AN EXPLORATORY STUDY OF AIRBNB CUSTOMER REVIEWS AND IMPACT OF COVID - 19

A Thesis
by
GOPIKA MALIK

Submitted to the School of Graduate Studies
at Appalachian State University
in partial fulfillment of the requirements for the degree of
MASTER OF SCIENCE

May 2022
Department of Applied Data Analytics
Walker College of Business
AN EXPLORATORY STUDY OF AIRBNB CUSTOMER REVIEWS AND IMPACT OF COVID - 19

A Thesis
by
GOPIKA MALIK
May 2022

APPROVED BY:

______________________________
Lakshmi Iyer, Ph.D.
Chairperson, Thesis Committee

______________________________
Indika Dissanayake, Ph.D.
Member, Thesis Committee

______________________________
Jason Xiong, Ph.D.
Member, Thesis Committee

______________________________
Neel Das, Ph.D.
Director, MBA

______________________________
Marie Hoepfl, Ed.D
Interim Dean, Cratis D. Williams School of Graduate Studies
Abstract

AN EXPLORATORY STUDY OF AIRBNB CUSTOMER REVIEWS AND IMPACT OF COVID - 19

Gopika Malik
M.S. Business Administration, Delhi University
M.S. Applied Data Analytics, Appalachian State University

Chairperson: Lakshmi Iyer, Ph.D.

Airbnb has experienced meteoric growth in the sharing economy business model since it launched in 2009. The Airbnb business model helps exploit underutilized assets and helps travelers stay in a home away from home to fulfill their travel needs at a much cheaper cost. Due to the outbreak of the COVID-19 pandemic, the operations impacted the overall decline in Airbnb activity. The depth study of how Airbnb activity and the rental market have changed during the COVID-19 pandemic using the comprehensive record of Airbnb listings, calendar, and reviews data. It will help understand how hard the business model is impacted by the pandemic. We propose an analytical framework of natural language processing on the text data of reviews left by the customers based on their experience with the property and host. Airbnb properties from the top five cities where Airbnb has the highest demand rate across the United States were considered for the research. The study found that COVID-19 had a catastrophic impact on the sector leading to massive cancellations and recorded low bookings with adverse effects on annual accommodations bookings for 2020.
globally. However, as anticipated, the suspension of global travel due to the pandemic was foreseeable, but the drift was temporary and merely shifted to new tourist locations. Even with the impact of a pandemic, the downsizing of the sharing economy will not last long. Even though the tourists believed the model was risky, the increased uptake in Airbnb listings was more affordable than equal accommodation offerings like hotels, amidst several cleaning and sanitization restrictions. With the lift in the restrictions, booking numbers for home rentals have picked up faster than hotels. We perform exploratory data analysis for the city-state; the impact of COVID-19 on the sharing economy in different cities; Analyze customer reviews using sentiment analysis to differentiate between positive and negative reviews and check the robustness of these reviews. We also found that the customer interaction with Airbnb experiences has increased compared to the pre-pandemic era. The reviews help spread the word about what tourists are expecting, help revert pandemic impact, and attract businesses to recover with higher survival rates. We point out implications and avenues for future research based on the findings.
Acknowledgments

I would like to express my sincere gratitude to my committee, Dr. Lakshmi Iyer, Dr. Jason Xiong, and Dr. Indika Dissanayake. This effort would not have been possible without their support and domain knowledge in data science, natural language processing, and data exploration. They encouraged me with constant guidance, patience, and insight. I would like to thank my thesis chair, Dr. Lakshmi Iyer, for her continuous guidance throughout the process. She has inspired me to continue to study and was influential in shaping my experiment methods and critiquing my results. I am especially indebted and would like to thank Dr. Indika Dissanayake, who involved me with the idea, topic for research, and support in designing the methodology. The project would not have been possible without the constant support of Dr. Jason Xiong. He has continuously poured his creative ideas and encouraged me to challenge myself to reach the pinnacles of the finest.
# Table of Contents

Abstract ............................................................................................................................ iv

Acknowledgments ........................................................................................................... vi

Chapter 1: Introduction .................................................................................................. 1

Chapter 2: Review of the literature .............................................................................. 5
  Sharing Economy ............................................................................................................ 5
  Boom in the Sharing Economy ...................................................................................... 8
  Outbreak of Global Pandemic COVID-19 ................................................................. 9
  Hard Hit on the Sharing Economy ................................................................................ 10

Chapter 3: Research methodology ............................................................................... 15
  Data Sources ................................................................................................................ 15
  Trend Amongst Cities ................................................................................................... 16
  Methodology ................................................................................................................. 20

Chapter 4: Results and discussion ............................................................................. 22
  Host with many properties ............................................................................................ 22
  Average reviews per listing ......................................................................................... 26
  Distribution of room types and average price per listing ............................................... 29
  Property type and Host about information .................................................................. 33
  Text Review Analysis ................................................................................................. 36
Chapter 1: Introduction

Short-term rentals (peer-to-peer platforms) are one of the perfect examples of the sharing economy. Airbnb has enabled individuals to earn income by mutually putting the under-utilized property for accommodation via fee-based sharing. Online platforms such as Airbnb, VRBO, and HomeAway have pioneered shared accommodation by providing a platform to connect people who own unused assets with those who require short-term rentals at cheaper rates. The sharing economy business model involves the usage of underutilized assets by applying communities with temporary ownership, which helps create a larger workforce, a collaborative economy, and the opening of job opportunities to more individuals (Vinod & Sharma, 2021). Since its inception in 2008, Airbnb has increased exponentially. The Airbnb platform was easily reached in more than 1,000 cities worldwide and was used by more than 500 million guests by April 2019 (Airbnb, 2019). Airbnb has received 250 million reviews from guests as of 2019, showcasing the customer's trust and faith in the model (2019). “In 2018, Airbnb had a market valuation of nearly $31 billion, $2.6 billion in profit, and $93 million in revenue.” (2019).

In January 2020, the Coronavirus Disease outbreak was seen worldwide, with more than 180 countries being affected (CDC, 2021). During that time, lockdown and some significant precautionary steps impacted many businesses, including the sharing economy.
Narasimhan et al., (2017), shared, “Forecasts indicate that the global sharing economy market can grow exponentially from US$15 billion in 2015 to US$335 billion in 2025.” However, COVID-19 has severely hit the sharing economy. According to a report by Statista, in comparison to the year 2019, the cancellations exceeded bookings for March and April 2020, leading to gross booking value crashing by more than 100 percent (Richter, 2020). In addition, the pandemic has disrupted the rental market – the peer-to-peer rental market since the World Health Organization declared the novel coronavirus outbreak a public health emergency of international concern on January 30, 2020.

We investigate how travelers changed perceptions of staying at an Airbnb during the pandemic. Specifically, we compare consumers' behavior and preferences with using Airbnb before (pre-pandemic) and during the pandemic. In our multi-city analysis, Airbnb properties from five different cities across the United States, where Airbnb has the highest demand rate, were used for the study. These cities are - Nashville-TN, Jersey City-New Jersey, Los Angeles-California, Austin-Texas, and Boston-Massachusetts. Due to the pandemic, sanitation, cleanliness, and hygiene have become more critical. Next, we analyze the travelers' reviews and study how hosts manage their listings and the top priorities. Finally, we are looking to compare customer perceptions, examine their influence on customer satisfaction, and explore the factors that influence guest satisfaction.
The following research question guided this study:

1) How has Airbnb activity changed due to the COVID-19 pandemic, and what have been the customers' preferences on rental activity?

2) How do the choices of room types/property types change with the price compared to the hotel industry during the COVID-19 pandemic?

3) What are the significant determinants of the count of customer reviews on Airbnb during the pandemic?

4) How do these determinants influence the count of customer reviews on Airbnb across different states and cities?

To answer the research questions, listing, reviews, and calendar data were extensively researched. The literature was reviewed to determine the rate of bookings cancellations and how the outbreak of COVID-19 paralyzed travel and tourism. We compared the listing data from 2019 with 2020 to assess the impact on bookings and the cancellation rates. We have used various publications - such as news articles, TV news items, YouTube videos, and blog posts, to study how the sharing economy phenomenon is coping with the changing environment due to the pandemic. The Airbnb platform allows users to post public reviews on their experience with Airbnb and help other users make better decisions based on the reviews. Studies have investigated how knowledge from their public reviews, even though customer-generated, plays a vital role in decision making (Troisi et al., 2018). Inspired by
this, this study attempts to mine public reviews and exploit knowledge to provide meaningful implications. Most reviews left by the customers after the end of their stay are a true reflection of their experience. Hence it is imperative to analyze the textual reviews to understand the gap between customer expectations and services offered.

Data analytics tools help interpret extensive data collected over time. Python, an open-source programming language, was used to derive insights and lead to conclusions. The Jupyter notebook compiled the code and created analysis and visualizations.

A natural language framework is used to text mine the reviews data to find out the themes in the reviews change over time. COVID-19 has raised the question of the survival of the sharing economy. To understand how COVID-19 has affected this sector, the various facets of the sharing economy were explored. Through the data analysis, the Airbnb data were explored to help identify the change in the behavior of the sharing economy and customers with respect to the response to the pandemic. Using natural language processing, it is believed that with the quantitative analysis of the textual data, one can better understand the experiences made with the Airbnb services. Hence, Airbnb’s textual reviews were chosen to be the unit analysis of this study for uncovering Airbnb's service attributes expected by the customer. The remainder of the paper is structured as follows: First, it presents the literature review on the sharing economy. The following section explains the research methods, including the data and its collection and analysis. Moving forward, we present the results of this study. Lastly, the highlights are some implications and future research avenues in the final section.
Chapter 2: Review of the Literature

The sharing economy is facing a boom in the short term rental market. Also known as collaborative economy, it is based on the pillars of information technology and network services development. This literature review chapter of the study is looking to articulate the review of the literature by other dignified scholars and authors to determine the understanding about the sharing economy, how it works, what are the factors that led to the boom in the sharing economy. In the following sections, there will be a discussion about the sharing economy, boom in the sharing economy, the outbreak of the pandemic COVID-19 and how pandemic impacted the sharing economy. The literature written by the other authors was studied, and reviewed to understand the key concept behind the sharing economy and the impact of COVID-19. In this section, the study efforts were made to determine and identify the crucial gaps on the existing literature on the relationship between digital platforms and the sharing economy and provide how sharing market was hit by the pandemic and provide future investigation.

Sharing Economy

Scholars have studied the exponential growth in benefits of the sharing economy in a range of industries, such as transportation(e.g., Uber), lodging (e.g., Airbnb, Vrbo), food delivery (e.g., Doordash, Uber Eats), and household services(e.g., Care.com). The navigation
for smart, sustainable cities has researched various ways technology can enable the efficient use of limited resources and "idling capacities" (underutilized physical assets), which will help reduce waste and improve the environmental sustainability of the cities (Bernardi & Diamantini, 2018). However, the author never explained how a sharing economy could help improve environmental sustainability. Through this study, the effort is to understand the critical pillars of the sharing economy. The sharing economy is a new business model based on peer-to-peer sharing of under-utilized resources. Sharing economy is an "activity of obtaining, giving, or sharing access to goods and services, coordinated through community-based online services" (Hamari & Ukkonen, 2013, pp. 2047-2059).

Access-based consumption, a peer-to-peer economy, and a platform economy are some of the ways to explain collaborative consumption in the sharing economy (Langley & Leyshon, 2017). Cheng conducted an extensive review of the literature on sharing economy published focusing on hospitality and tourism. They have identified two broad perspectives of the extant literature in this domain wherein the sharing economy (SE) impact is divided into two clusters; Figure 1 explains: 1) SE’s impacts on destinations and tourism services, and 2) SE’s impacts on tourists. Cheng emphasized “Tourism” as the domain theme as the sharing economy challenges people’s traditional ways of using tourism services. The concepts of stay, guests, and host indicate SE’s impacts on destinations. The idea of employment also shows SE’s impact on the traditional labor market. The theme “Airbnb” means the rapid growth of Airbnb has significant impacts on hotels and rentals in terms of market share and competition.

The literature aptly explains the theme “website” from the above is a relatively small theme (light blue) containing a website and home by connecting with hosts demonstrating the role of websites in facilitating host home rental. SE’s impacts on tourists include three
themes: “social,” “behavior,” and “members.” Social is the domain theme showing the role of technology in SE on tourists, social exchange and alternative experiences, and the moral dimension of the impacts on tourists. However, the underlying principles behind the sharing economy, the internet technologies, like smartphone apps and trusting strangers, are two key concepts that are not much emphasized. In this study, technology and trust are valued critical pillars of the sharing economy.

**Boom in the Sharing Economy**

A shared home is a peer-to-peer model of short-term rentals at lower rates, with more extensive accessibility than traditional hotels and lodges (Biswas et al., 2021). However, the actual booking data is not available publicly, so scholars have adopted the "count of consumer reviews" as an alternative measure (Langley & Leyshon, 2017). Airbnb's business model currently operates with minimal regulatory controls in most locations. Platforms might also conduct direct surveillance and encourage customers to leave ratings for other users. According to Akande, these activities cultivate a robust system of reputation and accountability throughout the network and may increase the individual participant's trust in the platform's ability to operate efficiently. However, the author should explain how the crime rate can impact trust, the leading pillar of the sharing economy. The sharing economy has affected many industries, such as transport (e.g., Uber) and accommodation (e.g., HomeExchange) (Akande, 2020). The literature by Guttentag investigated tourists’ motivations for using Airbnb and aggregated results that indicated five motivating factors - Interaction, Home Benefits, Novelty, Sharing Economy Ethos, and Local authenticity. They
found that the basic phenomenon of locals informally renting lodging to tourists has been in practice for many years. However, the new Internet applications have revolutionized this practice and allowed it to scale exponentially by facilitating virtual markets where communication and trust are established between hosts and guests (Guttentag et al., 2017). The literature should also discuss the issues that cropped up by subletting the property for short-term rental. In this study, the problems neighbors face due to the influx of so many new people in the Airbnb properties will be discussed briefly.

**Outbreak of Global Pandemic COVID - 19**

COVID-19 has adversely affected the sharing economy. There were various approaches to dealing with the pandemic as preventive measures, such as lockdown and social distancing. Airbnb has been one of the hardest-hit platforms by COVID-19 (Curry, 2022). Since the onset of the outbreak, Airbnb listings on offer have remained relatively steady. However, the number of bookings crashed due to travel bans and health concerns, according to the vacation-rental site AirDNA. The reduction in bookings was most significant in COVID-19 epicenters. This impact was felt in March 2020, with average Airbnb occupancy rates in May 2020 dropping to 11% in the United States (Average Airbnb Occupancy Rates By City [2022] | AllTheRooms, 2022). These lost bookings had a calamitous impact on Airbnb's revenue, as seen from the recent layoff of about a quarter of its workforce (Hu & Lee, 2020). “Airbnb CEO Brian Chesky noted in May of 2020 that revenues could decline more than 50% during the year, prompting him to cut costs, raise $2 billion in capital and let 1,900 employees go (25% of the total)” (Airbnb, 2020). Dolnicar et
al., 2021 explain how quickly hosts and co-hosts developed creative ways to continue operating. However, the detailed information about those steps was not mentioned. The literature about the capitalist hosts who viewed their properties as an investment could not continue paying all their expenses, exiting the short-term rental market, or moving short-term to the long-term rental market. The survey was conducted for the hosts and co-hosts from Canada, the UK, the US, New Zealand, and Australia in January 2021, and all widely reported trouble paying their bills, including mortgage payments. The missing detailed information about the survey would have helped understand the sample case behind the study and how far it is correct in applying to all the listings under Airbnb.

**Hard Hit on Sharing Economy**

The Roof and Carville, in their literature, discussed that as the summer of 2020 progressed, revenues for the sharing economy did decline. “Still, Airbnb experienced more booked nights for US listings between May 17 and June 3 of 2020 than during the same period in 2019 and saw a similar increase in domestic travel globally” (Roof & Carville, 2020). We will be guiding our study to find out the truth behind these projections- The researcher could see that Airbnb could thrive, as customers see the accommodations as being much less risky than traditional hotels concerning the spread of the virus (Clifford, 2020). However, this was a prediction and with no statistics support.

Tussyadiah (2016) mentioned, “The COVID-19 pandemic might have induced a higher need for personal ‘physical’ distance, especially with non-relatives. This tendency
might translate into a reduced quest for, or sensitivity to, social interaction/experience and the related low levels of required social/physical distance that characterize some segments of peer-to-peer accommodation customers.” The author mentioned that the spike in the safety protocols increased the need for physical distance to avoid infection, which directly impacted the accommodation choices. However, the literature doesn’t explain the impact on hosts with a change in the selection of the listings. This study will demonstrate the percentage change in the choice of listings type before and during pandemic comparison. In literature, the author explained that “the selection revolves around two criteria: i) studies exploring different types of accommodation (comparing hotels vs. peer-to-peer, but also considering different types within peer-to-peer), and ii) studies investigating the effects of social/physical distance in peer-to-peer models and traditional models. (Bresciani et al., 2021).

In his literature, the author highlighted the missing information on the type of accommodation (shared room vs. entire home) when he analyzed the Airbnb accommodation type (Bresciani et al., 2021). Traditionally, paid peer-to-peer models are classified into three broad groups: Shared room, private room, and entire home (Airbnb, 2020). A shared room generally means that the traveler sleeps in a common space. A private room usually indicates that the traveler is entitled to use a single room in a flat or house. Finally, an entire home generally implies that the traveler has exclusive use of the whole property. This represents an essential aspect of COVID-19. The author in her literature highlighted that the COVID-19 and the social/physical distancing requirements are imposing constraints on the first two groups because a portion of travelers is afraid of sharing a house or room with strangers.
Therefore, the structure of accommodations (shared house vs. the entire house), with impact due to the decreased social/physical interaction and an increasing fear of infection, dramatically alters the demand for paid peer-to-peer accommodations during COVID-19 (Bresciani et al., 2021). This study will share how these numbers changed due to the pandemic. Dolnicar et. al, 2021 explained that there were significant requirements and changes in the cleaning protocol as requested by customers and reported by the hosts, which in turn likely led to the evident spike in time and money spent on cleaning. In response to increases in the cost, 42% of the hosts had declined the booking requests based on the country of origin of the people requesting the booking, whether they are from the COVID-19 epicenters. The author also emphasized that the overwhelming financial pressure associated with the pandemic has left many hosts to leave the peer-to-peer accommodation market either temporarily or indefinitely as they fail to pay their monthly installments and bills. However, this literature didn’t highlight the new joiners’ hosts on Airbnb. These were the people who had lost their jobs due to the pandemic and now subletting their primary residence to earn income for survival. This study highlights the drop in the listings, how numbers changed from before COVID - 19 to during the pandemic and how fast these dropped bookings have picked up.
Figure 2

_Gross bookings reflections of Airbnb for the year 2020_


In Figure 2, this can be visualized that the COVID 19 has impacted the sharing economy; gross booking value crashed by more than 100 percent in March and April compared to the previous year, meaning that gross bookings were negative as cancellations exceeded bookings (Richter, 2020).
Chapter 3: Research Methodology

The exploratory data analysis study is based on carefully planned data collection. The data was collected and combined from the website insideairbnb.com. The Inside Airbnb website is managed by Murray Cox and provides archives of data gathered through web scraping of the Airbnb website. InsideAirbnb's website has reviews, listings, and calendar data for all the target five cities of study, combined using the listing ID for comparison purposes. The adverse effects of COVID-19 on travel activity due to 1) The lock-down, 2) New cases of COVID-19 infections 3) The government’s policies on travel bans. We conduct a comparative exploratory data analysis of different variables present in the dataset to find consumer behavior changes and their impact on the Airbnb business pre-COVID and during covid amongst the selected five cities in the U.S. The following sections are going to explain the information about the data sources, information about the Airbnb’s penetration in various cities and each city’s geographic, demographic, and socio-economic characteristics.

Data Sources

The data source for each state provides three types of data, Listings, reviews, and calendars. Listing data includes information on listing, host, neighborhood, listing price, size of the listing, and listing’s specific coordinates in longitude and latitude. The reviews data set includes the reviewer's name, listing id, and detailed text review. To advance the analytical framework and depth of research, the text reviews data was analyzed using techniques from natural language processing. The calendar data provides information about the availability calendar for 365 days in the future. The users are referred to as "hosts" and their properties
as their "listings." Each host is associated with attributes including a photo, a personal statement, listings, guest reviews of their properties, and Airbnb-certified contact information. Similarly, each listing displays characteristics including location, price, a brief description, photos, capacity, availability, check-in and checkout times, cleaning fees, and security deposits. The collected data set contains detailed information on the distinct hosts and their distinct listings of all five cities spanning from 2019 to 2020. This current data covers all listings for five cities scraped from the Airbnb website. While data for later periods was available, this time was selected to compare the Airbnb penetration in all five series and analyze the impact of the COVID-19 pandemic on the sharing economy. This information is summarized in Table 1 below.

### Table 1

**Source information of primary datasets used in this analysis**

<table>
<thead>
<tr>
<th></th>
<th>Austin</th>
<th>Boston</th>
<th>Los Angeles</th>
<th>Nashville</th>
<th>New York</th>
</tr>
</thead>
<tbody>
<tr>
<td>listings</td>
<td>11,374</td>
<td>3,349</td>
<td>33,329</td>
<td>391,498</td>
<td>38,277</td>
</tr>
<tr>
<td>Calendar</td>
<td>4,151,151</td>
<td>1,221,655</td>
<td>12,161,485</td>
<td>1,048,575</td>
<td>13,964,535</td>
</tr>
<tr>
<td>Reviews</td>
<td>382,362</td>
<td>128,311</td>
<td>1,144,105</td>
<td>391,498</td>
<td>891,964</td>
</tr>
</tbody>
</table>

*Note.* This table demonstrates the detailed records for all five cities of the study.

### Trend Amongst Cities

We study the relationship between Airbnb’s penetration in various cities and each city’s geographic, demographic, and socio-economic characteristics. We gather data about Airbnb’s listings for five U.S. Cities - Austin-Texas, Boston-Massachusetts, Los Angeles-California, Nashville-Tennessee, and New York City-New York. We have chosen
the cities ranked as the top cities where Airbnb has the highest demand rate (*The Best Cities for Airbnb: Compare Markets Side by Side* | AirDNA, 2019). All these cities vary in size, population composition, and cost of living. We propose a method of exploratory analysis of Airbnb listings and review data to study the relationship between Airbnb’s penetration and the impact of COVID-19 on the sharing economy in different cities. Let’s learn about these cities in the next section.

**Austin**

Austin is the state capital of Texas. It is an inland city bordering the Hill Country region. Austin is the home to the University of Texas flagship campus; Austin is known for its eclectic live music across the country. According to Airbnb, “In 2019, travelers spent $83 billion on travel in Texas. But during the pandemic, that dropped to only $56.9 billion, a decline of more than 31%” (Visit Austin, 2021). Deane mentioned that Austin is one of the most popular destinations on Airbnb and ranked top amongst the destinations wherein Airbnb is cheaper than hotels (Deane, 2022). It stands in third place amongst the most Airbnb listings, with an average of 168,247 listings amongst other states during 2021. Austin had the most significant average number of bedrooms per active Airbnb property in 2015, with 1.8 bedrooms per unit (Richter, 2020). In this research it was mentioned that Austin, Texas, came in third place with $342.4 million in gross revenues on Airbnb in 2021 (Richter, 2022).

**Boston**

Boston is ranked the 24th-most populous city in the United States and is full of historical places and heritage monuments. Boston is the capital city in the Commonwealth of
Massachusetts and is an ongoing center of scientific research. The Boston area is home to many colleges and universities, making it a world leader in higher education (Wikipedia). According to the Greater Boston Convention and Visitors Bureau, 22.7 million people visited Boston in 2019. Due to COVID-19 and the state’s stay-at-home advisory, the occupancy rate dropped from 72.7% in February 2020 to 5.3% in April. Overall, the lodging market in Boston and Cambridge’s occupancy rate dropped to less than 26% in 2020, driving revenue per available room down more than 80% percent, according to the Pinnacle Advisory Group (2021). Based on occupancy data, Boston ranks among the top twenty cities for Airbnb (Deane, 2022).

**Los Angeles**

Los Angeles is the largest city in the state of California. According to the information from Wikipedia, in the year 2020, the population of Los Angeles reached 3898747; Los Angeles is the second-largest city in the United States. Tourism in Los Angeles thrived in 2019, with 51 million visitors spending more than $27 billion. Tourism spending dropped by 67.2% of the 2019 amount. Airbnb ranks the greater Los Angeles area as “the nation’s fourth most profitable region for new Airbnb hosts, with one U.S. listing during the first six months of 2021” (Smith, 2021). Los Angeles hosts 70.751 listings in the year 2020. It ranked amongst the top five cities with the most Airbnb demand globally (1,218,884 nights stayed) and generated gross revenue of $203,650,634 in 2020; It is the home to vast short-term rental markets (AirbnbStatistics, 2022). Los Angeles is also amongst the cities with the highest number of commercial Airbnb listings. After Florida and New York, it is amongst
the most popular destinations on Airbnb. Based on occupancy data, Los Angeles is amongst
the top twenty cities for Airbnb (Deane, 2022).

**Nashville**

The area selected for study is the city of Nashville, Tennessee. The city of Nashville is home to over 689,000 residents, with nearly 2 million people living in the area. Nashville set a record of 16.1 million tourists visiting in 2019, according to data released on January 27 by the Nashville Convention & Visitors Corp. "Nashville is ranked No. 18 on financial company IPX1031's, Best Cities to Invest in Airbnb list", compiled from data made available by short-term rental analytics database AirDNA. By listing a property on Airbnb, a host can make $17,100 a year based on the average $209/night price (Medina, 2021). The city topped the charts as one of the best places to visit in the United States and set a tourism record in 2019 with 15.8 million visitors (*Nashville Receives $7 Billion in Visitor Spending*, 2019). The love of Country music drives tourism and travel toward the city of Nashville, boosting its housing market. All the people going on vacation need a place to stay, making this an excellent opportunity for Airbnb investors owning vacation home rentals in the region (Baker, 2018).

**New York**

New York City comprises five boroughs and is the most densely populated and one of the most famous cities globally. According to the comptroller, New York, “New York City hosted 66.6 million visitors in 2019, about 25% of the United State’s 265.5 million visitors, a tenth-consecutive annual record. However, in 2020, the 10-year period of record growth in
tourism ended, and the number of visitors to New York City fell by 67% to 22.3 million visitors” (About the Comptroller’s Office, 2021). According to state law in New York City, it is illegal in most buildings for an apartment to be rented out for less than 30 days unless the permanent tenant is residing in the apartment at the same time. According to Deane (2022), New York City is the next most famous city on Airbnb after Tokyo. Amongst the United States, New York City is the most popular destination on Airbnb after Florida. However, Airbnb users may only list one home at a time; in New York, 72% of Airbnb hosts use their revenue to remain in their homes. Amongst all the listings, 55% of Airbnb listings in New York year 2016 were illegal. New York City is also among the top 20 cities for Airbnb based on occupancy data; the median income for Airbnb hosts in New York City in 2016 was $5474 (Deane, 2022).

**Methodology**

For analysis purposes, the users are referred to as "hosts" and their properties as "listings." The collected data set contains detailed information on the distinct hosts and their distinct listings of the selected cities spanning from 2011 to 2021. For comparison between pre-COVID and COVID, the two years of data wherein the data from 2019 referred to as pre-COVID and 2020 as COVID. Airbnb implemented several changes in its services towards hosts, guests, and third parties as to the onset of COVID-19 and the restrictions put in place by nations worldwide to protect the people from the pandemic. We used natural language processing to analyze textual reviews data; perform sentiment analysis and topic modeling to understand the reviews by customers based on their experience with the property
and the host. The spatial data helped determine the listings spread across cities and understand the people's preferences during pre-pandemic and pandemics. Before delving into the datasets, the data was cleaned and normalized for exploration. The steps were taken to remove columns with more than 50% nulls and missing values for information provided in percentages. All strings were converted to floats, all true/false information was transferred to boolean type, and price variables normalized to float. Missing values were replaced with averages for better data handling.

Hosts are also categorized as superhosts if they meet the criterion. A Superhost goes above and beyond in their hosting duties and is a shining example of how a Host should be. For a comparison study across cities, the super host count was one way to determine how many hosts were doing their job efficiently.
Chapter 4: Results and Discussion

The chapter 4 presents the results for the research questions. The dataset from InsideAirbnb was used for analysis of Airbnb penetration and change in the behavior of the customers during COVID. The below mentioned sections specify the results and discuss the limitations as well.

Host With Many Properties

Data collection from insideairbnb.com produced information about unique hosts with how many of them are super hosts in the listing dataset, individual listings in the calendar dataset, and the number of reviews from the reviews dataset. Table 2 exhibits the information about the number of records from each city about hosts, super hosts, listings, and reviews.

Table 2

*Information about the count of listings & hosts across cities*

<table>
<thead>
<tr>
<th>City-State</th>
<th>Austin</th>
<th>Boston</th>
<th>Los Angeles</th>
<th>Nashville</th>
<th>New York</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hosts</td>
<td>7,252</td>
<td>1,205</td>
<td>18,475</td>
<td>2,971</td>
<td>25,904</td>
</tr>
<tr>
<td>Super Hosts</td>
<td>0.3101</td>
<td>0.3012</td>
<td>0.3087</td>
<td>0.5385</td>
<td>0.1927</td>
</tr>
<tr>
<td>Listings</td>
<td>11,373</td>
<td>3,347</td>
<td>33,325</td>
<td>2,874</td>
<td>38,259</td>
</tr>
<tr>
<td>Reviews</td>
<td>383,362</td>
<td>128,311</td>
<td>1,144,105</td>
<td>391,498</td>
<td>891,964</td>
</tr>
</tbody>
</table>
The descriptive statistics were applied to the data to recognize for all cities accurately the hosts with unique listings, the neighborhood group of the listings, and the property type for the top ten hosts in all the cities. Sample findings from Los Angeles can be viewed in table 3 on the next page. Table 3 below shows the list of the top five hosts and the count of their listings, wherein host_id 107434423 has the maximum of 110 listings in Los Angeles City.

Table 3

*Top five hosts and count of their listings*

<table>
<thead>
<tr>
<th>Host ID</th>
<th>Count of Listings</th>
</tr>
</thead>
<tbody>
<tr>
<td>107434423</td>
<td>110</td>
</tr>
<tr>
<td>891818</td>
<td>107</td>
</tr>
<tr>
<td>271118401</td>
<td>106</td>
</tr>
<tr>
<td>144214204</td>
<td>101</td>
</tr>
<tr>
<td>101537031</td>
<td>100</td>
</tr>
</tbody>
</table>

The top host was further analyzed to find the spread of the locations from the neighborhood group for this host and found that his listings are all spread across the most visited places in the Los Angeles area.
### Table 4

*Top five locations of listings by Host #1 (Host_id 107434423)*

<table>
<thead>
<tr>
<th>Neighborhood Group</th>
<th>Count of listings</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hollywood</td>
<td>15</td>
</tr>
<tr>
<td>Beverly Grove</td>
<td>10</td>
</tr>
<tr>
<td>West Hollywood</td>
<td>9</td>
</tr>
<tr>
<td>Westchester</td>
<td>8</td>
</tr>
<tr>
<td>Sawtelle</td>
<td>6</td>
</tr>
</tbody>
</table>

The spread of all the listings in the Los Angeles area can be viewed from the shapefile showcasing the Airbnb spread across the city. The latitude and longitude provided in the listings dataset helped visualize the spread of the listings across the city.
Figure 3

Listings of Airbnb in Los Angeles City

Note. This figure demonstrates the number of listings across the city of Los Angeles. This helps understand the collection of significant listings is all closer to the downtown and center of the city. However, during the pandemic, the behavior of the travelers...
changed, and many of the tourists preferred to stay in remote locations away from crowded places (Cox & Chon, 2020). This drift in the choice of destinations is yet to be verified with the data using advanced programming skills with geocoding and will be the future work aspect. Also note, that hosts with multiple listings belonging to different cities and locations can open different Airbnb accounts to manage their listings. However, the host with multiple listings at one destination is not allowed to create duplicate accounts just to target their listings and increase their probability of showing up more often on the website (iGMS, 2021). In this study, all the unique listings and hosts IDs are considered for analysis purposes.

Average Reviews per Listing

Customers give reviews based on their experience with the property and the host. The listings are divided into four categories - Entire home, Hotel room, Private room, and Shared room. The reviews data for the pre-COVID year 2019 was used to find out the spread of the reviews based on the listings. The visualization in the below table showcases the spread of the distribution of the reviews based on the types of listings for the city of Austin.
Figure 4

*Distribution of the listings based on the reviews for Austin during Pre-COVID*

For the city of Austin, there are 187 average reviews on entire homes, 145 for hotel rooms, 141 for private rooms, and 25 for shared rooms. Entire rooms are the favorite choice of listings for the customers during pre-COVID. Figure 5 displays that during COVID entire homes remained popular among the customers. There are 177 average reviews on entire homes, 138 on hotel rooms, 145 on private rooms, and 75 on shared rooms. The histogram below shows the spread of reviews across the city of Austin during COVID.
Figure 5

*Distribution of the listings based on the reviews for Austin during COVID*

Note. Distribution of listings across different room types based on the reviews for the city of Austin during COVID.

It helps explain how the preferences of the customers changed with respect to the room type selections. The Entire home/apartment is the first choice of the people. The gap between hotel rooms and private rooms is very thin and shared rooms are the least choice for the customers. The same pattern was observed in all the cities in the study. According to Dolnicar (2021), the choice between an Entire apartment and a shared room has always been attractive from 2008 to 2019. However, the choice of the customer changed with the onset of the pandemic and is leaner toward Entire homes than shared apartments, and this has been verified from the data analysis results in Figure 5.
Distribution of Room Types and Average Price per Listing

The listings provide the options of room types for customers to choose from while booking. There are four different types of rooms, customers may choose Entire homes or apartments, Private rooms, shared rooms, or hotel rooms. Figure 6 below shows the distribution of room types before COVID in the city of Nashville’s Airbnb listings. The data analysis shows that before COVID, there were 89 entire homes, 10 private rooms, 1 hotel room, and 1 shared room was the choice of the customer at the time of the booking.

Figure 6

*Waffle chart displaying the distribution of room types pre-COVID*

Note. Distribution of different room types for the city of Nashville during Pre-COVID.
The waffle chart below shows the distribution of room types during COVID. Where the listings for entire homes increased to 92 and for private rooms, it decreased to 7. This is obvious due to the fact that people were sharing fewer rooms due to the pandemic.

**Figure 7**

*Waffle chart displaying the distribution of room types COVID*

![Waffle chart](image)

*Note.* Distribution of different room types for the city of Nashville during COVID.

The distribution of average price during the pre-COVID era in Nashville’s Airbnb listing is compared with the average price of a hotel room to display how the price varies for the customers. This drift in customer choice was expected as an aftereffect of the pandemic. Airbnb’s offerings also include a change in room type. The surge in the request for entirely independent accommodations such as Entire houses/apartments. Figure 8, as shown below
demonstrates that the average price per hotel room is $223. An average Airbnb listing costs $175 pre-COVID.

**Figure 8**

*Distribution of the prices compared with hotel prices for Nashville during Pre-COVID*

Note. Distribution of prices across Airbnb and a hotel room for the city of Nashville during Pre-COVID

With the onset of the pandemic, the cleaning, sanitization related procedures became extensive and led to an extra burden on the average rental cost of both Airbnb and hotel rooms. The gap in the price between Airbnb and average hotel rooms gets wider. Wherein the average price per hotel room increased to $230 and the average on Airbnb listings was the same, however, they charged extra afterward in different types of charges. Airbnb pricing model is not regulated and is in the control of the hosts. In this study the data analyzes the change in the average listing price of Airbnb property compared to a change in the average
price of the hotel room in Austin. Hotels are expensive, and rooms are available at a high price in Austin but the prices increase further due to the pandemic outbreak. The Airbnb host increased the average price per room, which used to be between $127 - $306, to $197 - $313; in comparison to the hotel industry, the average price per night of a hotel increased from $223 to $230. The change in price of the listing was not much higher during the pandemic, however the overhead charges including cleaning and sanitization spiked the price and showed up at the time of the payment.

There was a change in the distributions of the listings across districts during COVID. Figure 9 below shows the spike in the distribution of listings across districts when entire rooms/apartments were in demand. Due to COVID, people wanted to stay at far-off locations with individual apartments. The distribution of the districts across different areas will be explored in the future work.
Figure 9

Distribution of listings in Nashville during COVID in neighborhood areas categorized by room types.

Note. Distribution of listings for the city of Austin during COVID.

Property Type & Host About Info

Property type is distributed amongst various different categories based on different cities. The availability of these property types vary from entire town homes, entire guest houses, house boats, private rooms, bed and breakfast. The figure below demonstrates different types of properties listed on Airbnb in Boston. This visualization in figure 10 depicts the choice of the customers and their preferences in choosing the property type during COVID. It is observed that more listings were available for the entire unit and private rooms.
**Figure 10**

*Distribution of property type in Boston during COVID*

![Distribution of property type in Boston during COVID](image)

*Note.* Distribution of property type on the Airbnb listings in the city of Boston during COVID

The analysis was conducted using data to find out different property types available under property types on the listings in Airbnb with average price per night. When property type varies from entire town homes to boutique hotels, price per night varies from $324 to
$216 in the city of Boston. The table 5 below shows the distribution of prices amongst varying types of property in the city of Boston.

**Table 5**

*City of Boston property type by room type with price comparison*

<table>
<thead>
<tr>
<th>Property_type</th>
<th>Room_type</th>
<th>Price</th>
</tr>
</thead>
<tbody>
<tr>
<td>Entire townhouse</td>
<td>Entire home/apt</td>
<td>324.454545</td>
</tr>
<tr>
<td>Entire serviced apartment</td>
<td>Entire home/apt</td>
<td>296.735537</td>
</tr>
<tr>
<td>Entire residential home</td>
<td>Entire home/apt</td>
<td>270.068468</td>
</tr>
<tr>
<td>Boat</td>
<td>Entire home/apt</td>
<td>250.476190</td>
</tr>
<tr>
<td>Room in boutique hotel</td>
<td>Entire home/apt</td>
<td>216.134189</td>
</tr>
</tbody>
</table>

Many hosts wrote descriptive and detailed descriptions regarding their property for the purpose of attracting customers and providing an insight on their properties, which helps their listing stand out from the competition. Figure 11 below shows examples of the descriptions written by hosts directed at the customers in the city of Boston.
Information posted by the hosts from Boston about their listings

Note. Information about the listings for the city of Boston posted by the hosts

In aggregate, this study clearly presents an association between property type, room types and its price. The elaborated description of the property also attracts customers as it provides better understanding of the property which otherwise was not easily available. The future work could look at applying the natural language processing techniques on property descriptions to filter the listing by host type whether hosts are investors, owners, property management consultants, or real estate agents listing their properties.

Text Review Analysis

Text analysis was performed using natural language processing. The sentiment analysis of the reviews from the reviews data found that people are more positive about the Airbnb experience during Pre-COVID (see Figure 12). Similar steps of sentiment analysis
were performed on the COVID data and customers were 45.39% positive, 34% neutral, and 20% more likely to be negative which still explains that customers were more positive towards their experience of staying in an Airbnb property.

**Figure 12**

*Sentiment analysis of Airbnb reviews for NYC*

*Note.* Negative, Neutral, Compound, Positive sentiment scores from the text reviews by the customers for the listings from the city of New York.
Next, the analysis concluded by conducting a Topic analysis wherein it found the top 10 topics people were discussing both pre- and during the pandemic. Both the comparisons show that people who are using Airbnb services were pleased and satisfied with the services offered. Most of the time during pre-covid the most common topics reviews were about comfortable location, decorated room, outstanding neighborhood, and beautiful broadway. The main topics people discussed are safety, cleanliness, sanitization, beautiful property, privacy, and convenient location. All these words that come up from topic analysis showcased that the customers found themselves safe in these properties as they gave more privacy and less contact with the staff or outsiders, unlike in a hotel.

**Figure 13**

*Distribution plot for the length of comments in reviews by customers*
Customers gave lengthy positive reviews after completion of their stay in the property in NYC. These reviews help bring business as they are a guide for other customers planning to stay in these properties. The figure above shows the distribution plot of the length of comments in reviews by the customers. It can be clearly seen that positive reviews experience larger word counts than negative ones.

Analysis of word clouds on positively tuned comments in figure 14 showed that during COVID, people were more focused on cleanliness, distance from downtown, and expectations of a decorated clean place with a great neighborhood. Commonly used positive words captured in word cloud include words comfortable, relax, deciorated, clean, great reflecting the theme of positivity from the reviews.
Figure 14

*Word cloud of the positive tuned comments from reviews for the NYC*

**Positively Tuned Comments**

---

*Note.* Word Cloud from the positive reviews from the customers for the listings from the City of New York

Word cloud is a visual representation of the text wherein words which appear more often appear bigger in size. It helps understand the unstructured data and getting insights and patterns.

Analysis of word clouds during COVID on negative reviews reflects word themes like con, problem, war, die. Most of these words are in different languages and also reflect on
the diversity of the people using Airbnb platform. NYC is the only city amongst all the cities with word clouds in so many diverse languages.

**Figure 15**

*Word cloud of the negative tuned comments from reviews for the NYC*

![Word Cloud](image)

**Note.** Word Cloud from the negative reviews from the customers for the listings from the City of New York

The bar chart below in Figure 16 is the term frequency bar chart and is the representation of the number of times a given term appears in the document. It helps
determine the relevance of a document for a particular theme. The top four words on the bar chart include words like great, stay, place and clean. This bar chart clearly explains the theme during COVID wherein people are more inclined towards staying in a great place which is clean.

**Figure 16**

*Frequency of most used words in review comments during COVID*

*Note.* Mostly used words from the review comments during COVID for the City of New York
Through the text analysis using word cloud, and term frequency it is very clear that the theme people are looking for a great clean place to stay is the most prominent amongst the customers and is reflected in the reviews during COVID. Reviews gave a true picture of the customers’ experience with the property and the host. It also reflects reputation of the host, property and Airbnb system.
Chapter 5: Conclusions, Implications & Future Research

Tourism is a human endeavor that has been evolving since the start of the twentieth century and is continuously expanding. The short term rental giant Airbnb plays a considerable role in today’s tourism trends. Airbnb supports tourism wherein hotels fail to meet demand and also allowing travelers to stay in destinations wherein either they could not afford or hotels do not exist. In this study the research concluded that Airbnb has a significant impact on hotels and rentals in terms of market share and competition.

Airbnb platform encourages customers to leave reviews and ratings on the basis of their experience with the property. It is the only mechanism to build in reputation and future business. It helps take the voice of the customers to the hosts and Airbnb platform but not sure what steps hosts take to improve their services in case of any complaints or suggestions by the customers. The Airbnb model runs on minimal regulatory controls in most locations. This study helps understand how reviews help fill the gap between hosts and the end users. With the help of natural language processing the study shaped the text analysis to quantitative analysis.

In this paper we have tried to answer the question: 1) How has Airbnb activity changed due to the COVID-19 pandemic, and what have been the customers’ preferences on rental activity? Adopting the comparative analysis of the data from the pre-COVID and
COVID era. The customers were more inclined in staying at a place wherein they can sublet entire apartment and house.

This work is exploratory, using one city as a case study to draft preliminary considerations of ways reviews when analyzed using Natural Language Processing can help explain customer preferences on rental activities. This research used text analytics as a core research method to analyze the success of the sharing economy model by analyzing customer reviews. Based on the analysis from the customers, they were very positive and optimistic about their experience with Airbnb. They were more focused on cleanliness and their expectations are of getting a clean and decorated place to stay with preference of an Entire Unit or home.

The findings of this study lead to suggested avenues for future research. To conduct this study, there were data sources from InsideAirbnb. However, these sources of data were not enough to accumulate answers to some of the questions like - What are the significant determinants of the count of customer reviews on Airbnb during the pandemic? and How do these determinants influence the count of customer reviews on Airbnb across different states and cities? Therefore there is a range of future research opportunities. For example, guidelines need to be established for the sharing economy service providers and receivers to ensure the sharing economy has good policies in place concerning tax policies, and how policies can help improve the situation that arises through these pandemic eras. Cancellations of booking had never happened before, and there was no clue how to handle such situations.
Trust in the sharing application is the pillar of the Airbnb business model. Trust maximises the likelihood of a successful booking. Host attributes and reviews in the study act as a reputation system and help create positive and significant effects. Sharing market has also been looked up with some safety issues. This has further lead to many doubts and questions as, if people are staying with entire place there is instep of many unknown people in these community. Neighbors are lacking trust and doubting identity of such guests. These findings provide additional insight into the potential risk or trust issues. People were uncomfortable with bringing strangers into their home. However at the time of the registration every customer has to provide an identity check, photo and social media integration to verify the identity of the customer. However it will be taken up as a future work to investigate and understand whether home sharing brings crime rate higher or not?

COVID-19 has also placed a need for social distancing, hygiene, and safety, but how we achieve this for the well-being of people needs for future exploration. The COVID-19 has also brought some positive outcomes; in this, they have also compelled various providers to improve their services and make it a more mature platform.

This study contributes to the society by addressing significant implications and presenting a robust methodology for comparing the impact of Airbnb across cities during COVID and Pre COVID era. This study provides the empirical evidence of textual analysis of the customer reviews and their importance in Airbnb data to understand customer behavior.
and scope of the sharing economy during this unprecedented times. This study also adds to the previous literature by exploring textual data and revealing insights using sentiment analysis, topic modeling and helping bring forward choice and voice of the customers during COVID times.
References


iGMS Inc. (2021, May 29). How to manage multiple airbnb accounts and should you have them? iGMS. https://www.igms.com/airbnb-multiple-accounts/


https://doi.org/10.1177/1847979018767776

https://doi.org/10.1016/j.ijhm.2016.03.005

https://doi.org/10.14453/aabfj.v15i1.4

https://www.austintexas.org/?gclid=CjwKCAjw9qiTBhBbEiwAp-GF0eZiV6_x4xRoV0w6BgDpEdFmOALgQ56jA7p8tnxlL8gSAkw4DT3g3xoC3W8QA0D_BwE

https://en.wikipedia.org/wiki/Los_Angeles
Vita

Gopika Malik received her bachelor’s degree from Delhi University, India in 2002 in Business Administration and Education. After completing her undergraduate she pursued Masters in Business Administration from Delhi University, India in 2004. Upon completing her degree she has taught as Statistics Instructor for ten years at Middle school level.

Her strong interest in statistics and data bring her back to school wherein she completed Associates degree in Business Analytics from Wake Technical Community College in 2018. She started pursuing her career as Data Analyst in Energy utility and sustainability sector. Along with her career she never stopped learning and continued studying Advanced Statistics and Data Management certificate course from NC State in the year 2020.

Gopika left her full time job to pursue her master’s degree in 2021 in applied data analytics at Appalachian State University. Her research besides this project involves studying about the University technology transfer wherein she contributed her efforts in analyzing how university policies, demographics, and location influence success. Her research has been accepted for presentation at the AMCIS conference to be held at Minneapolis, United States.