked them as unavailable. Therefore, the listings’ available status showed that they could not be reserved, but tourists had not actually reserved the houses.

The daily number of reviews in two time spans (January 22, 2019-June 10, 2019 and January 22, 2020-June 10, 2020, with the data in some cities having not been updated as far as 2020.06.10), the pandemic situation in the corresponding period of 2020 and the time of the policies implemented for sample cities are plotted in Fig. 4. The changes in the review numbers in the first part of 2019 and 2020 for each city were synchronous. However, with the development of the pandemic, each city’s review numbers began to drop sharply at different rates. Due to the pandemic, the proportion of the review numbers compared to the actual number of occupants may also decrease. Among these cities, Beijing was the first city where the pandemic broke out, and its review numbers began to decline first. No changes occurred in other cities during the same period. After some time, the review numbers in other cities began to drop. The timing was generally consistent with the implementation of external border restrictions or lockdown policy measures.

3.2. Changes of review languages ratio

Language detection was used to identify the language used in the reviews (comments). The language proportion of each city’s reviews from January 1, 2020 to June 10, 2020 is plotted in Fig. 5 (some cities lack data, and the missing data in the figure are filled with gray). The
yellow lines mark the implemented start time of policies concerning external border restrictions or lockdowns. After implementing the policies, the proportion of mother tongues in sample cities all increased significantly. Specifically, the ratio of Spanish in Barcelona (Spain), Chinese in Beijing (China), German in Berlin (Germany), Danish in Copenhagen (Denmark), French in Paris (France), Portuguese in Rio de Janeiro (Brazil), Italian in Rome (Italy), and English in London (UK), New York City (USA), Singapore (Singapore) and Sydney (Australia) all increased. Although Vancouver is a bilingual’s country with French and English as its mother tongues (Canada Government, n.d.), the proportion of English increased. Aside from the English, German, and Italian speaking countries (more than 50 countries where English is the official language, while German is in only 5 or 6 countries, Italian in 3 or 4), the ratio of foreign languages in the other countries were all decreased sharply or virtually disappeared after the policy implementations. In English, German, and Italian speaking countries, the proportion of foreign languages decreased as well.

The different colours represent different languages. The white color
in Beijing represents no comments were made on that day. The yellow dotted lines illustrate the implemented start time of the policies.

3.3. Changes in the spatial distribution of reservations

The longitudinal analysis of each listing’s occupancy rate in May 2018, May 2019, and May 2020 of sample cities is illustrated in Fig. 6. Every listing’s occupancy rate is expressed by dividing the number of occupied days in May by 31 (the number of days in the entire month). The higher the occupancy rate is, the more popular the listing is, and vice versa. It is worth noting that the booking condition in Beijing is special from the previous calendar map. Most of the ‘reserved listings’ in Beijing were units marked by the hosts (who had actually withdrawn their houses) and thus should be excluded when calculating the occupancy rate. Therefore, we excluded the listings occupied for the full 31 days, considered them as the house sources that the host withdrew, and mapped the remaining occupied listings reserved by tourists. Hotspot analysis was adopted to detect clusters with high and low occupancy rates with statistical significance. The distribution of the reservations with different occupancy rates was obtained. Three patterns of spatial distribution shifts for the occupancy rate in May 2020 compared with those in 2018 and 2019 were observed.

- The spatial distribution patterns of the high and low occupancy rate clusters remained the same. However, the areas of high occupancy clusters decreased while the areas of low occupancy clusters increased. The shrinking areas of high occupancy rate clusters and the expansion areas of low occupancy rate clusters were mainly concentrated in densely populated urban centers, such as Berlin, Copenhagen, New York City, Rio de Janeiro, Singapore, and Sydney.
- The areas occupied by high- and low-values clusters remained unchanged. However, the spatial distribution pattern of reservations with different occupancy rates changed significantly. High-value clusters shifted from city centers to the suburbs, while low-value clusters accumulated in city centers, e.g., Barcelona, Paris, and Vancouver.
- The spatial distribution patterns of different occupancy rate clusters and their areas all changed dramatically. For example, in Beijing, London and Rome, there are large areas of low-value clusters in the central area and scattered small-scale, high-value clusters on the peripheries.

3.4. Changes in tourists sentiment and theme focused

3.4.1. Sentiment analysis

After language detection, we screened out the comments written in English for the twelve cities and adopted sentiment scores to show tourists’ emotional variations during the pandemic. After the pandemic’s outbreak, all the cities’ emotional scores fluctuated and were not as stable as before (Fig. 7). The degree of emotional fluctuation in different cities was different. The sentiment scores in cities where English is a foreign language, such as Barcelona, Beijing, Berlin, Copenhagen, Paris, Rio, and Rome, fluctuated heavily, with mean sentiment scores mostly concentrated from 0.1 to 0.5 (Fig. 8a). In London, New York City, Singapore, Sydney, and Vancouver, where English is the native language, the sentiment scores had relatively fewer emotional fluctuations, with mean sentiment scores mostly concentrated from 0.25 to 0.4 (Fig. 8b). In the former cities where English is the foreign language, most of the comments were made by foreigners; in the latter cities where English is the mother tongue, most of the comments were made by locals. Foreigners showed more obvious mood swings than locals. This might be because they are unfamiliar with the foreign environment; hence they were more fear-prone in the event of a pandemic.

3.4.2. Theme focused

We conducted topic modeling to determine the changes of things people focused on during the pandemic period. The top 100 high-frequency words in the comments of the twelve cities made by tourists in February and April were filtered out in Fig. 9. Some words related to the pandemic appeared and were highlighted in blue, such as “Lockdown” and “Cancel(l)ed”. Tourists’ attention to “Location” has declined from February to April in cities such as Berlin, Paris, New York City, and Rio de Janeiro. Tourists in some cities focused on eating out with a high frequency of “Restaurant” and later shifted to eating at home with “Kitchen” appearing, such as in Barcelona, Paris, and Vancouver. The attention to tourism in some cities dropped significantly or even disappeared, especially in Rome. The theme words in February were all related to tourist attractions and food- or tourism-related things, such as “Havana,” “Miami,” “Tour,” “Guide,” “J.J.” “Culture,” and “Delicious,” while in April, these words disappeared and turned to things about daily living, such as apartment and stay. This situation also appeared in Singapore and Rio de Janeiro, with the attention to “Tour” and “Beach” declining.

4. Discussions

The outbreak of COVID-19 has led to devastating impacts on society, the economy, and urban systems. Vacation rentals—the upcoming trend in the accommodation sector—have not escaped this disaster’s impact. We illustrate data on vacation rentals in 12 highly internationalized megacities worldwide and find impacts on vacation rentals in many aspects, including rental demands, changes in review language ratios, distribution of reservations, and tourists’ sentiments.

The direct consequence of the pandemic and travel restrictions or lockdown policies is a change in bookings. Numerous studies have
Fig. 9. Topic model of reviews in February and April for each city.
illustrated the adverse effect of other disease-induced epidemics on hotel occupancy rates in the literature that all targeted one country (Kim, Chun & Lee, 2005; Tse, So & Sin, 2006; Wu, Law & Jiang Brianda, 2010). In our study, there were different models in twelve cities in terms of rental bookings, and their difference in the timing and degree of change was mainly due to the local trend of the pandemic. The pandemic first broke out in Beijing, and its booking rates decreased first. When Rio and Sydney adopted policies such as lockdown and external border restrictions—but the pandemic did not break out locally (Center for Systems Science and Engineering (CSSE) at John Hopkins University, 2020)—their booking rates basically stayed the same. This is because tourists primarily booked in advance, and most orders cannot be canceled; thus, measures such as lockdowns did not directly reflect the decline in booking rates. However, things were different with review numbers, which could reflect the actual occupancy rate. The implementation time of the lockdown and external border restrictions policy became another key factor causing the sharp decrease in review numbers. When the pandemic spread in most countries, reviews numbers decreased significantly in Rio de Janeiro and Sydney, although the pandemic did not break out locally. This result is consistent with a study showing that lockdowns were particularly effective in reducing long-range recreational trips (Pullano, Valdano, Scarpa, Rubrichi & Colizza, 2020). At the same time, we also found that in Beijing, which imposed stricter controls on residents’ movement on February 10 (with community blockades and residents forbidden to leave their homes for nonessential circumstances), most landlords took back their houses. As a result, Beijing has seen an unusual “reservation rate.” To better interpret this phenomenon, we compiled the total tourist volume trends before and after the epidemic from the government official website of each city/country with data on overnight visitor arrivals (Table 4). From the table, we can see that most of the cities (some cities lack the data of overnight visitors in 2020) had a large reduction in overnight visitor arrivals; Beijing had the highest drop rate of overnight tourists in 2020 compared with the former year; New York City retained the largest number of absolute overnight visitor arrivals. This is consistent with the overall trend of the decline in the number of vacation rental orders and reviews.

We assume that the language used in tourists’ reviews is their native language, helping us infer their original countries. Aside from the English, German, and Italian speaking countries (more than 50 countries where English is the official language, while German is in only 5 or 6 countries, Italian in 3 or 4), the ratio of foreign languages in the other countries all decreased sharply or basically disappeared after the policy implementations. This shows that the proportion of foreign tourists dropped significantly during the pandemic. Based on this, the proportion of English, German, and Italian in English, German, and Italian speaking countries increased, which can also indicate a certain extent the decline in the proportion of foreign tourists in these countries. This is because the external border restrictions and blockade policies implemented during the pandemic prevented foreigners from entering the country. In addition, the coronavirus pandemic may have impacted the thoughts and feelings of tourists and changed the way tourists travel (Zenker & Kock, 2020). This tendency may have triggered a change in the perception of tourists’ travel behavior, thereby avoiding overcrowded long-distance travel and mass tourism destinations and preferring destinations with small domestic populations (Faulkner, Schaller, Park & Duncan, 2004), which also caused a decrease in foreign travel.

## Table 4
The overnight visitors arrivals of the sample cities in 2019 and 2020.

| City          | Overnight visitor arrivals (million) | Percentage Change: 2020 vs. 2019 |
|--------------|-------------------------------------|----------------------------------|
|              | 2019      | 2020  |                                |
| Barcelona    | 20.3      | 5.0   | -75.4%                         |
| Beijing      | 3.8       | 0.3   | -92%                           |
| Berlin       | 14.0      | 4.9   | -64.4%                         |
| Copenhagen   | 6.1       | 1.4   | -77.1%                         |
| London       |           |       |                                 |
| New York City| 66.6      | 22.3  | -67%                           |
| Paris        |           |       |                                 |
| Rio de Janeiro|         |       |                                 |
| Rome         |           |       |                                 |
| Singapore    | 19.1      | 2.7   | -87.5%                         |
| Sydney       | 16.8      | 6.4   | -61.8%                         |
| Vancouver    | 11.0      | 3.1   | -72.3%                         |

## Table C1
Share of comments written in English for sample cities.

| City          | Number of comments(N) | Number of comments in English(N_en) | Percentage(N_en/N) |
|--------------|------------------------|------------------------------------|--------------------|
|              |                        |                                    |                    |
| Barcelona    | 45,500                 | 26,118                             | 57%                |
| Beijing      | 11,451                 | 10,509                             | 8%                 |
| Berlin       | 36,655                 | 22,091                             | 60%                |
| Copenhagen   | 14,917                 | 13,397                             | 69%                |
| London       | 107,823                | 89,371                             | 83%                |
| NYC          | 81,528                 | 72,255                             | 89%                |

| City          | Number of comments(N) | Number of comments in English(N_en) | Percentage(N_en/N) |
|--------------|------------------------|------------------------------------|--------------------|
|              |                        |                                    |                    |
| Paris        | 93,470                 | 77,652                             | 83%                |
| Rio          | 44,821                 | 30,930                             | 67%                |
| Rome         | 87,907                 | 71,541                             | 82%                |
| Singapore    | 60,307                 | 55,905                             | 92%                |
| Sydney       | 17,996                 | 16,672                             | 93%                |
| Vancouver    |                        |                                    |                    |
eating out to eating at home. The tourism theme words, including tourist attractions and food or tourism-related mentions, dropped significantly, replaced by those about daily living and normal life. These changes in focus are also a manifestation of the changes in housing demands.

We believe that this sudden shock disrupted people’s normal travel life, caused strong disturbances to the rental industry, and posed a challenge to the sustainable development of cities. Only when we first understand the possible impact of this kind of shock on vacation rentals, the key time points, and the degree of influence can we better deal with the negative outcomes the pandemic aroused. At the same time, we should be aware that some of the changes brought about by this pandemic may last for a long time, and some may become the norm in the future. Border closure and lockdown policies made people’s travel more difficult than ever and their willingness to travel also weakened. As a result, the overall number of tourists declined, and the proportion of foreign tourists dropped, and domestic ones increased, making domestic tourism become more popular in these countries than before. If the epidemic cannot be effectively controlled, this situation will be difficult to reverse in the short term. Surprisingly, we noticed new trends in the spatial distribution of rentals in some cities, namely Rome and London. As many famous attractions located in the city center were forced to close, and tourists search for natural and not crowded destinations in the suburbs, a notable decentralized form appeared.

It is necessary to carry out targeted potential mining based on the city’s own tourism resource endowment, and grasp the dynamic changes of user groups in real time to better support the dynamic development of the tourism industry. It should be ideal for this special period to give full play to the advantages of the domestic tourism market, tap the tourism resources of the city, and appropriately shift to a suburbanized, small-scale, and decentralized tourism model like Rome and London. The combination of the traditional crowded and tourist attraction oriented development with small-scale suburban decentralized tourism model will make the tourism industry more flexible. The research on the impact of the epidemic on 12 typical highly internationalized megacities helps us understand how cities respond to crises and provide a reference for building long-term resilience to promote the sustainable development of cities. There is still a long way to go to implement research results into practice, and more in-depth research is needed in the future.

5. Conclusions

This study analyzed the impact of COVID-19 on Airbnb vacation rentals in 12 highly internationalized megacities through data mining, illustrating the changes caused by the pandemic. The images of all the twelve cities were reshaped. The impact time was related to the break time of the pandemic and the time of the implementation of the external border restrictions and lockdown policies. Booking rates decreased to different degrees, except in Beijing, in which the landlords withdrew their housing, creating a “super high unavailable rate.” The tourists composition changed in the cities, and there were fewer foreign tourists. The spatial distribution of reservations showed a trend of suburbanization. Tourists’ sentiment fluctuated, especially among foreigners, and tourists showed less attention to tourism-related activities.

Indeed, the period selected in this study is relatively short in terms of the pandemic’s entire development process. However, it was during the most critical period after the pandemic first broke out when people were unprepared, and the impact on vacation rentals was also unexpected. It has great significance for enlightening future research on the impact of such crises. As the pandemic continues, a longer-term extended study is needed to explore the changes dynamically. In addition, the impact of the pandemic on vacation rentals is the result of complex external factors. To find the mechanism behind it, we need further details, and more potential factors should be considered. It is also vital to examine whether this contingency factor will permanently affect these results or which pattern may turn normal in the future, a prerequisite for the subsequent formulation of corresponding countermeasures.

This article explored the impact of the pandemic on vacation rentals—the upcoming trend in the accommodation sector. In the future, we can make comparisons with the impact on the traditional hotel industry to further explore the difference between these two rental models, better integrate rental resources, and maintain sustainable development. We studied the impact of highly internationalized megacities. While ordinary cities might be different, a comparative analysis of the impact on different types of cities is warranted. In addition, it is necessary to combine the big data and traditional data to explore the underling mechanism behind the impact results. These are all important issues that need to be further resolved.

Notes

① The average international tourist arrivals (million) of sample cities in 2017, 2018 and 2019. Source: Top 100 City Destinations: 2019 Edition. Euromonitor international. https://go.euromonitor.com/white-paper-travel-2019-100-cities.html
② United Nations. Department of economic and social affairs, population dynamics. World urbanization prospects 2018. Retrieved from https://population.un.org/wup/Download/
③ Airbnb—The world’s largest online rental platform. Airbnb is one of the world’s largest marketplaces for unique, authentic places to stay, offering over 7 million accommodations, all powered by local hosts. By August 2016, more than 2 million Airbnb listings were located in 34,000 cities and 191 countries across the world, with the highest numbers concentrated in Paris, France (78,000), London, United Kingdom (47,000), New York, United States of America (46,000), Rio de Janeiro, Brazil (33,000), Barcelona, Spain (23,000), Rome, Italy (23,000), Copenhagen, Denmark (20,000), Sydney, Australia (20,000). It has more than three quarters of a billion guest arrivals to date, and accessible in 62 languages across 220 countries and regions (Airbnb, 2020).
④ The reference for Table 4 The overnight visitors arrivals of the sample cities in 2019 and 2020 is below:
1 Barcelona. Statistics and surveys. https://ajuntament.barcelona.cat/turisme/en/estadistiques_enquestes
2 Beijing. The outline of Beijing tourism statistics. 2021.
3 Berlin. Tourism Stats. Facts & Figures. https://about.visitberlin.de/en/materialien/toolkit/tourism-stats
4 Copenhagen. Reports and Insights. https://www.wonderfulcopenhagen.com/wonderful-copenhagen/analyses-insights/reports-and-insights
5 NYC. The Tourism Industry in New York City. OFFICE OF THE NEW YORK STATE COMPTROLLER. https://www.osc.state.ny.us/files/reports/osc/pdf/report-2-2022.pdf
6 Singapore. Statistics & Market Insights Overview. Source: Singapore Tourism Board. https://www.stb.gov.sg/content/stb/en/statistics-and-market-insights/statistics-and-market-insights-overview.html
7 Sydney. Travel to Sydney Tourism Region Year ended December 2020. Source: National and International Visitor Surveys, TRA. https://www.visitbritain.org.uk/wp-content/uploads/2021/05/travel-to-sydney-time-series-dec-2020.pdf
8 Vancouver. Visitor Volume or Overnight Visitors to Metro Vancouver. Marketing research. https://assets.simpleviewinc.com/simpleview/image/upload/v1/clients/vancouverbc/YTD_Visitor_Volume_2020_4f985a5-5de4-4e22-9e9f-dace5ac5071.pdf

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.
Acknowledgements

The China Scholarship Council (Grant No.201906120308) supported this work. Special thanks are given to Dr. Filip and Ph.D. candidate Tanya from National University of Singapore for their help in providing access to the public rental data sources and language polishing respectively, and to the reviewers and the editors for their valuable comments and thorough editing efforts. Their contribution has indeed improved the work. Acknowledgements to Airbnb and Inside Airbnb for their public data source permission.

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.scs.2021.103121.

Appendix A

A Figure and an Excel file. Confirmed cases per million people in twelve countries. (Calculated with the number of confirmed cases and the countries’ population derived from the Center for Systems Science and Engineering (CSSE) at John Hopkins University and United Nations, respectively.

Appendix B

Policy announcements from governments of the twelve countries were derived from the large hand-coded data set of more than 15,000 separate policy announcements from governments released by the CoronaNet Research Project supported by over 500 scholars and research assistants from all over the world.
Appendix C

Share of comments written in English for each city is shown in the Table below. $N_t$ represents the total number of comments written in all languages. $N_{en}$ represents the number of comments only written in English. Percentage ($N_{en}/N_t$) represents the share of English comments in all comments.

Table C1.

References

Airbnb. (2020). Airbnb. Fast Facts. Retrieved from https://news.airbnb.com/about-us/.

Aliperti, G., & Cruz, A. M. (2019). Investigating tourists’ risk information processing. Annals of Tourism Research, 79(September 2018), Article 102803. https://doi.org/10.1016/j.annals.2019.102803

Araña, J. E., & León, C. J. (2008). The impact of terrorism on tourism demand. Annals of Tourism Research, 35(2), 299–315. https://doi.org/10.1016/j.annals.2007.08.003

Becken, S., & Santana-gallego, M. (2020). The effects of natural disasters on international tourism: A global analysis. Tourism Management Perspectives, 20, 195–203. https://doi.org/10.1016/j.tmp.2016.09.004

Beria, P., & Lunkar, V. (2021). Presence and mobility of the population during the first wave of Covid-19 outbreak and lockdown in Italy. Sustainable Cities and Society, 65 (September 2020), Article 102616. https://doi.org/10.1016/j.scs.2020.102616

Blake, A., & Sinclair, M. T. (2003). Tourism crisis management. Annals of Tourism Research, 30(4), 813–832. https://doi.org/10.1016/S0160-7383(03)00056-2

Byers, J., Proserpio, D., & Zervas, G. (2013). The rise of the sharing economy: Estimating the impact of Airbnb on the hotel industry. SSRN Electronic Journal, 1–45. https://doi.org/10.2139/ssrn.2366898

Cahyanto, I., Wiblishauser, M., Pennington-Gray, L., & Schroeder, A. (2016). The dynamics of travel avoidance: The case of Ebola in the U.S. Tourism Management Perspectives, 20, 195–203. https://doi.org/10.1016/j.tmp.2016.09.004
Wu, E. H. C., Law, R., & Jiang Brianda, B. (2010). The impact of infectious diseases on hotel occupancy rate based on independent component analysis. *International Journal of Hospitality Management, 29*(4), 751–753. https://doi.org/10.1016/j.ijhm.2009.07.001

Zeng, B., Carter, R. W., & De Lacy, T. (2005). Short-term perturbations and tourism effects: The case of SARS in China. *Current Issues in Tourism, 8*(4), 306–322. https://doi.org/10.1080/13683500508668220

Zenker, S., & Kock, F. (2020). The coronavirus pandemic – A critical discussion of a tourism research agenda. *Tourism Management, 81*(June), Article 104164. https://doi.org/10.1016/j.tourman.2020.104164

Zhang, C., Qian, L. X., & Hu, J. Q. (2020). COVID-19 pandemic with human mobility across countries. *Journal of the Operations Research Society of China*. https://doi.org/10.1007/s40305-020-00317-6