Politecnico di Torino

Engineering and Management

A.y. 2023/2024

Graduation session April 2024

Airbnb and their performances in Milan

Analysis of the performances of the Airbnb in Milan in relation with the distance from the “points of interests” pre, during and post Covid-19

Supervisors:

Elisabetta Raguseo

Francesco Luigi Milone

Candidate:

Andrea Calandra

Table of Contents

Abstract 1

Introduction 2

Literature review 3

Sharing economy 3

Definition *3*

An increasing phenomenon *6*

Differences of the Sharing economy *7*

Airbnb 9

History *9*

The business model *13*

A taxation problem *15*

COVID-19 impact on tourism. 17

Focus on Milan 22

Overview on Italian tourism sector *22*

A problem in Milan *26*

Numbers and areas of the listings *27*

Types of listings *30*

Literature review about the variables of distance 34

Research framework 37

Research questions 39

Data and Methodology 41

Airbnb's dataset 41

External datasets 43

Public parking *46*

Sport Facilities *47*

Degradation *48*

ZTL/Restricted Areas *49*

Metro stations *50*

Example of the creation of a dataset *52*

Descriptive analysis 55

General performances: 56

Performances variables 58

New variables 59

Seasonality analysis: 61

Performance related to the range of distance 64
Revenues *65*
Reservation days *67*
ADR *68*
RevPAN *69*
OCC *71*
Regression analysis 72
Univariate analysis 73
Revenues: *73*
Reservation days: *77*
ADR *79*
OCC *81*
RevPAN *83*
Multivariate Analysis 84
Revenues *86*
ADR *87*
OCC: *88*
RevPAN: *88*
Multivariate regression with SuperHost 90
Multivariate regression with LTR 93
Conclusions 95
Attachments 99
Multivariate regressions 99
Multivariate regressions with SuperHost 101
Multivariate regressions with long term rent 103
Acknowledgments 105
Bibliography 106

Abstract

In this master thesis will be studied how the distances between each Airbnb and its closest 'point of interest' affect the performances of the listings in the city of Milan.

As 'point of interest' have been identified the following parameters as the ones that can affect the choice of a customer in the booking: metro station, public parking, ZTL or pedestrian area, sport facility and area or building in degradation conditions.

The work has been conducted from June 2023 to March 2024 reviewing first the literature and then conducting the analysis using the software of Python, Excel, and STATA. In this script will be explored in depth the phenomenon in the city of Milan during the four-year period from 2019 to 2022.

Some of the trends that this research highlighted have been: the strong and positive relation between the nearness from the closest pedestrian area and the performance of an Airbnb; a positive relation between the listings and the closeness to a metro station and a public parking.

Introduction

Airbnb is part of a growing industry that, since its birth, has been able to gain more and more traction every year, slowing down only during the Covid-19 crisis in the tourism sector.

In the years, much research has been conducted to study the Airbnb's model from both customers and host point of view, other than to verify which variables affect the performances of the listings on the Airbnb's platform. These variables are affected by the years, the geographical position of the study and many other factors.

Only a few articles have studied the performances of Airbnb in relation to the distance from a point of interest, which can be the city center, the major square of a town, the coastline or the main tourist's attraction of a city. No research has been conducted to answer to the questions "are the performances of the Airbnb affected by the closeness to one of the following points of interest: a metro station, a public parking, a ZTL or a pedestrian area, a sport facility or an area or building in degradation conditions."

In the choice for an Airbnb, a customer may be influenced by some of these parameters of distance, because these can affect his or her stay in the following way: be closer to a metro station result in easier and faster travel in the city; be closer to a public parking may be a big incentive for customers arriving in the town with their own car; be closer to a ZTL or a pedestrian area may influence the choice of clients that want to enjoy the city walking; a customer that has to train during his/her stay may be affected by the distance to the closest sport facility and finally, a guest may want to stay away from the area in degradation conditions to avoid bad experiences.

Literature review

Sharing economy

Definition

Since Rachel Botsman and Roo Rogers coined the term “collaborative consumption” in their 2010 book What’s Mine Is Yours: The Rise of Collaborative Consumption, this economic model, known as the sharing economy or peer-to-peer economy, has gained widespread acceptance in business sectors including transportation (e.g. Uber), accommodation (e.g. Airbnb), food (e.g. EatWith), entertainment (e.g. WillCall) and even finance (e.g. LendingClub) (Wirtz et al., 2019), often facilitated through digital platforms or online marketplaces. The Sharing economy is an economic and social phenomenon that has gained traction in the last years. It entails a system that enables users to get a service or a good without owning it directly. In the last decade many researchers have gave a different definition to this growing phenomenon:

Sharing economy is a consumer granting each other temporary access to under-utilized physical assets (idle capacity) possibly for money (Frenken and Schor, 2017). consider that for a company to belong to the shared economy it must have the following characteristics: (i) the business should be focused on access to underutilized goods; (ii) consumers should benefit from access to goods and services; and (iii) the business should be supported by decentralized networks and marketplaces.

Sharing economy refers to a socio-economic system that involves a spectrum of activity based on maximizing the potential of our underused human and physical resources, from our skills to our things (Balaram B., 2016)

One of the characteristics of the shared economy is to make profitable assets that are little used and, in this way, to take advantage of market gaps (May et al., 2017).

Sharing economy refers to the use and access of shared physical or human resources or assets rather than the fact that there is no monetary exchange. The Sharing economy enables different forms of value exchange and is a hybrid economy. (Matofska, 2016)

The sharing economy has produced disintermediated industries: it allows people to transact directly by connecting them in unprecedented ways. (Caldieraro et al., 2018)

Collaborative consumption is an economic model of sharing, swapping, trading or renting products and services, enabling access over ownership. (Botsman R., 2010)

Sharing economy/Collaborative consumption/peer economy, individual participate in sharing activities by renting, lending, trading, bartering, or swapping goods, services, transportation solutions, space or money (Mohlmann, 2015)

Hamari et al. (2015) defines sharing economy as a peer-to-peer based activity of obtaining, giving or sharing the access of goods and services, coordinated through community-based online services.

Sharing economy is organized by the value in taking under-utilized assets and making them accessible online to a community, leading to a reduced need for ownership (Stephany A., 2015)

Sharing economy allows individuals or groups to make money from underused assets. (PricewaterhouseCoopers, 2015)

An economic system in which assets or services are shared between private individuals, either for free or for a fee, typically by means of the internet (Oxford Dictionary, 2015)

Sharing economy is the wide and varied range of old, revamped and new practices whose central characteristics are the ability to save or make money, provide a novel consumer experience, reduce ecological and carbon footprints and strengthen social ties. (Schor and Fitzmaurice, 2015)

Sharing economy provides value of the idle assets and makes them available to online community, that the sense of ownership is unnecessary. (Richardson (2015))

An umbrella term that, at the time of the recent global financial and economic crises, was given to an alternative economic and social model that has gained considerable attention. (Heinrichs, 2013)

Sharing economy describes a type of business model that builds on the sharing of resources between individuals through peer-to-peer services that allows customers to access goods and services when needed. (Böckmann, 2013)

Collaborative consumption is made by the activities of sharing, exchanging, and rental of resources without owning the goods. (Lessig, 2008)

A new socio-economic model that has taken off, thanks to the technological revolution, with the Internet connecting people through online platforms on which transactions involving goods and services can be conducted securely and transparently. (European Parliament, 2015)

Nevertheless, the key characteristic of sharing economy is to provide economic opportunity for the individuals to exchange their underutilized assets with strangers through intermediaries that match supply and demand in an efficient way and with the help of information technology (Petropoulos, 2017).

Therefore, we can combine all these definitions to point to the sharing economy as a peer-to-peer economic activity that provides comfortable lifestyle reachable to everyone thru the advancement of Internet and the platforms to it correlated.

The term sharing economy implies an increased utilization of assets with spare capacity (e.g. Benoit et al., 2017; Frenken and Schor, 2017; Hamari et al., 2016) and related reduced use of resources and ecological impact (e.g. Guttentag et al., 2018; Tussyadiah and Pesonen, 2018). This means, the sharing economy implicitly refers to the sharing of capacity constrained physical assets (e.g. cars, rooms, and bicycles) and the provision of performances and experiences that rely on shared assets and labor (e.g. a cooking or dining experience).

An increasing phenomenon

The Sharing economy phenomenon has gained interest since the last years 10’ in almost every sector, it has entered in our lives and right now is a part of our everyday lives, today’s actors are ones of the most known and profitable firms in the world. The first peer-to-peer sales website went online as early as mid-1990s when E-bay introduce online consumer bidding. From 2013/14 has gained more and more importance and more and more common showing an exponential growth and at the same time also the number of papers and research on this industry has grown.

The growth and impact of the sharing economy has changed different industries’ perception as it has generated billions of dollars for the past seven years. PricewaterhouseCoopers (PwC) has estimated its growth in the five major sectors – transportation (Grab, Uber, ZipCar), retail and consumer goods (Rent the Runway, ThredUp), accommodation (Airbnb, Couchsurfing, Homeaway), entertainment (Spotify), services (Task Rabbit) and finance (Lending Club). According to Roh (2016), by 2025 sharing economy could represent $335 billion in revenue worldwide or with an annual growth of 25-30% in the next five years. In fact, in 2011, Time Magazine has labeled sharing economy as “10 ideas that will change the world”. This huge phenomenon has gained a lot of traction, and more and more people are leveraging on it to make profit or to get access to resources, that can be identified in services of goods, at lower price, since the developed of this new system of collaborative consumption. More recent research suggests that global revenue generated by the sharing economy totaled $40.2bn in 2022 (Statista, 2022) and the projection is impressive for the next years. The two main horses of the sharing economy are the P2P accommodation and the P2P Transportation, that together generate almost 80% of the total revenues of the platforms leveraging on the sharing economy. (PwC, 2015)

As we can notice the difference from the prediction in the 2011 and the actual outcome evidence a large difference, this due to multiple factors that hit to the heart of the sharing economy principles. The most known one was certainly the COVID-19 pandemic that caused a lower growth and, in some circumstances, also a regress in the development of the activities. The transportation (e.g. Uber) and the accommodation (e.g. Airbnb) industries were the ones hit the most during the pandemic, and since they are the ones that bring more revenues to the sharing economy's sector, it had a slower growth than the predicted one.

Differences of the Sharing economy

Different types of business models aimed at profit entail different interactions between users, owners and public figures involved in the process. These different platforms attract different users depending on their main goals and how the technology has been developed. (Wirtz, Jochen, et al., 2019)

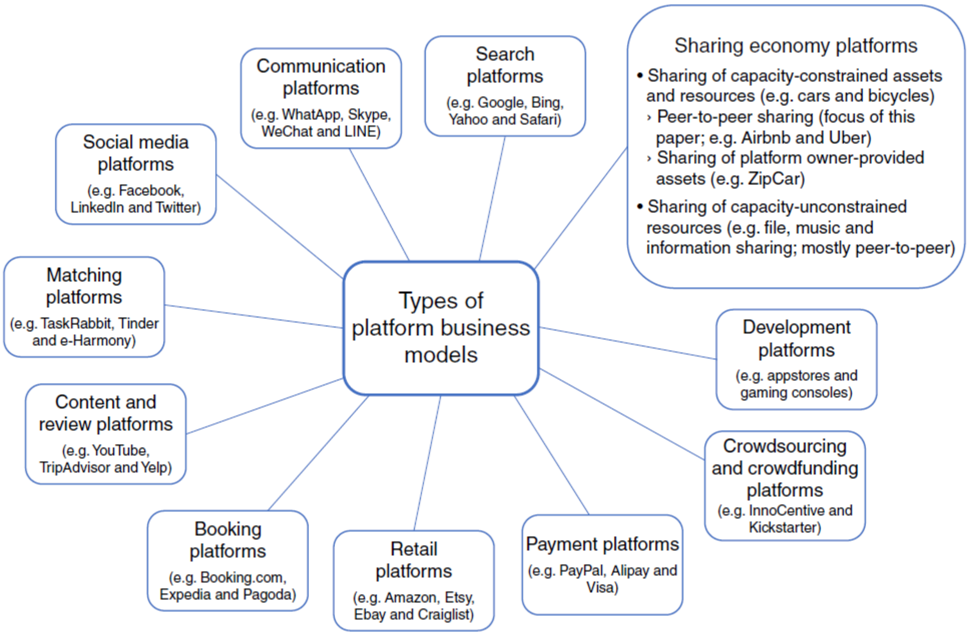


Figure 1

In *figure 1* Writz, Jochen et al. (2019) show the different types of platform business model, evidencing the sharing economy platform.

peer-to-peer platforms differ significantly from platforms that own the provided assets as they do not need to manage two-sided markets with issues such as signing up, rewarding, and managing service providers, or deal with their typically highly heterogeneous assets (e.g. Airbnb postings differ widely on key attributes).

The business models of the sharing economy rely on the peer-to-peer relationship, but they differ from the actors and the way they interact with each other. The following list briefly describe the possible models of peer-to-peer used by the different models:

1) B2C (business-to-consumer): firms companies use new technological possibilities to make their products available to customers.
2) B2B (business-to-business): used by sharing economy's platforms that provide services to companies rather than final users.
3) No-profit: the no-profit business model is used by platforms that have social and environmental objectives rather than profit ones.
4) C2C (consumer-to-consumer): this kind of model enables direct interaction between the final users. Ones of the most known and profitable firms used this model: Airbnb (which will be the focus of this study) and Uber.

Another division must be made for the provision of the resources. The axes taken into consideration in the study of Writz, Jochen et al. (2019), as shown in *figure 2*, are the ownership of the assets and the way these resources are provided. The focus of this study will be on Airbnb, indeed in a platform that relies on the concept of peer-to-peer provided resources and access-based, indeed it does not entail the transfer of the ownership. The distinction between platforms as Airbnb, Uber and Zipcar, WeWork relies on the ownership of inventory of assets, indeed the last two are able to offer to their customers their own assets by using an app that allow them to reserve and use these resources. On the other side peer-to-peer platforms link the two users and provide them with a platform on which they can interact to establish a connection. The

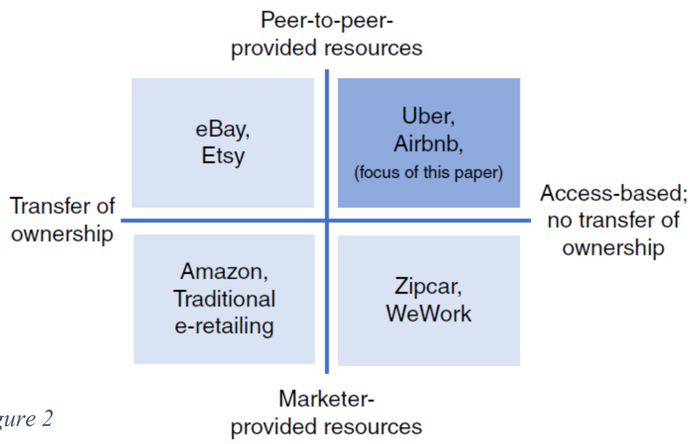


Figure 2

difference between eBay and Uber, instead, relies on the transfer of ownership. Traditional e-retail and eBay or Etsy usually are not considered as sharing economies because as assets are being sold rather than shared, therefore is possible to consider sharing economy only the platforms that provide access to services, resources or assets without implying the transfer of the ownership.

Airbnb

History

Accommodation (short-term letting) is one of the five crucial sectors (together with passenger transport, household services, professional and technical services and collaborative finance) of the sharing economy in the European Union, where it generated revenues of €3.6 billion in 2015 (European Commission, 2016).

Peer-to-peer (P2P) (Pizam, 2014) or consumer-to-consumer (C2C) (Täuscher and Laudien, 2017) accommodation sharing is a major trend shaping the tourism and hospitality industries (Guttentag, 2015; Guttentag and Smith, 2017). This phenomenon, also known as P2P property or accommodation rental, is not completely new, because room rentals have existed for a long time. However, the main difference between P2P accommodation sharing and traditional room rentals is the presence of digital platforms to connect renters and owners (Pizam, 2014). Such platforms are characterised by some distinctive features: they connect independent actors from the supply and the demand sides, and these actors directly interact; they provide an institutional and regulatory frame for transactions; and they do not substantially produce or trade goods or services themselves (Täuscher and Laudien, 2017). P2P accommodation sharing is commonly perceived as offering high value at a reasonable cost (Tussyadiah, 2016), thus making price an important factor for its success.

Among P2P short-rental companies, Airbnb is one of the most prominent (Edelman and Luca, 2014). It is an online platform where people can publish, discover and book accommodations worldwide. The idea was born from two flat mates: Brian Chesky e Joe Gebbia. The two students had troubles in paying the rent and at the moment a conference about industrial design was taking place in San Francisco and all the hotels were full. They decided to collect air mattresses and to offer a bed and breakfast to the designer that was looking for accommodation to participate to the conference. To sponsor their initiative, the two students created a web site called airbedandbreakfast.com. Afterwards, with the help of their old tenant Nathan Blecharczyk (with a degree in informatic), they were able to develop a more complex platform for the sharing of houses between users. The website was online for the democratic national convention held in Denver in 2008, same year of foundation of the company which has headquarters in San Francisco, California. One year after the company changed its name to "Airbnb" and in November 2010 the app was online for the first time ever. As the platform gained more and more customers were introduced also some reliefs like the $1M host guarantee or the "disaster relief tool" to face the damages of hurricane Sandy. As years go by, they introduced more and more functions to the platform, for example in November 2016 the experiences were added to the platform to differentiate the offer, in the same year also started the program Airbnb for business, that enable the platform to also reach travelers for work and not only for leisure. After the expansion has been developed all over the world, with different marketing campaigns (Aibiying is the name of their Chinese brand), and numbers of partnership, among which the nine-year one with the International Olympic Committee, in December 2020 the company went public by initial public offering (IPO), and this led to a valuation of 87 billion dollars. Some of the most important funds have been invested in the company since it went public, some examples are Vanguard Group, FMR LLC, Capital Research Global Investors and Blackrock Inc. Airbnb's "home away from home" concept has become a success story, and the future projections are bright for the company of San Francisco. Right now, the three founders still have an important role in the company, indeed Brian Chesky is the CEO, Joe Gebbia is the CPO and Nathan Blecharczyk holds the position of CSO.

The accommodations available on Airbnb range from a simple room to an apartment, a castle, or a villa, as years go by more and more houses have been listing on the platform, offering unusual accommodations, for example in the US right now is possible to spend the night in the Shrek's Swamp, or to stay in school bus adapted to become a cozy house. Airbnb is able to connect people with accommodations at any price in more than 100,000 cities and 220 countries; currently, more than 1,00,000,000 hosts have used the service. Airbnb's biggest night to date was 5 August 2017, when over 2.5 million people stayed at accommodations booked on the platform. In 2015, the role of Airbnb as an alternative to the hotel industry became official when it was named as the official alternative accommodation services supplier for the 2016 Rio Olympic Games (https://press.atairbnb.com). Several studies have analyzed the motivations for choosing Airbnb and have identified both economic and social benefits (Guttentag, 2015; Guttentag et al., 2017; Tussyadiah and Pesonen, 2016). This is consistent with Airbnb's mission 'to create a world where people can belong when they travel by being connected to local cultures and having unique travel experiences' (https://press.atairbnb.com).

The insufficient hotel rooms during peak period are one of the main reasons why Airbnb was formed. Airbnb became the biggest pioneer of sharing accommodation as it complements the on-going lack of supply in rooms. Its vision is to help the community to earn money in a flexible way and help strengthen local economies. Airbnb has transformed travel accommodation in a unique way enabling travelers to feel home away from home and exchange experiences with local community. It provides adventure and a unique home access, experience and be part of local community around the world. In addition, Airbnb provides innovative solutions to consumers by providing an online platform to reach its community. (Chua et al., 2019)

According to some observers (Guttentag and Smith, 2017; So et al., 2018), Airbnb has the potential to become a disruptive innovation (Christensen and Raynor, 2003), that is, to move from attracting low-end consumers to satisfying the demand of mainstream customers by acting as a substitute for, and not as a supplement of, hotels (Blal et al., 2018). Airbnb is the most prominent and its disruptive impact has driven the extremely rapid growth of the accommodation sharing phenomenon (Guttentag, 2015). Other studies claim that Airbnb is destined to remain a niche player and that the traditional hospitality companies do not consider it a significant disruptor (Varma et al., 2016). The debate surrounding whether Airbnb represents a true disruptive innovation or remains a niche player in the hospitality industry is complex and continues to evolve. It's essential to consider that both perspectives have valid points.

On one hand, proponents argue that Airbnb has the potential to become a disruptive innovation, following Clayton Christensen's model of moving from serving low-end customers to attracting mainstream customers. Airbnb's ability to offer unique, personalized accommodations at competitive prices has allowed it to capture a significant share of the market. This shift from niche to mainstream may indicate disruptive potential.Moreover, the rapid growth of the accommodation sharing phenomenon, largely driven by Airbnb, is undeniable. This surge in popularity and market influence supports the argument that Airbnb's impact is profound and transformational.

On the other hand, critics and some traditional hospitality companies do not view Airbnb as a significant disruptor. They argue that the hotel industry, despite the challenges posed by Airbnb, continues to thrive. Hotels offer a level of consistency, service, and reliability that many travelers value. Furthermore, regulatory challenges in various regions have hindered Airbnb's growth, suggesting that it may not replace hotels entirely.

The outcome of this debate is not black and white. It is possible that Airbnb's impact varies from market to market, and it may not entirely replace traditional hotels. The two can coexist, with Airbnb serving a specific niche of travelers seeking unique experiences while hotels continue to cater to mainstream travelers.

Airbnb development was inspired and contributed to the spread of the idea of the sharing economy or collaborative consumption (Dredge & Gyimóthy, 2015). Airbnb enables hosts to use their underutilized assets (apartments or rooms), while the platform employs sharing narrative to market its service as a novel and sustainable form of economic activity. Airbnb is not the only platform providing peer-to-peer accommodation in private homes and rooms. Table 1 lists major global platforms offering similar services. Airbnb appears to be one of the youngest (excluding Chinese competitors), yet the largest network of peer-to-peer hospitality. (Adamiak, Czesław, 2021)

Platform	Established	Headquarters	Rental model	Platform revenue	Platform size
HomeExchange	1992 HomeExchange, 2011 GuestToGuest	París, France	Guest points ear-ned for hosting for paying for stays	Membership fee paid by users	450k homes in 187 countries
Vrbo	1995 VRBO, 2005 HomeAway	Austin, USA	Paid	Subscription paid by hosts or commissions from bookings	Over 2M offers
Booking.com	1996	Amsterdam, Netherlands	Paid (mainly hotels, but also homes and apartments)	Commissions from bookings	About 2.4M offers of apartments, homes, homestays etc.
TripAdvisor Rentals	1999 HolidayLettings, 2004 VacationHome-Rentals 2005 Niumba 2007 FlipKey 2009 HouseTrip	Needham, USA	Paid	Subscription paid by hosts or commissions from bookings	More than 830k properties in 190 countries
Couchsurfing	2004	San Francisco, USA	Free	Membership fee in some countries; optional verification fee	14M members
Airbnb	2008	San Francisco, USA	Paid	Commissions from bookings	4M hosts, 5.6M offers in over 220 countries and regions
Wimdu	2011	Berlin, Germany	Paid	Commissions from bookings	Over 350k offers
Tujia	2011	Beijing, China	Paid (also hotels)	Commissions from bookings	Over 1M offers, mainly in China
Xiaozhu	2012	Beijing, China	Paid	Commissions from bookings	Over 500k offers, mainly in China

Table 1, Adamiak Czesław elaboration based on Hajibaba & Dolnicar, 2018; Feng, 2019; platform websites.

Although Airbnb is not the first mover in the industry, it was able to gain traction and fill an empty space in the market, becoming year after year more and more important in tourism worldwide.

The business model

Airbnb, indeed, is a two-side platform aimed "to create a world where anyone can belong anywhere, and we are focused on creating an end-to-end travel platform that will handle every part of your trip" (Airbnb.com). On the two sides of the platform there are: the part offering a service, the host, and the other one that is demanding a service, the traveler. The sharing experience led to the hosts to monetize unused spaces of their own apartment or second houses and led the travelers to save money finding better suitable places for different wallets, offering cheaper rooms than traditional hotels or bed and breakfast accommodations. Other than the monetary factor, another aspect that pushed customers to use the platform was the component of the experiences, indeed splitting the house with a local can increase your real experience immersing you in the local tradition and customs. The mission of the Californian's company was to lead travelers to feel like they were at their own house even if they were far away, here the concept enhanced year by year of "belong anywhere".

An aspect that was and nowadays is crucial for the platform is the system of reviews that led the users to express their own opinion about the service offered by the host. Year by year the system of reviews increased the details and led to a better judgement splitting category by category, right now after renting a house or a room on the platform, is possible to give feedback about 6 different categories: cleanliness, accuracy, check-in, communication, location and value. Matching the score of each category is possible to find the overall grade of the accommodation and by making a mean between all the reviews a mark is given to the listed house/room. On the other side also, the hosts have the possibility to evaluate the behavior of the guests, indicating also for them a review and a number of stars, on a scale of 5. This led to better assurance on the hosts' side to identify the person who they are renting the house from.

In order to attract and keep the customers, the platform studied a system of discounts and offers that grants the public better involvement and incentives to continue to use the platform. Airbnb started to offer promotional codes, credits and offers to create a loyal community and attract each day more customers. The Californian colossus focused its attention also on the hosts by trying to convert them into travelers with a system of credit that kept the money inside the platform giving the possibility to use the credit for the next accommodation.

Airbnb was able to create a loyal community and to create a strong and reliable brand, known all over the world. The combination of these actions led the platform to scale up in few times and nowadays counts more than 6 million of global active listings and more than 4 million of hosts worldwide, arriving at a current valuation of 113 billion dollars.

But how does Airbnb make money? The platform has structured a system of % fee that it takes from hosts and travelers to give them the interaction, the platform, a secure method of payments, customer service and many other functions. Airbnb recognizes revenues at the time of check-in to account for cancellations. Guests typically must pay a service fee based on the value of the accommodation; it is around 14% of the value of the booking. On the other side hosts are charged a fee of 3% of the booking subtotal (the nightly rate plus the cleaning fee), the fee rises to 20% for the experiences. Airbnb, by being listed, provide a public report of the company stating among a lot of information also the costs of the company. In 2021 the costs as a percentage of the revenues were: (ANNUAL REPORT PURSUANT TO SECTION 13 OR 15(d) OF THE SECURITIES EXCHANGE ACT OF 1934)

- cost of revenues = 19%
- Operation and support = 14%
- Product development = 23.5%
- Sales and marketing = 19.5%
- Others = 16%

A taxation problem

"Hosts should understand how the laws work in their respective cities. some cities have laws that restrict their ability to host paying guests for short periods. These laws are often part of a city's zoning or administrative codes. In many cities, hosts must register, get a permit, or obtain a license before listing a property or accepting guests. Certain types of short-term bookings may be prohibited altogether. Local governments vary greatly in how they enforce these laws. penalties may include fines or other enforcement. hosts should review local laws before listing a space on Airbnb" (Terms of Service (Airbnb June 30, 2014), online at http://www.airbnb.com/terms).

A crucial aspect for Airbnb has been the taxation and regulation problems. By having scaled up in a short time, governments of all nations have faced problems in understanding how to regulate the sector and how to tax the profits of hosts. The rapidity, that distinguished Airbnb, led to a quick globalization of the phenomenon, but different states applied different regulations and taxation to the accommodation industry. Government have also faced difficulties in identify the sector where Airbnb can be collocated, by being a new and innovative platform different from hotel industry and bed and breakfast, a new form of taxation had to be applied for the profits and in many countries this regulation has been unclear for a lot of time.

Considering the cost-efficient nature of the growing peer-to-peer lodging sector, there have been two main taxation proposals for platforms like Airbnb (2019b). The first idea involves implementing a service fee that would be applied to both guests and hosts. Airbnb currently earns between 9% to 12% from guests for each booking, with the exact percentage varying depending on the length of the stay, and a 3% fee from hosts. Alternatively, governments also have the option of introducing a local tax that is focused on the services provided by such platforms.

In the Italian framework the taxation in Airbnb's hosts profits has changed multiple times, indeed since 2017 a flat tax of 21% was set on short-term rental income, instead of a taxation based on the personal income brackets, qualifying a short-term rent as a rent that last for less than 30 days. Although was not specified that this taxation was launched for facing Airbnb growth, it was extended to all the short-term rental income. In 2020 the regulation in Italy has changed, indeed from now on a landlord who short rent more than four residential units will automatically be considered as a business and will be subjected to business accounting and tax rules. The new law legislation that is currently discussed in Italian's parliament will introduce more strict regulations in the short-term rent framework of the country. The new bill will introduce a minimum of 2 nights for renting a house on Airbnb, the reduction from four to two of the number of listings owned before becoming taxed as a business, and last, also the apartments have now to meet the requirements of safety requested to hotels like a fire extinguishing system, a CO detecting system and so on.

Every country can apply a different taxation rule to the profit of hosts that make profit in their country, indeed some difficulties have risen also for hosts who managed multiple listings in different regions. The platform, in 2022, has registered a collection of more than 7 billion US dollars of taxes and this number has growing predictions as taxation becomes harsher and number of listings higher.

COVID-19 impact on tourism.

In the last two decades several major disruptions have affected international tourism, such as terrorist attacks (e.g. in New York, 2001; in Bali, 2002), the global economic crisis in 2008–2009, the eruption of the volcano Eyjafjallajokull in 2010 or the 2004 tsunami in South Asia (Hall, C.M. 2010; Lim, J. and Won, D. 2020). The most important disease outbreaks that had effects on tourism and hospitality industry were the bovine spongiform encephalopathy ("mad cow disease") in 2002–2003, the severe acute respiratory syndrome (SARS-CoV) in 2003, the avian flu in 2004, swine flu in 2009 (H1N1), Middle East respiratory syndrome-related coronavirus in 2012 (MERS-CoV) and the Ebola outbreak in 2014. All of them have been analyzed and discussed intensely in the academic literature, focusing on the effects of the diseases, presenting post-crisis management perspectives and highlighting the importance of precaution (Sharpley, R. and Craven, B. 2001; Baxter, E. and Bowen, D. 2004; Henderson, J.C. and Ng, A. 2004; Gu, H. and Wall, G. 2006; McAleer, M. et al. 2010; Wu, E.H.C. et al. 2010; Rassy, D. and Smith, R.D. 2013).

Certainly, the most recent and disruptive case has been the Covid-19 pandemic, started in 2020. Society had to face a huge economic crisis, other than social and public disease. In this context one of the most affected industries has been the tourism sector due to the restrictions imposed to slow down the diffusion of the virus. The World Tourism Organization estimates that revenues from tourism could fall by $910 billion to $1.2 trillion in 2020 and between 100 to 120 million direct tourism jobs are at risk (UNWTO, 2020). According to Eurocontrol (2020), flight numbers decreased by 88% in April relative to last year. Tourism also came to an almost complete halt in Europe: occupancy in hotels decreased by 85% in April year-to-year, despite significant drops in hotel prices (STR, 2020). Similarly, to the traditional hotel industry, the short-term home rental market was also significantly affected, with occupancy rates of Airbnb falling strongly below the levels of 2019 (AirDNA, 2020). The UNWTO (2021) recently reported that international tourist arrivals were down 83% in the first quarter of 2021 as widespread travel restrictions remained in place across the world. Platform providers such as Airbnb and Uber were no exception. Thousands of people lost their jobs, the value of sharing firms plummeted and many service providers had no other option but to stop working (Hossain, 2021).

Companies in all sectors had to find solutions to adapt to the situation, continue to operate and not to fail. Public interventions, digitalization and working from home were crucial in this period to permit the firms to stay alive. In particular, in the tourism sector flexibility and agility emerge as two crucial factors for the tourism industry to adapt and mitigate the effects of travel restrictions (Ugur & Albiyik, 2020). Airbnb supply is more flexible than hotel's one and led to a better reaction to the fluctuation of the demand, indeed hotels have fixed capacity that cannot be expanded during periods of high demand. Zervas et al. (2017) also found that the supply of Airbnb listings increases during high demand periods.

Although the better flexibility, Airbnb suffered a lot the crisis tied to the covid-19 pandemic, indeed the CEO of Airbnb, Brian Chesky, qualified these times as "the most painful crisis of our lives" (Airbnb, 2020). Indeed, the Californian giant has to fired 25% of its workforce in 2020, this added to the cut to expenses of marketing and advertising led the company to save more than a billion dollars in the same year. On the other side many superhosts, that have as only sources of income the rent using Airbnb, seek for economic help from Airbnb. To face this problem the company from San Francisco was able to create a fund of 250 million dollars for the hosts that have suffered the most from cancellation due to the pandemic, another second fund of 10 million dollars was instituted for superhosts and hosts of experiences in troubles.

During the period of the pandemic, value creation changed in at least three ways. First, new protocols and certifications were introduced to ensure safe operations in the hospitality industry. Second, hospitality businesses reduced their capacity and increased the quality of their services. Third, employment in the hospitality industry changed drastically due to devastating effects of the COVID-19 crisis. Value capture allows the firm's value proposition to be translated into revenue (Clauss, 2016). Hospitality business models experienced three main changes during COVID-19. First, greater promotion of domestic rather than international tourism. Second, implementation of flexible cancellation and changes to tourism arrangements. Third, offer better quality and more individualized services.

Airbnb suggested to hosts to adopt flexible cancellation to increase the trust in their customers, and to lead available the calendar for longer stay with a discount to better predict the variability of the demand of the market. Many hotels (the Hilton Hotel, the Marriott Hotel, etc.) and airlines (Air France, British Airways, Eurowings, Lufthansa, etc.) have updated their change and cancellation policies to make them more flexible to respond to the COVID-19 crisis. The outbreak of COVID-19 resulted in massive flight cancellations worldwide. In order to support passengers, airlines are offering refunds, vouchers, or ticket exchanges. Most travel – including originally non-refundable bookings - that was suspended in the first months of the COVID-19 crisis was cancelled or refunded. This policy has remained in place; in the case of the need to cancel a trip, a refund is provided, and many airlines, hotels, and travel agencies are not imposing change or cancellation fees. (Youssef, A. B., Redzepagic, S., & Zeqiri, A., 2022)

During the covid period people were interested in the safety of the apartment they were going to book, this takes into consideration the possibility to make a "self-check-in" that allow not to have a physical interaction with the owner of the Airbnb, the type of the house (if the customer is renting an entire house or only a room in a shared house) and finally the cleaning of the flat. Regardless of the last one, in the case of Airbnb-type accommodation, hosts can register for the "Cleaning Protocol" certification which includes training in how to prepare accommodation for guests. The training includes provision of information on preventing COVID-19 infections such as use of face masks and gloves by hosts and cleaners, use of appropriate disinfectants and cleaning materials, etc. (Youssef, A. B., Redzepagic, S., & Zeqiri, A., 2022) All these precautional interventions were aimed to infuse more trust in the customers and allow them to travel again under safe conditions. Since the pandemic crisis more and more studies on the tourism sector have been made. In particular Airbnb, which has been growing rapidly since its creation, has suffered a lot from the restriction imposed by each country. Dolnicar and Zare (2020) described the pandemic's impact on the sharing economy as "disrupting the disruptor" and suggested that the trading of space on Airbnb and similar platforms would recover but not to pre-pandemic levels. They further predicted that the proportion of investor-hosted Airbnb listings would drop. Similarly, Bresciani et al. (2021) found that, compared with hotel rooms and full flats, travelers are now more reluctant to book shared flats on Airbnb out of a high need for physical distance. In the fig. 3 Bugalski, Łukasz. (2020) shows in his study the impressive growth in the number of Airbnb actively listed after 2014 and how with the pandemic this number fall suddenly.

Not only Airbnb was affected by this disruptive event, but all the hotel industry had seen the demand drop tragically compared to the years before. Alongside the lockdown impact, the availability of hotel rooms has dropped 40% compared to 2019. The demand of customers has decrease in February, in March arrived to a -50% in almost every Italian city and in April touched peaks of -90%. The situation for hotels was critical, indeed they registered a huge decrease in the revenues that didn't stop during the summer, the period when they typically registered most of their cash flow. Even if most of the hotels decided to open again to tourism, after the restrictions of the government, in summer the demand didn't meet the offer of the hotels, pointing out that the situation wasn't going to improve. (Gyody, K., 2021)

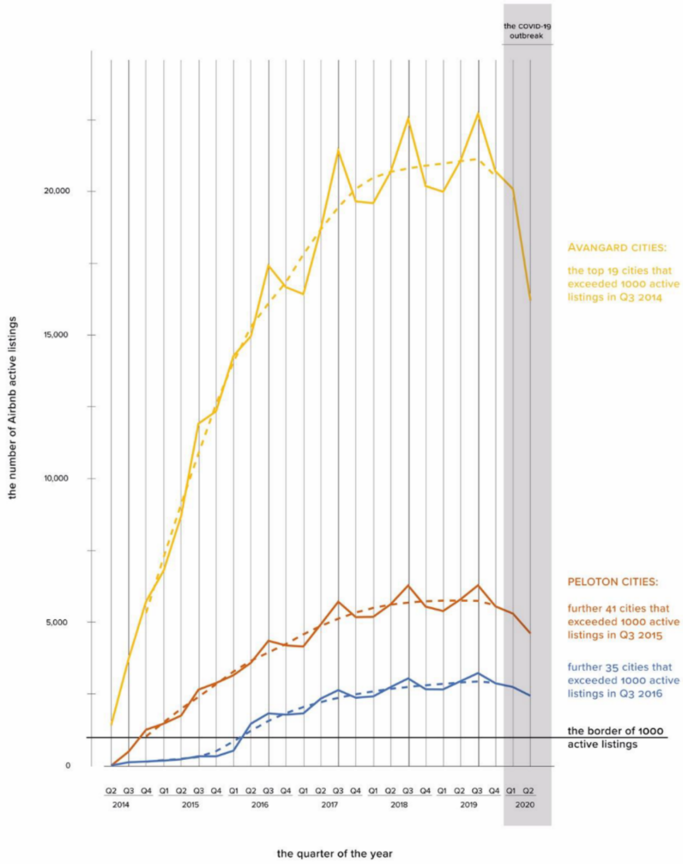


Figure 3

Figure 4 shows the dramatic percentage change in the demand for the hotel industry making a comparison with the same month of last year. As shown the peak is in April where a lockdown was imposed in almost all the Europe area causing a global stop in the travel and consequently in the revenues of hotels, but also restaurants and many other areas were affected as much as the travel sector during this dark period.

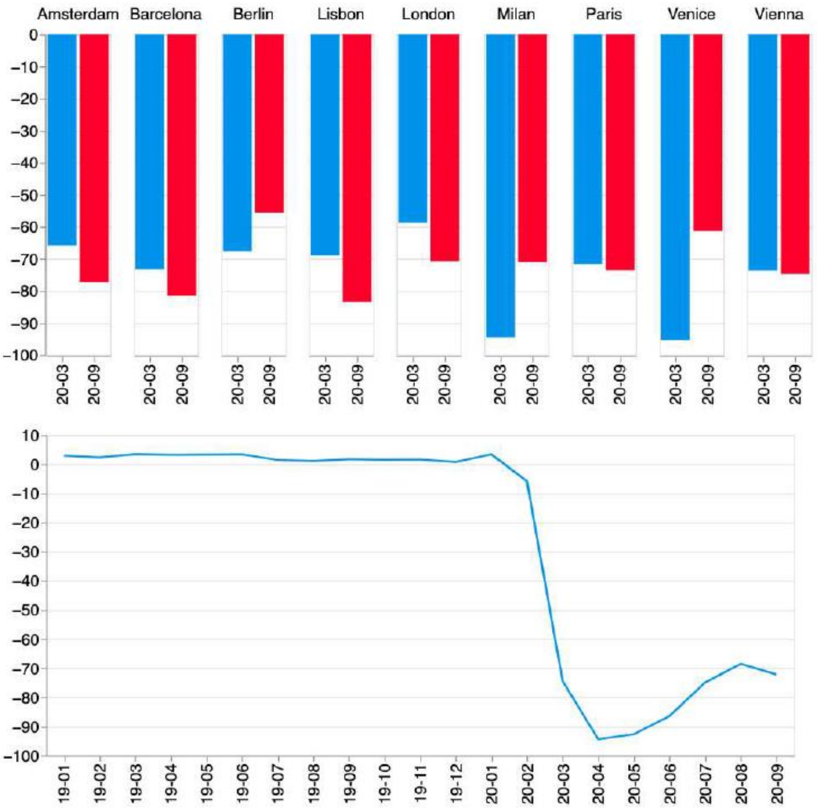


Figure 4, Hotel demand, percentage change from same month previous year.

Focus on Milan

Overview on Italian tourism sector

Tourism in Italy is a sector of remarkable global significance and plays a pivotal role in the national economy. The Italian Peninsula, located at the heart of the Mediterranean, is renowned for its rich cultural history, artistic and architectural heritage, extraordinary landscape diversity, and, not least, its celebrated cuisine. These elements, combined with a favorable climate and a wide range of tourist activities, make Italy one of the world's most sought-after destinations for travelers from around the globe.

Italy, an icon of international tourism, was ranked fifth by the World Tourism Organization (UNWTO) before the COVID-19 outbreak in terms of arrivals and sixth for receipts (UNWTO, 2020). As the third-largest national tourism destination in terms of overnights after Rome (30.9 million overnights) and Venice (12.9 million), Milan (12.4 million) is an interesting case for exploring listing determinants (ISTAT, 2020, p. 628). While the other large Italian cities (Rome, Florence, and Venice) are mainly focused on the leisure market segment, Milan is the national economic capital and the headquarters of the Italian stock market. The city has many attractions, including business firms and investors, the second-largest European trade fair center, and numerous points of interest, such as the Duomo and Leonardo da Vinci's The Last Supper. Other important attractions include the city's nightlife and the design and fashion industry. The city hosted the Expo in 2015, a mega-event that increased hotel revenues and reduced demand seasonality. The Milan Expo has partially changed the destination's image and increased the leisure market segment.

The focus of this analysis is on the city of Milan. Nestled in the heart of northern Italy, is a city renowned for its rich history, culture, and innovation. It's often hailed as the country's premier financial, cultural, and fashion hub, and this reputation draws visitors from all corners of the world for various reasons. In your thesis on tourism and Airbnb in Milan, you're diving into a crucial aspect of hospitality in this captivating city.

Milan's roots run deep, tracing back to Roman times when it was known as "Mediolanum." This historical significance is tangible in the city's architecture, with iconic landmarks like the Duomo and majestic historic palaces gracing the city center. But Milan isn't merely about its past; it's also a forward-thinking city, a hub of industry and innovation in Italy. Tourism in Milan is on a steady incline, fueled by its dynamic cultural scene, world-class museums, design and fashion events, remarkable art, and exceptional cuisine. Airbnb has played a significant role in Milan's hospitality sector, offering a diverse range of accommodation options, from traditional homestays to modern facilities right in the heart of the city. As you delve into the relationship between tourism and Airbnb, your thesis can shed light on how this form of lodging has influenced Milan's tourism sector, economy, and local community. Milan is a unique city where tradition and innovation harmoniously coexist, making it an enthralling destination for travelers from across the globe. Your research on Airbnb in Milan contributes to a deeper understanding of how emerging hospitality trends impact the visitor experience and the local economy.

Milan is the second-largest city by listings in Italy, with 17,659 registered properties in 2019, 60% of which were single listings. Around 73% of listings are for entire flats or houses, with 23.6% concentrated within the central city area (Cerchia dei Bastioni) (Inside Airbnb, 2020). The presence of Airbnb in Milan mitigated the effects of the financial crisis on the housing market of the city. Regional and local authorities viewed Airbnb as a win-win solution to support 'new public and private sharing services and products in a more sustainable and inclusive economy' (Aguilera et al., 2019, p. 12) and to reposition the city of Milan as a leading tourist destination. However, this has led to early signs of tourism gentrification. Italy, and in particular Milan, plays a key role in the European tourism sector, due to the high number of tourists for leisure and the high amount of business ones. Airbnb, focus of this study, has developed a lot in the peninsula, indeed in his study, Adamiak Czesław. (2021) shows how a relatively small country as Italy has a huge amount of Airbnb with at least one review positioning above China which has a population 24 times bigger (*figure 5*). Only two countries have a higher amount of Airbnb listed with at least one review: France, having as main attractions Paris, Disneyland and the French riviera; United States, where the Company was born in 2008 and having a huge number of tourists especially in the two costs. Even in the country where Airbnb was born there are continuous changes in the policies applied by each state, indeed the last important fact happened in New York, where

Airbnb hists cannot list no more an entire flat on the platform, but the owner must be inside and share the flat with the customer.

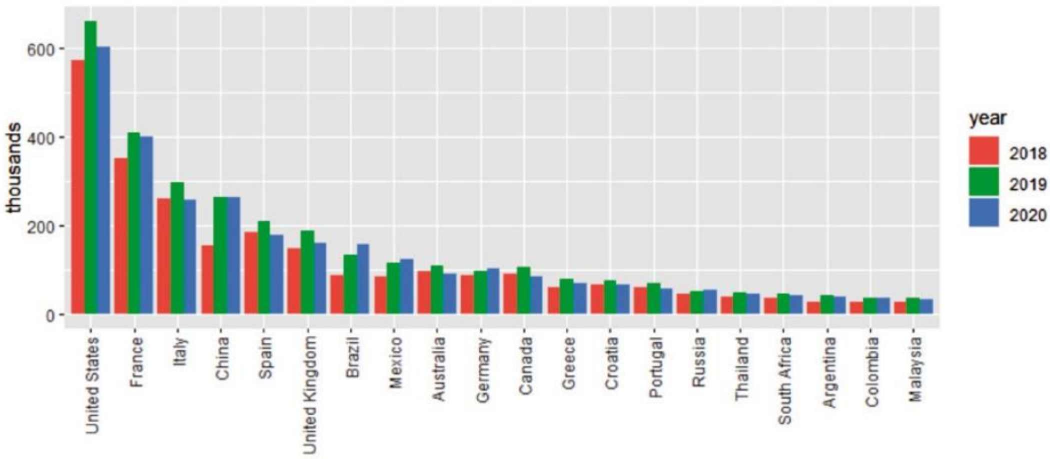


Figure 5

Going more in deep analyzing city by city Adamiak Czesław. (2021) showed the importance of Milan in the Airbnb system, by being at the 13th position in the world for amount of Airbnb listed with at least one review. *In figure 6* is represented the number of Airbnb listed in each city having at least one review in 2021. In the Italian framework, only Rome has a higher amount than Milan, overall, Paris and London dominate the scene followed by the most important cities in the States.

Most of the research about the tourism in Italy during the covid period, and in particular on Airbnb in the city of Milan, focused their research on the number of Airbnb listed before, during and after the Covid-19 pandemic.

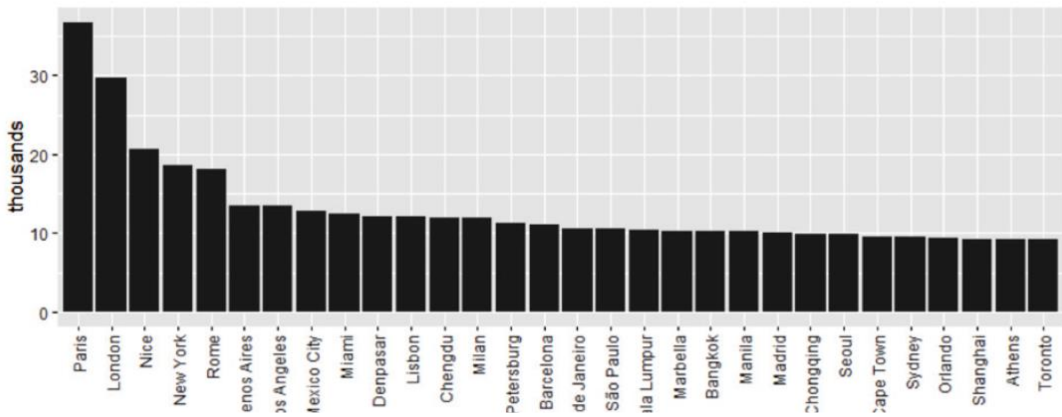


Figure 6

In all the world the Covid disease led to a decrease in tourism and therefore short-term rental bookings were affected to the same extent. A way to obtain this data can be done through the reviews, indeed Airbnb estimates that these are left approximately by 70% of the guests. Romano Antonello (2021), in *figure 7*, showed in his research the percentage change on short-term rental booking in all Europe, making a division in each province. The result illustrate how different cities were affected in different extent, from Italy, where the first European case of Covid-19 was discovered, that suffered a huge number of cancellations, to Luxemburg that registered even an increase in the bookings.

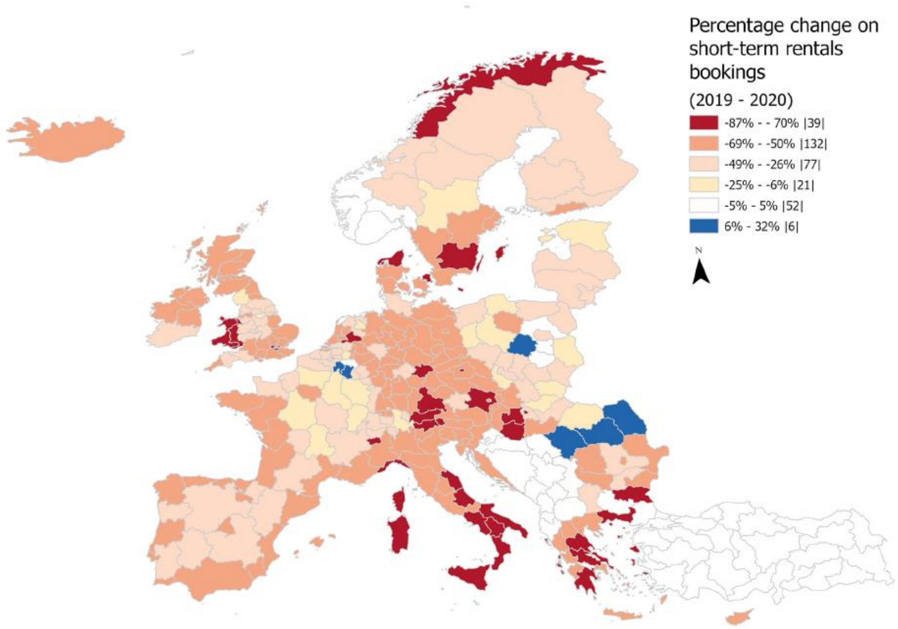


Figure 7

Analyzing only the city of Milan, is it possible to notice that the phenomenon has taken on a growing trend in the last ten years before the disruption of Covid-19, following the areas and times of the most important events in the city, an example is the Expo of 2015.

A problem in Milan

The emergence of Airbnb in Milan has given rise to a host of complex problems that have had significant repercussions on the city's housing market. What was originally conceived as a platform to help homeowners make some extra income by renting out their properties for short stays has evolved into a situation with far-reaching consequences.

One of the most evident impacts of Airbnb's growth is the surge in short-term rentals, particularly in the heart of the city. Tourists, whether they're visiting for leisure or business, are drawn to these central locations, making them highly lucrative for property hosts. This popularity has led to a noticeable scarcity of properties available for long-term rentals, which has, in turn, driven up the demand and, consequently, rental prices.

The scarcity of long-term rental options in Milan is particularly challenging for families and university students. These groups often require stable and affordable housing for extended periods. However, the shortage of long-term rental properties has made it increasingly challenging for them to secure suitable accommodation. This has left many students with limited options, such as moving farther from the city center or sharing accommodations with multiple roommates. These compromises can have a direct impact on their quality of life and, potentially, their ability to concentrate on their studies.

In response to these housing challenges, some university students have resorted to symbolic actions, including sleeping in tents outside government offices and university buildings. These acts of protest are a clear demonstration of their frustration and a call for immediate action from the authorities to address the housing crisis.

These trends in Milan are not unique to the city. Research and studies have consistently shown a strong link between the growth of short-term rentals and the increase in long-term rental prices in various urban areas. In the first

channel, home-sharing increases rental rates by inducing some landlords to switch from supplying the market for long-term rentals to supplying the market for short-term rentals. The increase in rental rates through this channel is then capitalized into house prices. In the second channel, home-sharing increases house prices directly by enabling homeowners to generate income from excess housing capacity. This raises the value of owning relative to renting, and therefore, increases the price-to-rent ratio directly (Barron, K., Kung, E., & Proserpio, D., 2019). The sharing economy, which includes platforms like Airbnb, has garnered significant academic attention, with researchers delving into its effects on housing markets, urban development, and local communities.

To counter these housing challenges, the Milanese government and local authorities have begun implementing policies and regulations designed to address the housing crisis. These measures often involve placing restrictions on short-term rentals and offering incentives for long-term renting. However, the effectiveness of these policies can vary, and addressing this multifaceted issue requires a holistic approach that considers the needs of all residents and aims to establish a sustainable and equitable housing market.

When considering how local governments should handle the growing trend of home-sharing, Barron, K., Kung, E., & Proserpio, D. (2019) believe that regulations should primarily focus on preventing a significant shift of housing from long-term rentals to short-term rentals. The aim is to maintain a healthy housing market while still encouraging homeowners to engage in home-sharing. One potential regulatory approach involves imposing occupancy taxes only on those who rent out their entire homes for extended durations or requiring proof of owner-occupancy to be exempt from these taxes.

Numbers and areas of the listings

Anselmi, Guido, and Veronica Conte (2021) in their study, have analyzed the changes of Airbnb due to the pandemic in Airbnb. Beginning from the trend (*fig. 8*) of Airbnb in

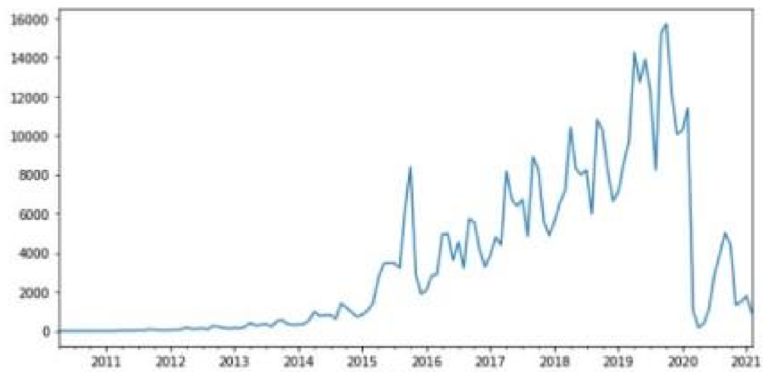


Figure 8

the city from its birth to 2021. They notice a peak in correspondence of the Expo and an overall growth until the covid breakdown.

The analysis pre-Covid, taking in consideration the year 2019, evidenced that in the city of Milan:

- Sixty-seven percent of the outstanding ads are entire property rather than private rooms and/or shared ones.
- Forty-two percent of review represent "*multi-listing*", residential assets held by owners that manage more than one single property.

Regardless of the territorial distribution of Airbnb, in 2019 the offer focalized in the so called "circonvallazione esterna" of Milan, in particular in the neighboring of Buenos Aires – Venezia, Centrale, Garibaldi Repubblica, Isola, Breba and Sarpi. This distribution is strongly affected by seasonality and the main events of Milan; indeed, the demand is higher in areas tied to the style industries during the "Milano fashion week". In *fig. 8* an analysis of Anselmi, Guido, and Veronica Conte (2021) shows the number of Airbnb in the city of Milan pre-Covid, with a partition for neighborhoods. As have been done in other studies, also here it is used the number of reviews left by the customers, hypnotizing that approximately 70% of the customers leave a review of the flat rented.

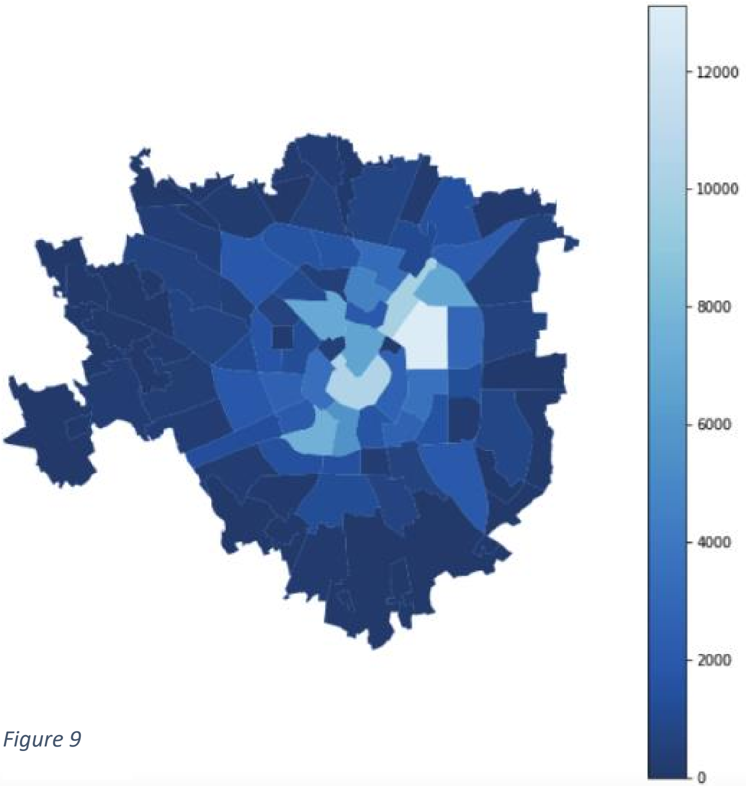


Figure 9

As easily understood, 2020 and the pandemic led to a huge reduction in the tourists flows (-95% registered in March 2020) and the booking in the short-term rental (Comune di Milano 2020). This led to a general reduction in the stay in flats on the short-term rental market between 2019 and 2020 as shown in *fig. 10.*

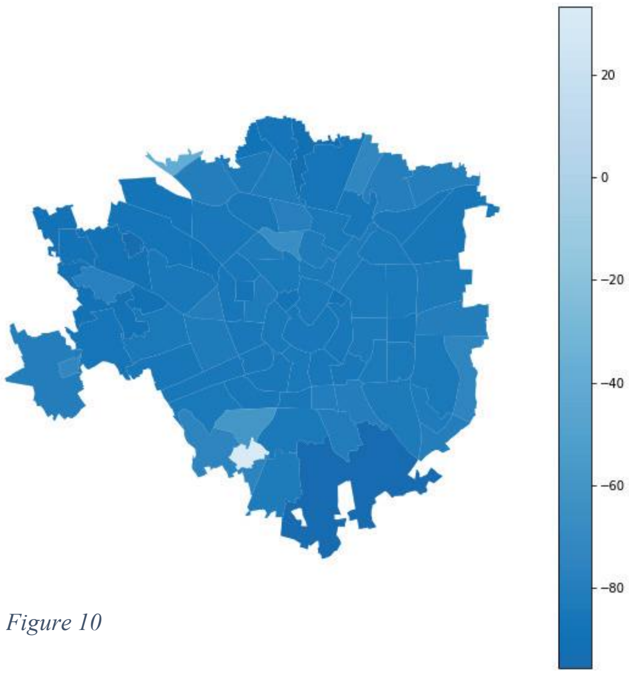


Figure 10

In 2021, the city of Milan counts 17650 listings, of which 72.5% are entire properties, 25.5% are private rooms and the remaining 2% are shared rooms. 39.7% of hosts hold multiple listings, a trend that is constant with respect to the years pre-pandemic. The territorial distribution doesn't seem to have suffered many changes. (Anselmi, Guido, and Veronica Conte, 2021)

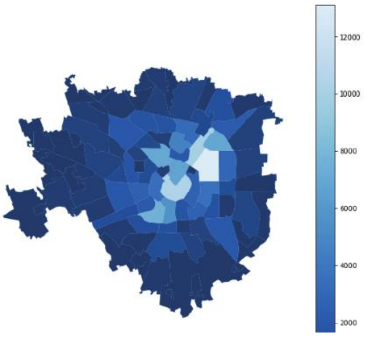


Figure 9, 2019

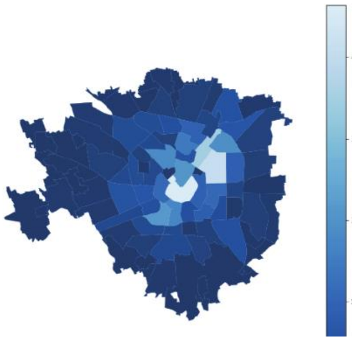


Figure 12, 2020

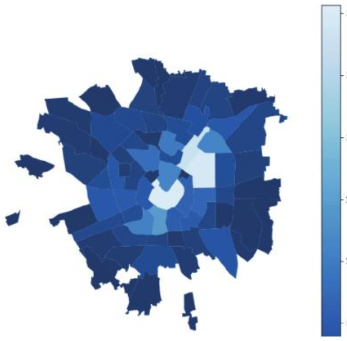


Figure 13, 2021

Types of listings

Other kinds of analysis focus their attention on the type of accommodation and on the number of listings by each host. Here is the division between the "common" host and the superhost. A typical normal host often starts with a single listing, primarily motivated by the prospect of earning extra income or sharing their extra space. Their commitment to Airbnb may be more casual, as hosting remains a side activity alongside their regular job or daily responsibilities. Superhosts, on the other hand, demonstrate a higher level of commitment. They often see Airbnb hosting as a primary or full-time occupation, and their motivation may extend beyond financial gain. Reputation, providing quality experiences, and building a brand are significant drivers for superhosts. The distinction between normal Airbnb hosts with one listing and superhosts managing multiple listings extends far beyond the number of properties they oversee. It encompasses varying motivations, commitments, guest interactions, approaches to quality, and their overall influence on the Airbnb ecosystem. While both categories contribute to the platform's appeal, the rise of superhosts with their focus on consistency and scalability has transformed Airbnb into a formidable player in the global accommodation industry. Understanding these differences is essential for travelers and prospective hosts, as it informs the expectations and experiences within the diverse world of Airbnb.

An analysis of Ruggero Sainaghi and Rodolfo Baggio, on the concept described above, showed the numbers pre-covid (2015-2018) about the number of listings of each host in the city of Milan, which is also the same one taken into consideration to conduct our analysis.

Figure 14 reports the host and listing distribution, which shows a clear power-law pattern. The graph illustrates the long tail with a strong concentration on the right side of the horizontal axis. Essentially, a handful of hosts manage a wide number of listings.

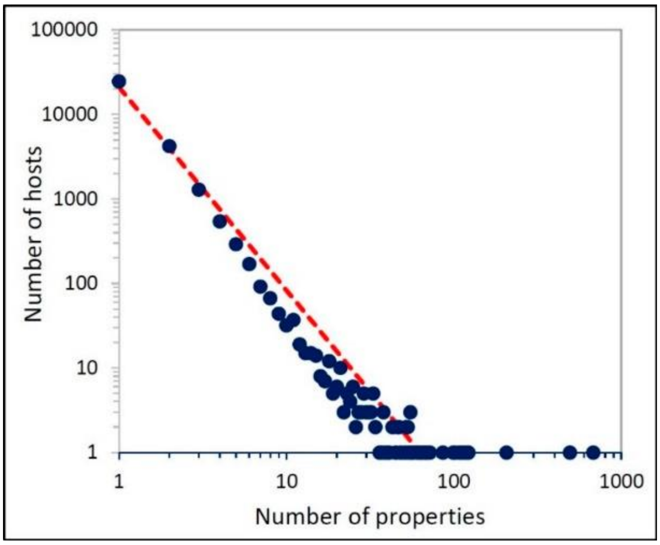


Figure 14

The study then analyzed the numbers by making a cauterization splitting the hosts who have 1, 2, 3, between 4 and 10, or more than 10 listings. 35% of all the revenues are made by the hosts having listed only one house/room and most of the hosts (78%) have only one item listed on the platform. They overall represent 48% of all the listing present on the platform. Furthermore, the study evidenced how before the disruption of Covid-19 the revenues were distributed as follow:

Clusters	Clusters'weight		
	Rex%	List%	Host%
P1	35%	48%	78%
P2	13%	17%	13%
P3	7%	8%	4%
P10	15%	13%	4%
P>10	29%	15%	1%
Pall	100%	100%	100%

Table 2

As the data show: the revenues are in the hands of hosts with more than 10 units listed (29%) and those who have only one listing (35%). The two categories were affected in different way by the pandemic:

Superhosts, who often rely heavily on consistent bookings across multiple listings, faced uncertainties as travel restrictions and lockdowns disrupted travel plans. This instability affected their income, and with multiple listings to manage, the financial impact could be significant.

Common hosts, who typically have one listing, also grappled with booking uncertainties. With fewer properties, the economic impact might be less severe, but the income loss was still keenly felt.

Other than the uncertainty regarding the bookings of customers, they had to face difficulties like cancellation and refunds, changes in the safety measure and government regulations. In summary, the challenges faced by superhosts and common hosts during the COVID-19 pandemic were marked by financial uncertainties, cancellations, health and safety concerns, adapting to shifting guest demands, and navigating a complex web of government regulations. While superhosts, with their multiple listings, had to manage larger-scale operations, common hosts faced many similar issues, albeit on a smaller scale. The pandemic underscored the resilience and adaptability of Airbnb hosts as they collectively navigated the stormy waters of the travel industry during one of its most challenging periods.

Moreover, in their article "Airbnb Host Scaling, Seasonal Patterns, and Competition", the two Italians analyzed, in the city of Milan, important divisions that can be done in study of the listings. These spaces include the entire home (EH) where the host is not at home, or private room (PR) and shared room (SR).

Types	n	%
EH	12809	73%
PR	4500	25%
SR	350	2%
Total	17659	100%

Table 3

Moreover, has been analyzed the numbers of single listings (SL) and multiple listings (MLs) in the Fashion Capital.

Types	n	%
SL	10643	60%
MLs	7016	40%
Total	17659	100%

Table 4

Finally matching together, the two analyses, they have given a better view of the distribution of the types of listings in Milan pre-covid.

Listing types	Listing forms	n	%
EH	SL	7751	44%
	MLs	5058	29%
PR	SL	2701	15%
	MLs	1799	10%
SR	SL	191	1%
	MLs	159	1%
Total		17659	100%

Table 5

The table provides valuable insights into Airbnb's offerings and the preferences of hosts and travelers:

Popularity of Entire Home Listings: It's clear that "Entire Home" listings are the most preferred by hosts, making up 73% of all listings. This type is popular among travelers looking for a private and comfortable space.

Diverse Listing Types: Airbnb caters to a wide range of traveler needs with its variety of listing types, including "Private Room" and "Shared Room." "Private Room" listings constitute 25%, while "Shared Room" listings are less common, at just 2%. This diversity ensures that Airbnb can accommodate different preferences and budgets.

Economic and Market Diversity: Airbnb's wide range of listing types and price points caters to diverse budgets and preferences. Travelers can find anything from high-end "Entire Home" rentals to budget-friendly "Private Room" options, highlighting Airbnb's ability to serve a broad spectrum of travelers.

Literature review about the variables of distance

Only a few papers have studied the variables of distance in correlation with the short-term rent market. none of those papers has studied how the performance of the Airbnb change in relation to the variables studied in this thesis: distance to the closest metro stations, to the closest public parking, to the closest ZTL or pedestrian areas, to the closest sport facility and finally to the closest area or building in degradation conditions.

Kirkos, E. (2022) studied the correlation between the performances of the Airbnb in the city of Salonicco, in Greece, in correlation with the distance to the central square, also known as “Aristotelous”. Using the regression has found that by going further from the central square the performances of the Airbnb decrease, in particular the Revenues and the occupancy have been studied in this paper.

Zhang, Z., Chen, R. J., Han, L. D., & Yang, L. (2017) studied the prices per nights of the Airbnb in the city of Nashville, Tennessee. They studied how the prices pe nights change depending on the distance to the highway and the convention center. The results showed that going further from the convention center, identified as the city center, is strongly related to a decrease in the price per night of the listings. the same happens for the highways, indeed going further from them is related to a decrease in the price per night, even if lower than in the case of the correlation with the convention center.

Bakker, M. (2021) conducted an analysis closer to this one where have been studied how the prices of the Airbnb are affected by the distances to the closest tourist spot. The first hypothesis was: “Closer access to touristic spots is positively associated with Airbnb listing prices per night”, founded to be true, indeed the author in his paper states “tourist spots have an important factor in the price determination of Airbnb, especially in areas between 0,2 km and 1 km”. The analysis has been conducted in a similar way of the one showed in this thesis, splitting the distances of each listing in the city of Amsterdam into 9 ranges of distance, to have a better overview of the phenomenon. The author identified as the main turists spots: the Rijksmusuem, Anne frank House, Van Gogh Museum, The Jordaan Area, Amsterdam Lookout, Body worlds, Vondelpark, Moco Musuem, Museum Het Rembrandthuis and Artis Zoo. A same analysis can be conducted in every city to study the dependence between the prices of the Airbnb and the distance to the closest tourist spot.

Perez-Sanchez, V. R., Serrano-Estrada, L., Marti, P., & Mora-Garcia, R. T. (2018) conducted an analysis in four different cities in the coastlines of Spain taking as a reference to compute the distance from each listing the city center and the distance from the coast. Evidence shows that “accommodation prices increase incrementally by 1.3% per kilometer from the tourist area, which in all four cases are situated in the historic area of the city. However, at the same time, accommodation prices decrease incrementally as distance from the coastline increases”.

Boto-García, D., & Leoni, V. (2022) studied how the performance of the short-term rent are affected from their distance to the coast in the Balearic Islands. They have taken as a reference sample in different ranges of distance from the closest point to the sea stating “we consider subsamples of listings that are located up to 500, 750, 1000, 2000, 3000, 4000 and 5000 meters away from the shoreline” going in depth also in how the coastline affects the performance of each listing.

“The purpose of this study is to identify the price determinants of sharing economy-based accommodation offers in the digital marketplace (specifically Airbnb.com). A sample of 180,533 accommodation rental offers from 33 cities listed on Airbnb.com is examined” Wang, D., & Nicolau, J. L. (2017). Between the huge number of variables studied is present also the distance of each listing to the city center. The analysis concluded with the authors stating, “The variable “distance,” representing accommodation location, has a significant negative effect, consistent with the findings of previous studies of hotel price determinants […]. The farther the accommodation from the city center, the lower its price.

Moreover, reviewing all the papers that studied the performances or the prices per nights of the short-term rent, we notice that the most majority of them focus their attention on the distance to the city center or to the central square, or the distance to the coast for the cities near by the sea. Everyone of them, indeed, computed the distance from a single point, this last one is the same for every listing. Only in the case of the analysis conducted in Amsterdam, a more complex algorithm has been presented, indeed having more tourist points led to a computation of the distance from every point and in a second phase there will be the selection of the closest tourist spot to each listing, saving the distance between the two. A similar analysis, but with different points of interest will be conducted in this thesis, where instead of using the tourist spots of Milan, different point of interest for the tourist have been studied, like the proximity to the metro station, the closest sport facility, the closest degradation area, the closest public parking and finally the closest ZTL or pedestrian area.

Finally, in a paper published in 2019, Buhalis analyzed how the booking probability of Airbnb listings in the city of London was correlated with the distance between the various properties from the city center and the subway, in addition to the performance variables. In this study, Buhalis considered 41,124 listings in the time period between March 5, 2017, and April 4, 2017. This study demonstrated how the composition of the listings and their geographical location are important for the performance of the various Airbnb listings. Indeed, from these analyses, some interesting results emerged. For example, it was found that moving 1 km away from the nearest subway station decreases the booking probability by 5.39%, while moving 1 km away from the city center results in a 6.78% reduction in the booking probability. This study demonstrated also that the signal attributes, which are designated to indicate the quality of the listings, are found to be important, especially for the listings without online reviews.

Research framework

The analysis of this thesis focused on the performances of the Airbnb of the city of Milan in relation to some variables of distance calculated from external datasets.

In our research framework the Y axe identifies the performance variables as Revenues, Reservation days, ADR, RevPAN and occupation rate. Those have been identified as the variables of performance that can describe how a listing is performing over the year considered in our analysis (from 2019 to 2020).

- The **revenues** Are expressed in USD and indicate how much a host has earned from a listing in a month.
- The variable **reservation days** indicates how many days in a month the listing has been booked by some client.
- **ADR**, that means average daily rate, represents the average daily price of the property. It is a useful indicator because it lets us get an idea about the profitability and of the trend of tariffs over time.

$$ADR = \frac{Revenues\ (USD)}{Reservation\ days}$$

- **RevPAN**, meaning Revenue Per Available Night, is useful is useful to get an in depth focus about the performance of a listing, it is defined as:

$$RevPAN = Occupation\ Rate * ADR$$

$$RevPAN = \frac{Reservtion\ days}{Reservation\ days + available\ days} * \frac{Revenues\ (USD)}{Reservation\ days}$$

$$RevPAN = \frac{Revenues\ (USD)}{Reservation\ days + Available\ days}$$

- **Occupation rate** (OCC) explains the ratio between the night booked by the clients and the total available days that can be booked in a month.

$$Occupation\ rate = \frac{Reservation\ days}{Reservation\ days + Available\ days}$$

In the X axe, instead, our framework presents the independent variables, computed using Python and making calculations explained in the next chapter, about the distance from each listing to the closest "point of interest".

A "point of interest" is referred as: a metro station, a site in a degradation condition (defined in that way by the municipality of Milan), a sport facility, a public parking and finally an area with restriction, which can be a ZTL (zona traffico limitato) or a pedestrian area. From the join between the external datasets, found on the website of the municipality of Milan (https://dati.comune.milano.it), have been computed the variables of distance used in the X axe of our analysis.

Finally, as control variables for our model have been chosen some variables present in the original dataset of the Airbnb of Milan that can describe appropriately the model. On those variables some adjustments have been made to make them becoming dummy variables:

- **MaxGuest** didn't change its essence, representing still the maximum number of guests that a listing can host.
- The **ListingType** has now become a dummy variable that identifies an entire home or an apartment with 1 and any other option with 0, as private room, shared room or hotel room.
- **LTR** (long term rent) has become a dummy variable that identifies a long-term rent, so a listing that requires a minimum stay of at least 28 days is now identified with 1 and 0 otherwise.
- The **IstantBook** variable didn't change its essence too because it is already a variable that identifies with 1 the case in which the listing can be booked instantly, and 0 otherwise.
- The variable **SuperHost** didn't change its essence because it identifies with 1 the listings managed by a "super host" and with 0 the ones not.

Together with these variables has been implemented also the variable describing the NIL in which a listing has its coordinates (computed using another external source on the website of the municipality of Milan, in particular a map in the format .geojson that matched on Python with our dataset return the NIL for each listing), the year, and the month of the registration of the data.

Research questions

In the literature review section, we analyzed Airbnb's activity, discussing its business model and the effects this platform has had on the real estate market. We then focused on the city of Milan, starting with a study on the subdivision of various zones, discussing the main events in the city, analyzing the seasonality of tourism in the Lombardy capital, and observing how the Covid-19 pandemic and related restrictions influenced the platform's performance.

While reviewing all articles related to the city of Milan, we noticed a lack of papers that delved into analyzing the most determining factors for Airbnb hosts' performances in detail. Therefore, this study aims to fill a gap in academic literature through the acquisition and processing of data taken from the Municipality of Milan, specifically regarding the presence of metro stations, degraded areas, sports facilities, restricted zones, and parking spaces.

Analyses will be conducted on this data from the period between 2019 and 2022 in the city of Milan. These data will be concurrently compared with the performances of the platform, with the goal of identifying relationships and trends existing between strategic variables and performance indicators (revenues, occupancy rate, and certain metrics like ADR and RevPAN).

In conclusion, the research question we posed were how do the performances of Airbnb properties in the city of Milan vary based on qualitative variables such as the presence of subways, degraded areas, sports facilities, restricted zones, and the availability of parking?

Before starting, we formulated 5 hypotheses that will be confirmed or refuted following the analyses:

Hypothesis 1: The presence of a metro station near the property positively influences the performance of such Airbnb.

Analyzing the structure of the city of Milan and the location of various attractions within it, we deemed the presence of a metro station crucial for a potential customer in choosing accommodation. Proximity to subway stations, according to our reasoning, would make lodging more accessible and convenient for visitors, facilitating city exploration.

Hypothesis 2: The presence of a degraded area near the property negatively influences the performance of such Airbnb.

The presence of degraded areas can impact guests' perception of safety, potentially discouraging potential visitors, especially if they perceive a higher risk of crime or other security issues nearby. According to this reasoning, guests might prefer staying in safer and well-maintained areas to ensure a better-quality experience during their stay. For this reason, we believe there is a negative correlation between this factor and the property's performance.

Hypothesis 3: The presence of parking near the property positively influences the performance of such Airbnb.

Parking availability would make the stay more convenient for guests arriving by car. It could be a crucial deciding factor for those who prefer using their vehicle or renting one during their stay. Therefore, we anticipate a positive correlation between this factor and the property's performance.

Hypothesis 4: The presence of ZTL/Pedestrian Areas near the property positively influences the performance of such Airbnb.

The presence of ZTL and pedestrian areas is associated with reduced noise and traffic congestion, likely located in more central parts of the city. Therefore, we believe there is a positive correlation between this factor and the property's performance.

Hypothesis 5: The presence of sports facilities near the property positively influences the performance of such Airbnb.

Proximity to sports facilities is advantageous for those interested in physical activities or attending sports events. Guests participating in or observing competitions might find it convenient to stay near sports facilities. For this reason, we believe there is a positive correlation between proximity to sports facilities and the performance of the property.

Data and Methodology

In this chapter will be explained the datasets that have been used in our analysis, where those were taken and the process through which the new variables have been computed.

Airbnb's dataset

The analysis began from the dataset of the listings of all the Airbnb in the city of Milan. For each row of the dataset there is the description of a single listing for a specific month collecting different variables qualitative and quantitative. The Dataset presents 696565 records across the pre, post and pandemic years, indeed from 2019 to 2022. This led us to a better view on the phenomenon of the sharing economy, in housing field, in Milan.

The records of the Dataset are the following ones:

- **Property ID**: the identification string that characterize univocally a single listing making it different from every other one.
- **Reporting Month**: in the format of a date, it points the month to which the record is referring to.
- **Year**: in the format of a number, it refers to the year considered.
- **KEY_YEAR**: string concatenation between the variables "Property ID" and "Year".
- **Revenue (USD)**: refers to the revenues in Dollars of the Airbnb in the determined month considered.
- **Reservation Days**: it refers to the number of days in the considered month in which the listing has been booked.
- **Available Days**: it refers to the number of days in the considered month in which the listing has not been booked.

- **Blocked Days**: it refers to the number of days in the considered month in which the listing has been blocked (by the host?)
- **Occupation Rate**: variable computed as fraction between the variable “Reservation Days” + “Blocked Days” and the total days in the month considered.
- **ADR**: variable computed as fraction between “Revenue (USD)” and “Reservation Days” pointed indeed to the average revenues per night booked.
- **RevPAN**: variable computed as fraction between “Revenue (USD)” and “Reservation Days” + “Available Days” that pointed indeed to the average revenues per day of the listing in the determined month.
- **Listing Type**: it entails all the different possibilities of the listings from a point of view of type of the listing. Can be: Entire home/apt; Hotel room; Private room; Shared room.
- **Bedrooms**: refers to the number of bathrooms in the Airbnb.
- **Bathrooms**: refers to the number of bedrooms in the Airbnb.
- **Max Guests**: this field indicates the maximum number of guests that the property can accommodate.
- **Latitude**: this field indicates the property's latitude, which is the north-south location of the Airbnb on the surface of the Earth.
- **Longitude**: this field indicates the property's longitude, which is the east-weast location of the Airbnb on the surface of the Earth.
- **Neighborhood**: this field indicates the neighborhood within which the Airbnb is located.
- **Cancellation policy**: this field indicates how willing and available a property is regarding the cancellation policy and its timing.
- **Instantbook Enabled**: this field indicates whether the reference Airbnb provides an instant booking function, the values that can be assumed are only True or False
- **Number of Photos**: this field indicates the number of photos relating to the property inserted on the Airbnb portal by the Host.
- **Number of Reviews**: this field indicates the number of reviews left within the Airbnb portal by users who have stayed in the specific structure.

- **Minimum Stay**: this field indicates the minimum number of nights that can be booked for the specific property. This parameter is chosen by the structure host.
- **Published Weekly Rate (USD)**: (this field indicates an average weekly rate for the specific Airbnb)
- **Published Monthly Rate (USD)**: (this field indicates an average monthly rate for the specific Airbnb)
- **Airbnb Superhost**: this field indicates whether an Airbnb host is a Superhost. The criteria to be part of this category are: having completed at least 10 stays or 3 reservations, for a total of at least 100 nights, having maintained a minimum response rate of 90%, having maintained a cancellation rate of less than 1% and maintaining an overall rating of 4.8. The values that can be assumed are only True or False
- **Airbnb HostID**: this field consists of a unique numeric code associated with a single Airbnb Host; a host can have multiple properties but will always be characterized by the same code.

External datasets

Other than Airbnb's one there has been done an analysis on the certified website of the municipality of Milan to identify external sources for this study of Lombardia's capital. The "Geoportale of Milan" and the website of the municipality of Milan have been interesting sources for external datasets to be matched with Airbnb's one.

Five different datasets have been identified that can bring an important contribution to the research and can lead to different types of analysis for a better understanding of the Airbnb's phenomenon in the city of Milan. Each of them shows a different type of phenomenon in Milan as follows:

- Public parking
- Metro stations
- Pedestrian area
- Sports facilities
- Degradation areas and buildings

Each of them has been downloaded as a file .csv to work on it using Python. For the datasets that do not present a column that univocally identifies each record there has been created an ID using Python.

Deling with geographical coordinates means that to compute the distance between a source and each Airbnb a transformation of the parameters must be done. The Haversine formula calculates the distance between two points on the surface of a sphere (such as the Earth) given their latitude and longitude. Here's the Haversine formula in mathematical notation.

$$a = \sin^2\left(\frac{\Delta lat}{2}\right) + \cos(lat_1) * \cos(lat_2) * \sin^2(\frac{\Delta lon}{2})$$

$$c = 2 * atan2(\sqrt{a}, \sqrt{1-a})$$

$$distance = R * c$$

Where:

- lat_1, lon_1 are the latitude and longitude of each Airbnb.
- lat_2, lon_2 are the latitude and longitude of the closest source considered.
- $\Delta lat = lat_2 - lat_1$
- $\Delta lon = lon_2 - lon_1$
- R is the radius of the Earth (median radius = 6,371 kilometers)

This formula gives the distance in the same units as the radius of the Earth (e.g., kilometers). Therefore, the multiplication by 1,000 must be done.

For each external dataset two analyses have been done to identify the closest source to each Airbnb and also to identify the intensity of each source. For this last one there has been decided to use the NIL (Nuclueo identità locale) rather than the neighborhood, that is already present in the Airbnb's dataset, because a NIL identifies a stricter area, and the analysis would be more precise. The NIL are specified in the website of the municipality of Milan, there is a .csv file stating for each NIL important data as latitude and longitude as long as the ID, the name and the shape, but even more important has been the file .geojson that is actually a map of the city with the subdivision of the NIL with the shape. This Map has been crucial for the analysis because with the libraries of Pandas, GeoPandas, shapely.geometry and requests was possible to import it to python and work with it. In *Fig. 15* there is shown the map of the NIL released by the municipality of Milan, below the list of all the NIL associated to the numbers in *figure 15* split by “Municipio”.

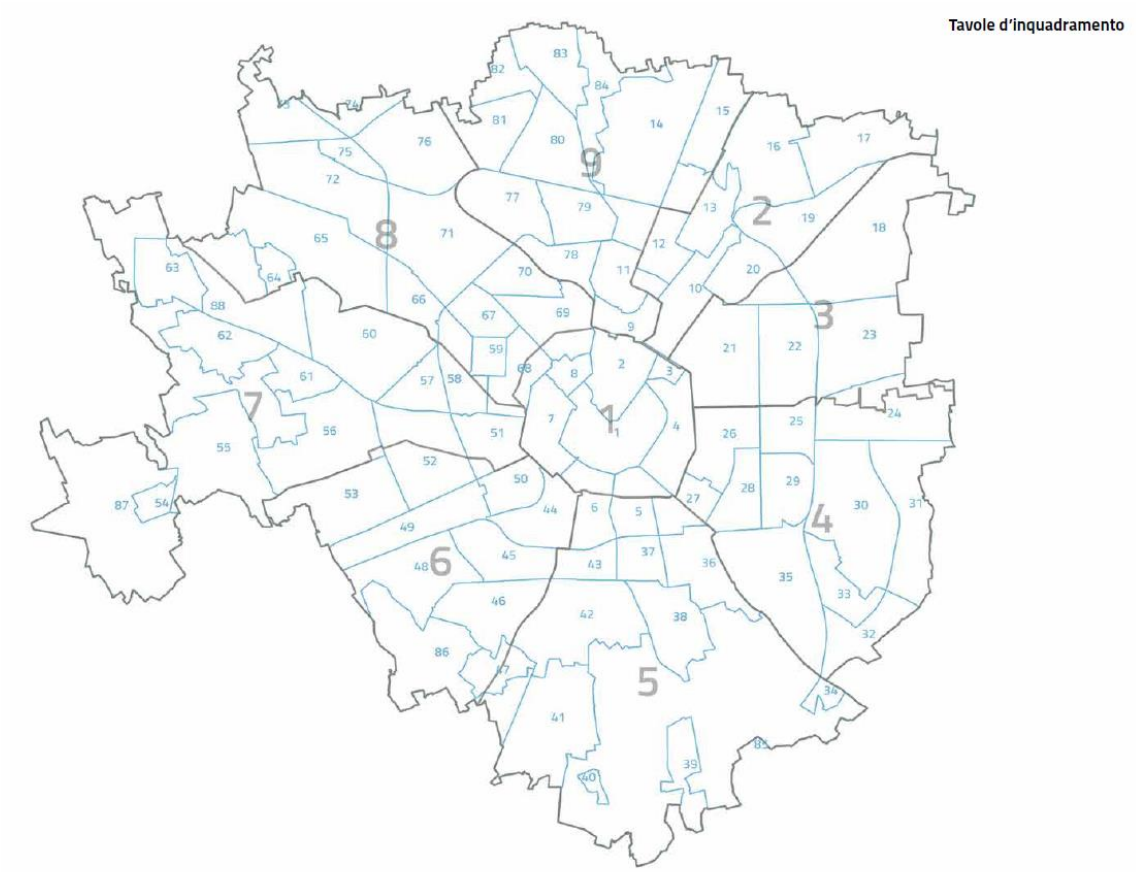


Figure 15

Municipio 1
1. Duomo
2. Brera
3. Giardini Porta Venezia
4. Guastalla
7. Magenta- San Vittore
8. Parco Sempione
(5. Vigentina)
(6. Ticinese)
(68. Pagano)
(69. Sarpi)

Municipio 2
10. Stazione Centrale - Ponte Seveso
16. Gorla - Precotto
17. Adriano
19. Padova - Turro - Crescenzago
(11. Isola)
(12. Maciachini-Maggiolina)
(13. Greco - Segnano)
(20. Loreto - Casoretto - NoLo)

Municipio 3
18. Cimiano - Rottole - Q.re Feltre
21. Buenos Aires - Porta Venezia - Porta Monforte
22. Città studi
23. Lambrate - Ortica
(20. Loreto)
(24. Parco Forlanini - Cavriano)

Municipio 4
25. Corsica
26. XXII Marzo
28. Umbria - Molise - Calvairate
29. Ortomercato
30. Taliedo - Morsenchio - Q.re Forlanini
31. Monluè - Ponte Lambro
32. Triulzo Superiore
33. Rogoredo - Santa Giulia
35. Lodi-Corvetto
(27. Porta Romana)

Municipio 5
5. Porta Vigentina - Porta Lodovica
6. Porta Ticinese - Conca del Naviglio
36. Scalo Romana
34. Chiaravalle
37. Morivione
38. Vigentino - Q.re Fatima
39. Quintosole
40. Ronchetto delle Rane
41. Gratosoglio - Q.re Missaglia
- Q.re Terrazze
42. Stadera - Chiesa Rossa - Q.re Torretta
- Conca Fallata
43. Tibaldi
85. Parco delle Abbazie
86. Parco dei Navigli
(47. Cantalupa)

Municipio 6
44. Porta Ticinese - Conchetta
45. Moncucco - San Cristoforo
46. Barona
47. Cantalupa
48. Ronchetto sul Naviglio
- Q.re Lodovico il Moro
49. Giambellino
50. Porta Genova
52. Bande Nere
53. Lorenteggio
86. Parco dei Navigli
(51. Washington)

Municipio 7
51. Porta Magenta
54. Muggiano
55. Baggio - Q.re degli Olmi - Q.re Valsesia
56. Forze Armate
57. San Siro
58. De Angeli-Monte Rosa
60. Stadio - Ippodromi
61. Quarto Cagnino
62. Quinto Romano
63. Figino
87. Assiano
88. Parco Bosco in città
(68. Pagano)

Municipio 8
59. Tre Torri
64. Trenno
65. Q.re Gallaratese - Q.re San Leonardo
- Lampugnano
66. QT8
67. Portello
68. Pagano
69. Sarpi
70. Ghisolfa
71. Villapizzone - Cagnola - Boldinasco
72. Maggiore - Musocco - Certosa
73. MIND - Cascina Triulza
74. Roserio
75. Stephenson
76. Quarto Oggiaro - Vialba - Musocco
(88. Parco Bosco in città)

Municipio 9
9. Porta Garibaldi - Porta Nuova
11. Isola
14. Niguarda - Ca' Granda - Prato Centenaro
- Q.re Fulvio Testi
15. Bicocca
77. Bovisa
78. Farini
79. Dergano
80. Affori
81. Bovisasca
82. Comasina
83. Bruzzano
84. Parco Nord
(12. Maciachini-Maggiolina)
(13. Greco)

In the next paragraph a more precise explanation of how each dataset has been implemented to create the final one having in each record the listing and the data of the external datasets.

Public parking

Like many other big cities, Milan has one of the most intense traffic and number of cars all over the world. The importance of having a garage or a parking spot has increased since the city has grown exponentially and the number of people using cars in the city too. It is very important, indeed, have a public parking as close as possible to your house or hotel if you are visiting the town and you will get there by car, in this study there will be an analysis of how having a parking close to the Airbnb affects the performances of this.

The datasets of the public parking on the website of the municipality of Milan has in each record a public parking with qualitative and quantitative information about the specific parking as:

- **_id**: that is automatically generated by the dataset that identifies univocally the record.
- **id**: an identifier number of the public car park.
- **nome**: a string that states the name of the parking.
- **n_posti**: a number stating the number of spaces of the parking.
- **indirizzo**: address of the parking
- **comune**: municipality in which there is parking (for each record is specified "Milan").
- **tipo**: type of the parking that can be "Autorimessa convenzionata", "Pubblici", "Pubblici/Residenti" or "Residenti/Pubblici".
- **LONG_X_4326**: variable that describes the longitude of the parking spot.
- **LAT_Y_4326**: variable that describes the latitude of the parking spot.
- **Location**: a string variable stating all latitude and longitude of the parking spot as following (Latitude, Longitude).

Sport Facilities

Municipally owned sports facilities form the backbone of the city's sports system.

The management of the facilities is distributed to the Milanosport spa company and to the sports federations, sports promotion bodies, associations, or amateur sports clubs.

There are different types of sport facilities, from the data collected we find for example: Athletics, Dance, Basketball, Football, Fitness, and many others.

Sports activity is one of the main and fundamental activities for leading a healthy lifestyle. In fact, in the following chapters we will analyze whether and how the presence of these sports facilities can influence the performance of Airbnb.

The sports areas dataset was found on the website of the municipality of Milan reporting more variables for each sheet (which in this case it is a sports center) like:

- **Longitude**: the longitude of the sports center.
- **Latitude**: the latitude of the sports center.
- **Type**: indicates the type of sport for which the center is equipped
- **NIL**: Indicates the NIL in which the sports facility is located

Degradation

Although in recent years reconversion processes have been initiated for significant abandoned areas, equipment and systems have still been present for several years in conditions of decommissioning and functional underuse. Their redesign represents an extraordinary opportunity to redevelop significant areas of the urbanized city in terms of land and the development of new services.

These areas largely coincide with the areas of the heritage of state bodies such as the railways and the military state property, as well as private and municipally owned areas. Entire railway yards not in operation, located within the city, which at the time of their construction influenced the homogeneous development of urban systems, forming enclaves and large open spaces constituting physical barriers and separations between areas belonging to the same urban area.

The other areas concern parts of the city on which unfinished transformation procedures have been started or buildings and systems on which redevelopment or enhancement proposals have been made.

In this study there will be an analysis of how having a degraded area nearby the Airbnb affects its performance.

The degradation dataset was found on the website of the municipality of Milan reporting more variables for each sheet (which in this case it is a area of degradation) like:

- **LONG_X_4326**: the longitude of the area of degradation.
- **LAT_Y_4326**: the latitude of the area of degradation.

- **OBJECTID**: indicates the unique ID of the degradation zone.
- **TYPE_MACRO**: indicates the type of area of degradation.
- **NIL**: Indicates the NIL in which the area of degradation is located

ZTL/Restricted Areas

Zones with Restricted Traffic Access (ZTL) and pedestrian areas are urban planning strategies implemented in various cities, like Milan, to regulate vehicular traffic and enhance the livability of specific zones. These measures are designed to address concerns such as air quality, noise pollution, and pedestrian safety.

Zones with Restricted Traffic Access (ZTL):

- **Definition**: ZTLs are areas where access by vehicles is restricted or regulated during certain times or altogether. The primary goal is to reduce congestion, improve air quality, and create more pedestrian-friendly environments.
- **Implementation**: ZTLs are typically marked by specific signage and access control points. Only authorized vehicles, such as residents or those with special permits, are allowed to enter these zones during restricted hours.
- **Purpose**: ZTLs aim to promote sustainable transportation, decrease pollution levels, and create more pleasant urban spaces.

Pedestrian Areas:

- **Definition:** Pedestrian areas are zones within a city where vehicular traffic is entirely prohibited, prioritizing pedestrians and non-motorized modes of transportation.
- **Characteristics:** Pedestrian areas often feature wide sidewalks, public spaces, and amenities to encourage walking and social interaction. They are commonly found in city centers, shopping districts, and cultural or historical sites.
- **Benefits:** Creating pedestrian-friendly zones can lead to improved air quality, increased foot traffic for local businesses, enhanced safety for pedestrians, and the promotion of a more sustainable and active lifestyle.
- **Design:** Urban planners may incorporate features like benches, green spaces, and public art to make pedestrian areas more attractive and enjoyable.

It is important to note that information on ZTLs and restrictions may change over time, so it is advisable to check the latest provisions with local authorities or on the official website of the Municipality of Milan for updated information.

In this study there will be an analysis of how having a restricted area nearby the Airbnb affects its performance.

The degradation dataset was found on the website of the municipality of Milan reporting more variables for each sheet (which in this case it is an area of degradation) like:

- **LONG_X_4326**: the longitude of the restricted area.
- **LAT_Y_4326**: the latitude of the restricted area.
- **ID_AMAT**: indicates the unique ID of the restricted area.

Metro stations

Milan, a major city in Italy, features an extensive metropolitan subway system known as the Milan Metro. The metro network plays a crucial role in facilitating transportation within the city and its surrounding areas. Here's an overview of the Milan Metro stations:

Line M1 (Red Line):

- **Characteristics**: The Red Line, Line M1, is one of the oldest metro lines in Milan, connecting the northwest and southeast areas of the city.
- **Key Stations**: Duomo (city center), Cadorna, Loreto, and Rho Fiera (connecting to the exhibition center).

Line M2 (Green Line):

- **Characteristics**: The Green Line, Line M2, intersects with Line M1 and serves the northeastern and southwestern parts of Milan.
- **Key Stations**: Centrale (central railway station), Porta Garibaldi, Cadorna, and Assago.

Line M3 (Yellow Line):

- **Characteristics**: The Yellow Line, Line M3, runs from the northwest to the southeast, intersecting with Lines M1 and M2.
- **Key Stations**: Duomo, Centrale, Porta Romana, and San Donato.

Line M4 (Lilac Line):

- **Characteristics**: Line M4, currently under expansion, will connect the city center with the southwestern areas, including Linate Airport.
- **Key Stations**: Forlanini FS and Linate Airport (future extension).

Interchanges and Connectivity:

- **Centrale**: Milan's central railway station is a major interchange, connecting multiple metro lines and serving as a transportation hub for trains, buses, and taxis.
- **Cadorna**: A key interchange station connecting Lines M1 and M2.

The Milan Metro system has undergone modernization efforts and expansion over the years to accommodate the growing transportation needs of the city.

Navigating Milan's metro system is convenient for both locals and visitors, providing efficient access to key landmarks, business districts, and transportation hubs. The continuous development and expansion of the metro network underscore Milan's commitment to sustainable urban mobility.

The dataset of the metro has been founded in the website of the municipality of Milan stating multiple variables for each record (which in this case is a metro station) as

- **id_amat**: the number that identificate univocally the metro station.
- **Nome**: the name of the metro station as a string.
- **Line**: the number of lines that pass through this specific metro station.
- **LONG_X_4326**: the longitude of the metro station.
- **LAT_Y_4326**: the latitude of the metro station.
- **Location**: a string that states the position of the metro station as follows: (latitude, longitude)

Example of the creation of a dataset

To explain better what has been reported before is possible to see below the code used to create the variables and the datasets. The first code refers to the creation of the variables of the closest public parking to each Airbnb, with the distance expressed in meters, the name of the closest parking and the number of spaces of the public parking in case:

```
import csv
import math

def convert_to_float(value):
    try:
        return float(value)
    except ValueError:
        return None

def convert_to_int(value):
    try:
        return int(value)
    except ValueError:
        return None

def haversine(lat1, lon1, lat2, lon2):
    # Raggio della Terra in metri
    R = 6371000.0

    # Conversione delle coordinate da gradi a radianti
    lat1_rad = math.radians(lat1)
    lon1_rad = math.radians(lon1)
    lat2_rad = math.radians(lat2)
    lon2_rad = math.radians(lon2)

    # Differenze nelle coordinate
    dlat = lat2_rad - lat1_rad
    dlon = lon2_rad - lon1_rad

    # Formula di Haversine
    a = math.sin(dlat / 2)**2 + math.cos(lat1_rad) * math.cos(lat2_rad) * math.sin(dlon / 2)**2
    c = 2 * math.atan2(math.sqrt(a), math.sqrt(1 - a))

    # Distanza in metri
    distance_meters = R * c

    return distance_meters
# Load Parcheggi data
with open("C:/Users/Andrea/Desktop/poli/Magistrale/Tesi/Dataset esterni/parcheggi_pubblici.csv",
newline="", encoding="ISO-8859-1") as fileParcheggi:
    lettoreParcheggi = csv.reader(fileParcheggi, delimiter=",")

    # Skip the header line
    next(lettoreParcheggi, None)
```

```
    datiParcheggi = [(riga[2], riga[3], convert_to_float(riga[8]), convert_to_float(riga[7]))
for riga in lettoreParcheggi if riga[0] != ""]

# Load Airbnb data
with open("C:/Users/Andrea/Desktop/poli/Magistrale/Tesi/DB_MILAN.csv", newline="",
encoding="ISO-8859-1") as fileAirbnb:
    lettoreAirbnb = csv.reader(fileAirbnb, delimiter=",")

    # Skip the header line
    next(lettoreAirbnb, None)

    datiAirbnb = [(riga[0], convert_to_float(riga[15]), convert_to_float(riga[16]), None) for
riga in lettoreAirbnb if riga[0] != ""]

# Calculate the distance to the closest degraded point for each Airbnb listing
for i in range(len(datiAirbnb)):
    min_distance = float('inf')  # Set initial min_distance to positive infinity
    TIPO_MACRO = None
    ID_NIL = None
    NIL = None

    for j in range(len(datiParcheggi)):
        if None in (datiAirbnb[i][1], datiAirbnb[i][2], datiParcheggi[j][2],
datiParcheggi[j][3]):
            continue  # Skip rows with missing values

        d = haversine(datiAirbnb[i][1], datiAirbnb[i][2], datiParcheggi[j][2],
datiParcheggi[j][3])
        if d < min_distance:
            min_distance = d
            NOME_PARCHEGGIO = datiParcheggi[j][0]
            NUMERO_POSTI = datiParcheggi[j][1]

    datiAirbnb[i] = (datiAirbnb[i][0], datiAirbnb[i][1], datiAirbnb[i][2], NOME_PARCHEGGIO,
NUMERO_POSTI, min_distance)

# Save the results to a new CSV file
output_file_path = "C:/Users/Andrea/Desktop/poli/Magistrale/Tesi/airbnb_with_parking_info.csv"
with open(output_file_path, mode="w", newline="", encoding="ISO-8859-1") as output_file:
    scrittore = csv.writer(output_file)

    # Write the header
    scrittore.writerow(["Listing_ID", "Latitude", "Longitude", "NOME_PARCHEGGIO",
"NUMERO_POSTI", "min_distance_parking"])

    # Write the data
    scrittore.writerows(datiAirbnb)

print(f"Data saved to {output_file_path}")
```

The second code, instead, refers to the creation of the dataset for the intensity of public parking per NIL, specifying the number of public parking per each NIL of the municipality of Milan:

```
import geopandas as gpd
import pandas as pd
from shapely.geometry import Point
import requests

# Load the parking dataset (CSV file)
parking_data = pd.read_csv("C:/Users/Andrea/Desktop/poli/Magistrale/Tesi/Dataset
esterni/parcheggi_pubblici.csv", encoding="ISO-8859-1")

# Create GeoDataFrame from parking_data
geometry = [Point(xy) for xy in zip(parking_data['LONG_X_4326'], parking_data['LAT_Y_4326'])]
parking_gdf = gpd.GeoDataFrame(parking_data, geometry=geometry, crs="EPSG:4326")

# Download the NIL boundaries GeoJSON file
nil_geojson_url = "https://dati.comune.milano.it/dataset/e5a0d956-2eff-454d-b0ea-
659cb7b55c0b/resource/af78bd3f-ea45-403a-8882-91cca05087f0/download/nilzone.geojson"
nil_geojson_local_path = "C:/Users/Andrea/Desktop/poli/Magistrale/Tesi/nilzone.geojson"
response = requests.get(nil_geojson_url)
with open(nil_geojson_local_path, "wb") as f:
    f.write(response.content)

# Load the NIL boundaries from the local GeoJSON file
nil_boundaries = gpd.read_file(nil_geojson_local_path)

# Spatial join to associate each parking facility with a NIL
joined_data = gpd.sjoin(parking_gdf, nil_boundaries, how="left", op="within")

# Group by NIL and calculate the sum of parking facilities and total number of spaces
nil_parking_stats = joined_data.groupby("NIL").agg(
    Parking_Count=('id', 'count'),
    Total_Spaces=('n_posti', 'sum')
).reset_index()

# Display or save the results
print(nil_parking_stats)

# Save the results to a new CSV file if needed
nil_parking_stats.to_csv("C:/Users/Andrea/Desktop/poli/Magistrale/Tesi/nil_parking_stats.csv",
index=False)
```

Dealing with the different datasets, for each record of the dataset of the Airbnb of Milan, has been created new variables describing the distance and the intensity of "the points of interest". For each of the "point of interest" considered has been calculated:

- The distance, in meters, between each listing and the closest "point of interest" (for example the closest metro station for each Airbnb)
- The number of "point of interest" in the same NIL (Nucleo di identità locale) of the Airbnb (for example a listing in the NIL "duomo" has now a new column with the number of metro stations in the NIL "duomo")

- Finally, a discrete variable has been computed splitting the distances in meters in multiple ranges: 7 ranges have been created splitting the distance as follows:

in the first range you have the point of interest closer than 200m, then in then in the second range you'll have the point of interest between 200m and 500m and so on.

Descriptive analysis

In the context of this research, descriptive analysis emerges as a crucial phase in understanding and exploring the collected data. This methodology aims to provide a detailed overview of the fundamental characteristics of the variables involved, outlining an initial framework within which more in-depth analyses (uni and multivariate regressions) will be conducted.

Descriptive analyses focus on clear and detailed presentation of the collected information, offering an initial observation of distributions, trends, and relationships present in the data. Through basic statistical calculation, creation of graphical visualizations, and exploration of key features, this phase provides a fundamental starting point for understanding the available data.

The primary goal is to provide an initial insight into the dynamics present in the data, enabling the identification of significant phenomena that may influence the overall interpretation of the analysis.

This initial process of data exploration not only provides a foundation for more advanced analyses but also offers initial feedback regarding the hypotheses formulated in the preceding section.

The ultimate purpose of this analysis is to verify how the number of Airbnb listings within the created ranges (metro, parking, ZTL, sports facilities, and degraded areas) has changed during the pre, during, and post-Covid periods.

To verify how the performance variables vary for the different ranges created (metro, parking, ZTL, sports facilities, and degraded areas).

For the descriptive statistical analysis of this paragraph, it is decided to proceed simultaneously on multiple time horizons, having data available from 2019 to 2022. The following time intervals are considered: pre-Covid (2019), Covid (2020/2021), and post-Covid (2022).

General performances:

We started by analyzing how certain general variables varied over the years.

Below is a summary table:

	2019	2020	2021	2022
# active listings	30595	24426	21217	24721
% changes YoY	-	-20%	-13%	+17%
% changes from 2019	-	-20%	-31%	-19%
Avg reservation days	10.34	5.90	6.69	12.46
Std. Err.	0.023	0.021	0.023	0.027
Max [Reservation days]	31	31	31	31
Min [Reservation days]	0	0	0	0

Table 6 General performance trend

As evident from the table above, the number of active listings, i.e., properties with bookings, sharply declined starting from 2020 due to the pandemic. We observed a 20% reduction compared to the previous year, which further decreased by another 13% in 2021, reaching the lowest value. In 2022, we see this figure increase by 17%, surpassing the value recorded in 2020 but remaining far below the pre-pandemic level measured in 2019 (-19%).

Another data point we monitored was the average reservation days. In 2019, for the 30,595 active listings, we had an average of 10.34 days reserved per month. This figure decreased to 5.90 days in 2020 due to significant restrictions imposed by the government to combat the pandemic. In 2021, the value increased to 6.69 days and rose to 12.46 days in 2022. This increase is primarily attributed to a significant decrease in restrictions, thanks to the decline in the number of infections, which facilitated the recovery of tourism, in this case in the city of Milan.

Below are the graphs representing the above trends:

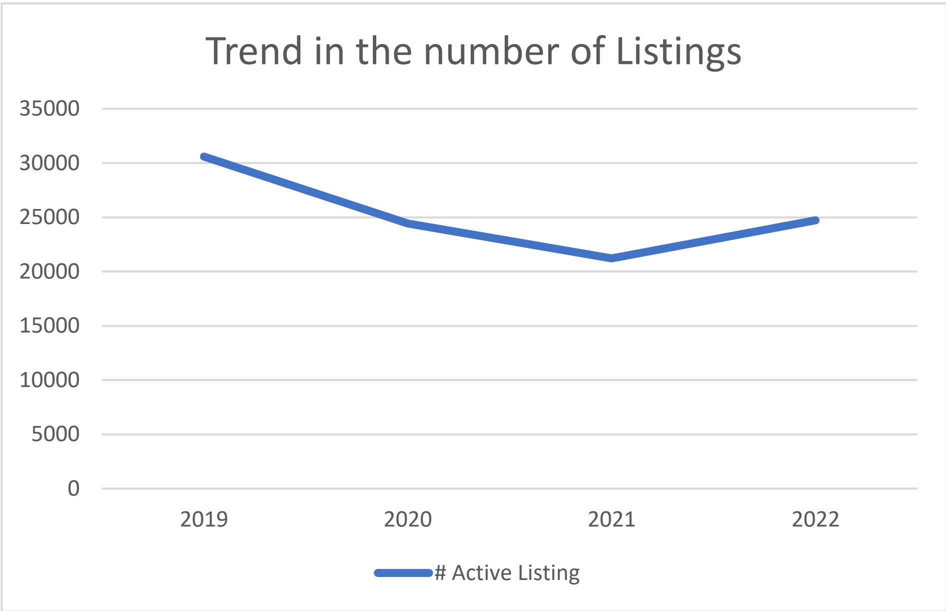


Figure 16, number of listings by year in the city of Milan

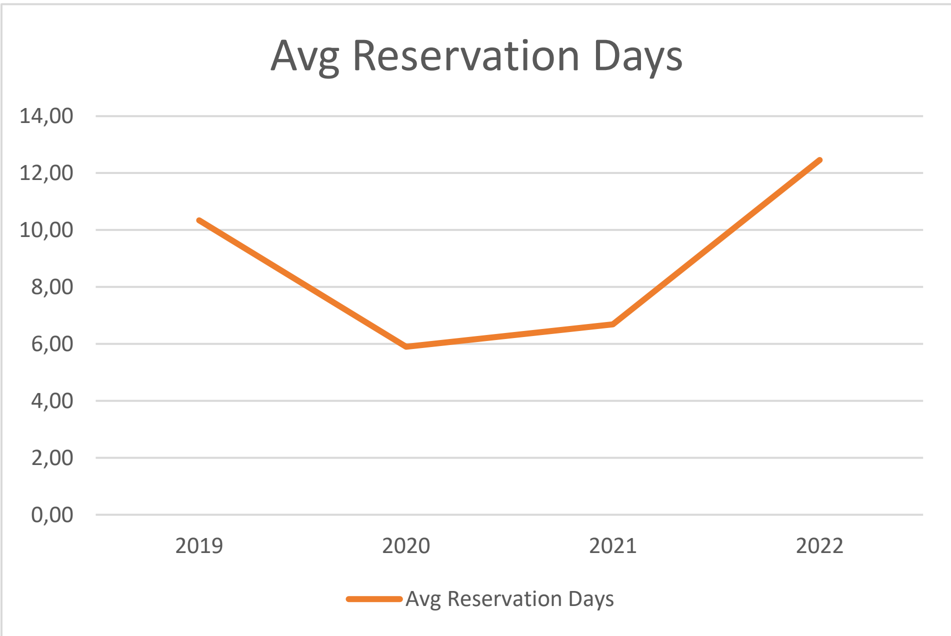


Figure 17, average reservation days by year in the city of Milan

Performances variables

Subsequently to the analyses on general performances, we moved on to analyze the average values assumed by the performance variables in the various time intervals to appreciate the differences brought about by the pandemic and understand if there has been a recovery post-Covid.

Below is a summary table:

Average	Pre-Covid	Covid		Post-Covid
	2019	2020	2021	2022
Revenues	1100.32	596.62	807.37	1769.26
Reservation Days	10.34	5.90	6.69	12.46
OCC	0.41	0.26	0.29	0.52
RevPAN	44.65	26.65	35.58	73.05
ADR	118.59	116.86	128.05	149.83

Table 7 Trend of performance variables

As evident from *Table 7*, all the performance variables follow a similar trend. In 2020, they experience a decline due to the onset of the pandemic. For instance, the host revenues nearly halve (from 1100 to 596) between 2019 and 2020. In 2021, there is a small sign of recovery that raises the values of the performance variables. An especially interesting data point is the ADR value in 2021 compared to 2019. Considering that ADR is the ratio between Revenues and Reservation Days, we can observe that the ADR value increased in 2021 (from 118 to 128). Therefore, hosts in 2021 were able to earn more per each reserved night. In summary, we can observe fewer reserved days at higher prices.

On the other hand, we notice how in 2022, the values of all these performance variables surpassed the pre-pandemic 2019 values.

New variables

We now determine the descriptive statistics for the new variables created. The objective is to initially analyze the number of properties located within the 7 ranges created for the 5 new categories (metro, parking, degradation, sports facilities, and ZTL).

Below are the 5 summary tables:

Listing Airbnb in the **Metro** ranges	Pre COVID	COVID		Post COVID
	2019	**2020**	**2021**	**2022**
1	4.659	3.896	3.356	3.702
2	13.142	10.699	9.379	10.778
3	5.678	4.404	3.770	4.446
4	3.114	2.395	2.046	2.497
5	2.928	2.211	1.946	2.365
6	758	564	509	689
7	316	257	211	244
Total	30.595	24.426	21.217	24.721

Table 8, Listing Airbnb in the metro ranges.

Listing Airbnb in the **Degradation** ranges	Pre COVID	COVID		Post COVID
	2019	**2020**	**2021**	**2022**
1	3.960	3.063	2.715	3.218
2	12.235	9.524	8.221	9.631
3	7.954	6.394	5.577	6.479
4	4.090	3.376	2.939	3.351
5	2.115	1.769	1.561	1.794
6	168	136	120	139
7	73	164	84	109
Total	30.595	24.426	21.217	24.721

Table 9, Listing Airbnb in the degradation ranges.

Listing Airbnb in the **Parking** ranges	Pre COVID	COVID		Post COVID
	2019	**2020**	**2021**	**2022**
1	3.141	2.510	2.223	2.512
2	9.453	7.620	6.724	7.685
3	6.579	5.279	4.679	5.458
4	4.050	3.162	2.718	3.176
5	3.410	2.628	2.222	2.608
6	1.956	1.539	1.296	1.563
7	2.006	1.688	1.355	1.719
Total	30.595	24.426	21.217	24.721

Table 10, Listing Airbnb in the parking ranges.

Listing Airbnb in the **Sport** ranges	Pre COVID	COVID		Post COVID
	2019	**2020**	**2021**	**2022**
1	1.926	1.517	1.321	1.546
2	10.610	8.280	7.254	8.440
3	10.491	8.317	7.181	8.415
4	5.104	4.100	3.532	4.054
5	2.121	1.779	1.587	1.895
6	335	315	291	306
7	8	118	51	65
Total	30.595	24.426	21.217	24.721

Table 11, Listing Airbnb in the sport ranges.

Listing Airbnb in the **ZTL** ranges	Pre COVID	COVID		Post COVID
	2019	**2020**	**2021**	**2022**
1	8.342	6.815	6.013	6.841
2	9.269	7.473	6.471	7.531
3	5.892	4.592	3.995	4.648
4	3.865	2.987	2.597	3.040
5	2.682	2.007	1.741	2.100
6	305	239	194	272
7	240	313	206	289
Total	30.595	24.426	21.217	24.721

Table 12, Listing Airbnb in the ZTL ranges.

The data in the tables above show how the various properties are distributed within the ranges. Starting from the restricted areas (ZTL, pedestrian areas, etc.), we see that in the initial data analysis year, 2019, range 2 is the most populated with 9269 listings. Additionally, we notice how this number has significantly decreased over the years, dropping to 7473 in 2020, 6471 in 2021, and reaching 7531 in 2022. This trend is repeated for the other 6 ranges as well.

Regarding the metro, degraded areas, and parking, we observe the same situation as mentioned previously, with range 2 leading in 2019 with 13142 properties for the metro, 12235 for degraded areas, and 9453 for parking. This lead is maintained by this range in 2020, 2021, and 2022.

Concerning degraded areas, we can notice how for range 7 (i.e., for apartments located more than 2 km from the nearest degraded area), the number of listings increases significantly from 2019 to 2020 with the onset of Covid (from 73 to 164).

The last area analyzed is related to sports facilities. We observe that ranges 2 and 3 are predominant with 10610 and 10491 listings, respectively. The trend over the years remains very similar to what was observed for other areas, showing a decrease in the number of structures in all ranges except for range 7, where we have seen an increase from 8 to 65 Airbnb between 2019 and 2022.

Seasonality analysis:

As explained in the literature review section, Milan is a city rich in events that attract a high volume of tourists. In this section, we aimed to analyze the seasonality of tourism in the Lombard capital. Tourism seasonality in Milan can be influenced by various factors, including cultural events, weather conditions, and national holidays. Tourism seasonality may also vary based on visitor segments (e.g., business tourism, cultural tourism, shopping) and global tourism industry trends. However, overall, Milan is an attractive destination for tourists throughout the year, offering a rich variety of cultural, artistic, gastronomic, and entertainment experiences.

To conduct these analyses, we used Stata, grouping bookings from various years by month and analyzing which months maximized the performance variables' values. We then created graphs to highlight the results obtained.

Below are the graphs:

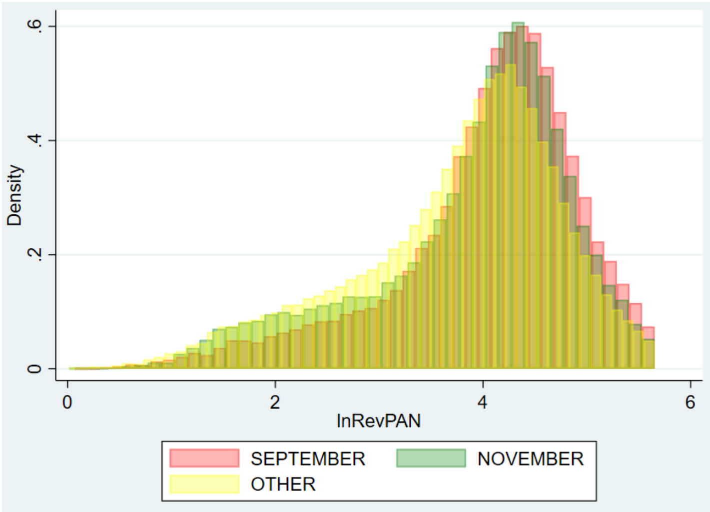


Figure 18, ln(RevPAN)

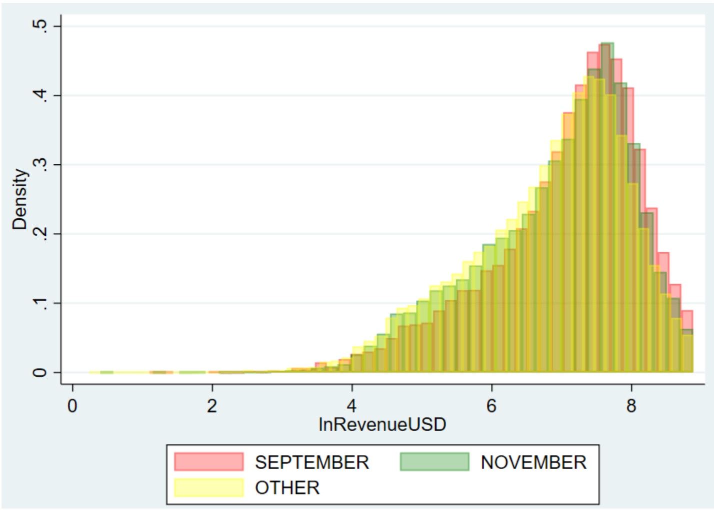


Figure 19, ln(RevenuesUSD)

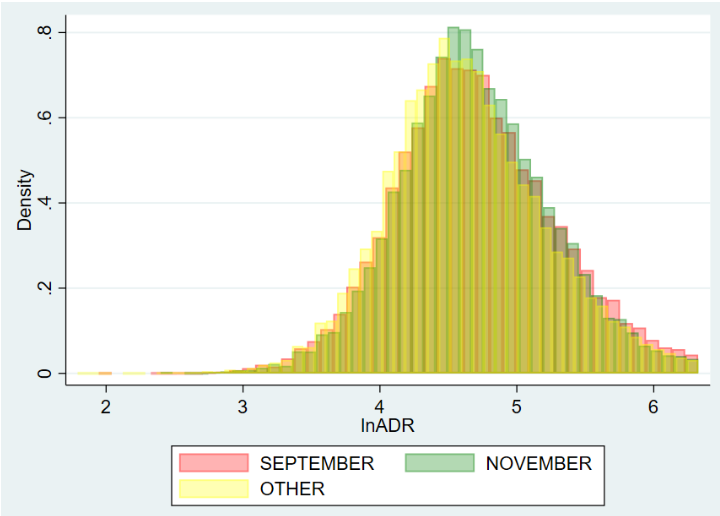


Figure 20, ln(ADR)

In the graphs above, we analyzed the seasonality concerning some performance variables, namely:

- Natural logarithm of RevPAN
- Natural logarithm of Revenues
- Natural logarithm of ADR

We used the natural logarithm of the performance variables because it can offer several advantages, such as better statistical stability, increased linearity in models, and improved interpretability of percentage changes in the data.

For each of these variables, we examined the density in three different periods of the year: September, November, and the remaining months. The choice to separate the months of September and November was driven by the presence of numerous events that occur annually and are potential sources of tourism for Milan.

Here are some of the most remarkable events that take place in those months:

September:

1. Milan Fashion Week: September is the month of fashion in Milan, with Milan Fashion Week showcasing the latest spring/summer or autumn/winter fashion

collections. This event attracts thousands of designers, buyers, journalists, and fashion enthusiasts from around the world.

2. MiArt: This is a fair for modern and contemporary art held in September. MiArt provides a platform for emerging and established artists, art galleries, and collectors to exhibit and purchase artworks.
3. Milan Film Festival: A significant film event held in September, featuring screenings of independent films, short films, and documentaries from around the world.

November:

1. Festival of Science: Organized by the National Museum of Science and Technology Leonardo da Vinci, this festival promotes scientific culture through interactive exhibitions, conferences, workshops, and labs.
2. Milan Wine Week: An event dedicated to wine lovers, where you can taste a wide selection of wines from various Italian and international regions. It takes place in various wine bars and venues across the city.
3. Milan Jazzin' Festival: A jazz festival held annually in November, featuring concerts by national and international jazz artists at various locations throughout the city.

These events, as we can see from the graphs above, influence the performance variables, registering peaks and superior performance in the months of September and November.

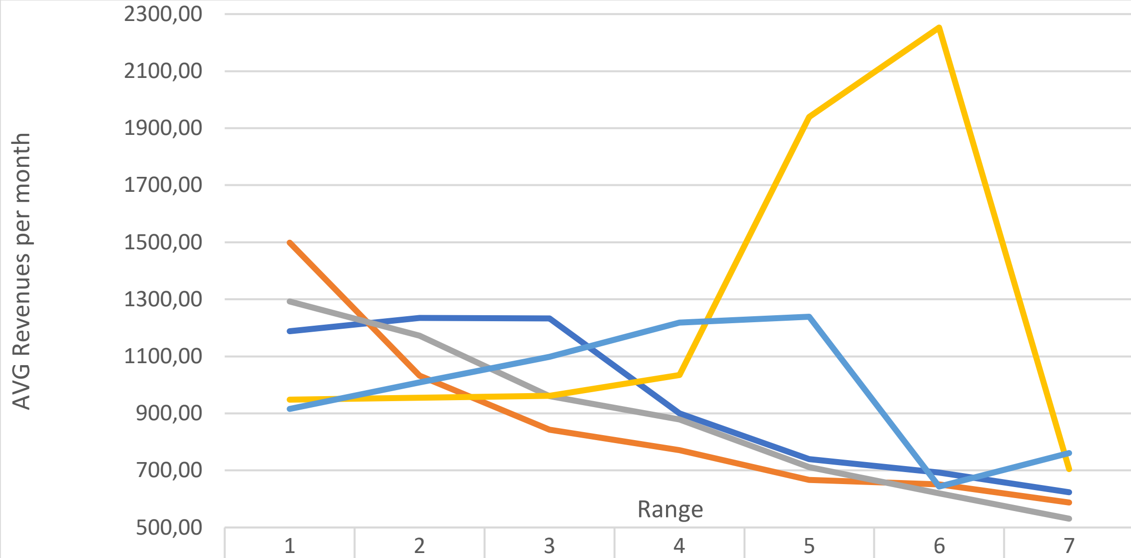
Performance related to the range of distance

The descriptive analysis continued using Stata and Excel in order to understand if the closeness of a "point of interest" affects the performance of an Airbnb in the period considered, indeed from 2019 to 2022.

This analysis has been done using the seven ranges computed and explained in the chapter above. After calculating the presence of those "point of interest" in each NIL, the performances of the Airbnb have been studied in relation to the range of distance from the closest "point of interest" for each of them.

Following an analysis of the performances split by each parameter.

Revenues



	1	2	3	4	5	6	7
Parking	1187,63	1234,74	1233,42	899,85	739,34	692,76	623,52
ZTL	1498,67	1032,40	843,33	771,08	667,09	651,15	587,61
Metro	1292,38	1171,99	960,74	879,08	711,47	620,15	530,85
Sport	948,72	954,62	961,47	1034,70	1939,35	2252,85	704,70
Degradation	915,79	1008,96	1098,88	1217,87	1238,96	643,50	761,13

Figure 21, average revenues per month for each range

In *figure 21*, in the X axe are presented the ranges of distance from the Airbnb and the closest "point of interest", instead in the Y axe the average of the revenues per month of the listing in the same range of distance.

Some macro-trends can be deducted from the graph and the data, indeed, as is possible to imagine, having the Airbnb closer to a public parking, a metro station or also a ZTL or pedestrian area results in a better performance of the listing, to higher revenues. In the next pages will be deducted the fact that enhanced the revenues of these listings. Higher revenues can be associated with higher prices, or a higher number of nights booked by the clients of the platform.

Regarding the Sport facilities there is a pick in correspondence to the range 5 and 6, meaning that listings that are between 1km and 2km from the closest sport facility will get higher revenues than other areas. This data is explained by the fact that in the NIL of the city center there are not sport facilities, the closest facilities to Duomo and Brera are

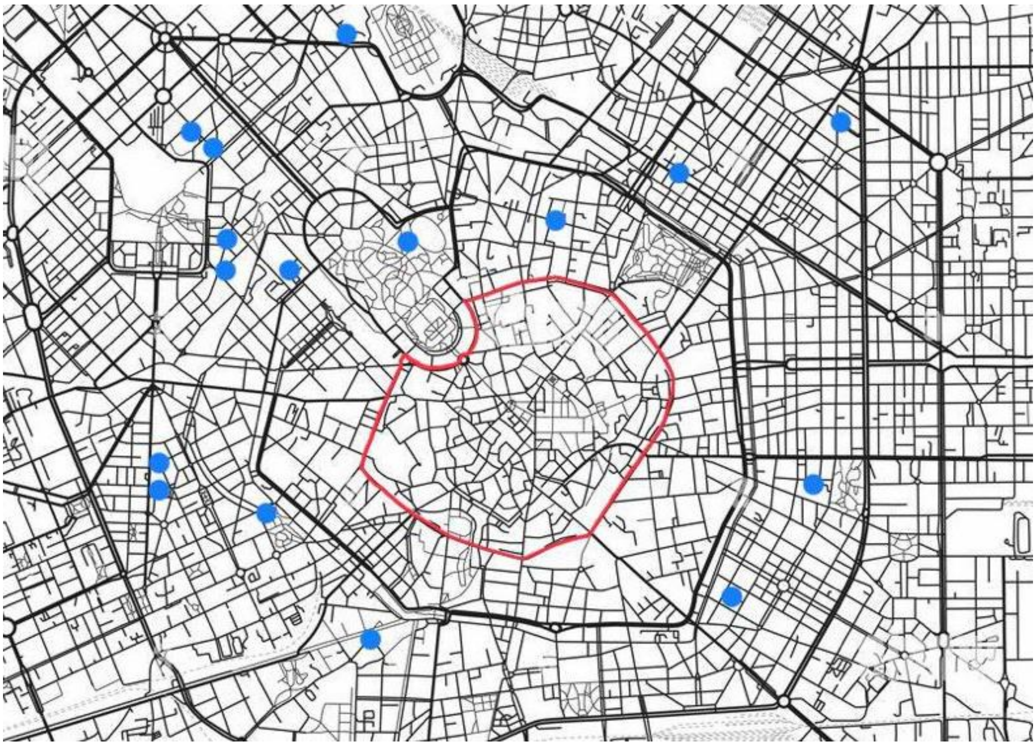


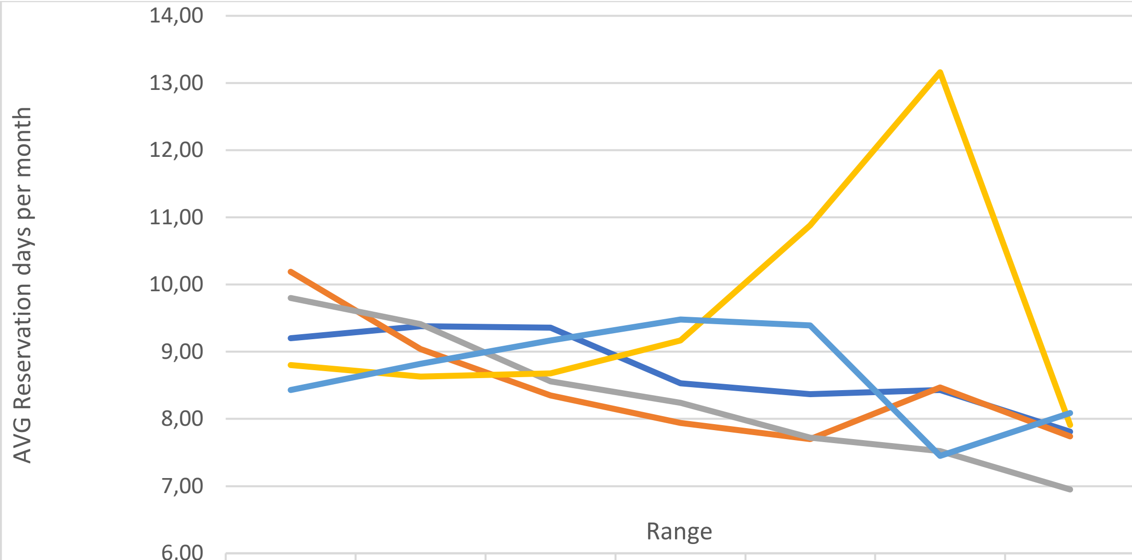
Figure 22, map of the city center of milan with the sports facilities

just outside them.

in the *figure 22* in blue are evidenced the sport facilities that are closer to the city center. In red, instead, the center of Milan, as we can see there are no sport facilities in the center and the distance between the closest and the center is between 1km and 2km explaining the pick in the range 5 and 6. Different things happened for the range 7, instead, where the sports facilities are further distant than 2000m. in this range are present only the listings at the border Milan that also show worst performances.

For the degradation, instead, identifying a trend using the ranges of distance is difficult. Indeed, there is a building in a degradation condition in the NIL “Duomo”, meaning that the listings in “Duomo” and “Brera” are in the range of degradation 1, 2 or 3. What we expected, instead, happened till the range 5, where the revenues increase as we move away from the buildings and area in degradation conditions, an anomalous trend, instead, is shown by the data in the range 6 and 7 where there are the listings further than 1500m from the closest building in degradation conditions and presents the lowest average revenues between every range considered.

Reservation days



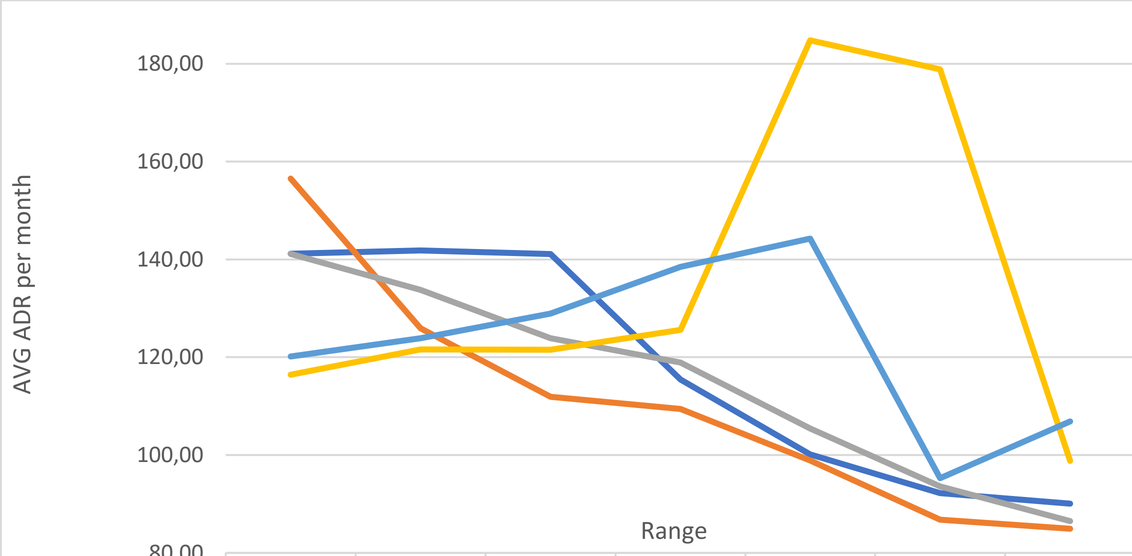
	1	2	3	4	5	6	7
Parking	9,20	9,38	9,36	8,53	8,37	8,43	7,81
ZTL	10,19	9,04	8,35	7,94	7,70	8,47	7,74
Metro	9,80	9,41	8,56	8,24	7,72	7,52	6,95
Sport	8,80	8,63	8,68	9,17	10,88	13,16	7,91
Degradation	8,43	8,82	9,17	9,48	9,39	7,45	8,09

Figure 23, average reservation days per month for each range considered

Figure 23 shows the average reservation days per month, dividing the listings into the ranges explained before. As expected, the reservation days were lowered by moving further from the closest metro station, the closest ZTL and the closest public parking.

For the reservation days the same discourse explained for the revenues applied. Indeed, the pick are in correspondence of the range five and six, where comprehends part of the area of “Duomo” and “Brera”. The same happened also for the degradation, with the reservation days that increases by moving further form the closest degradation area, but for the range 6 and 7.

ADR



Range	1	2	3	4	5	6	7
Parking	141,17	141,84	141,13	115,43	100,14	92,25	90,07
ZTL	156,53	125,89	111,91	109,42	98,87	86,81	84,94
Metro	141,12	133,76	123,86	118,89	105,43	93,58	86,50
Sport	116,40	121,63	121,56	125,56	184,80	178,89	98,81
Degradation	120,14	123,90	128,96	138,46	144,26	95,27	106,85

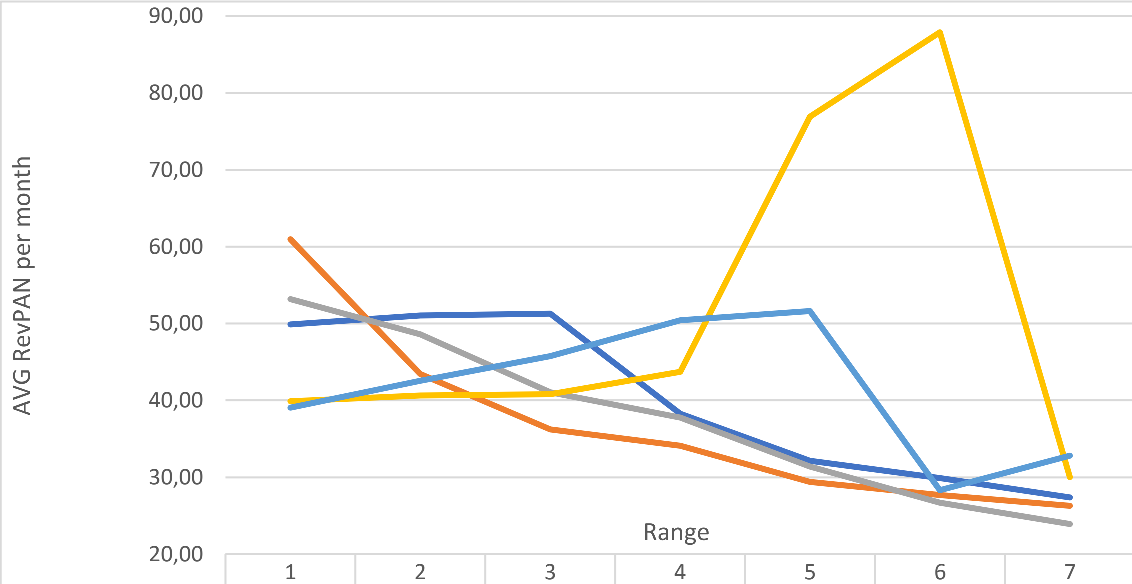
Figure 24, average ADR for each range

In *figure 24* is shown how the ADR (Average daily rate) changed depending on the ranges of distance from each "point of interest". The ratio between the revenues and the days reserved by the client drops for ZTL, metro and public parking by moving further from these points.

For the sport facilities is possible to notice a pick in the range five and six with an average daily rate that overtake 180$ per night. This because the listings in "Duomo" and "Brera" are at a distance between 1km and 2km from the closest sport facility.

The degradation line follows the same as the one for the revenues and reservation days, increasing the average daily rate by moving further from the degradation area, until the range six and seven where there is a drastic drop.

RevPAN



	1	2	3	4	5	6	7
Parking	49,90	51,05	51,29	38,29	32,16	29,89	27,38
ZTL	60,97	43,42	36,23	34,09	29,42	27,71	26,29
Metro	53,18	48,61	41,06	37,76	31,41	26,68	23,92
Sport	39,91	40,64	40,80	43,71	76,91	87,90	30,02
Degradation	39,06	42,56	45,77	50,45	51,63	28,32	32,83

Figure 25, average RevPAN per each range

In *figure 25* is shown how the RevPAN changes depending on the ranges of distance. As before the line of parking, ZTL and metro drop by going further from them. The two picks for the sport facility are still present in range 5 and 6, with the pick in range 6 that exceeds the double of range 1, 2, 3 and 7, showing the importance of having an Airbnb in the city center. While the degradation follows the same path until range six where has a drastic drop. As shown in the picture below, the degradation points of the city of Milan are distributed all over the territory.

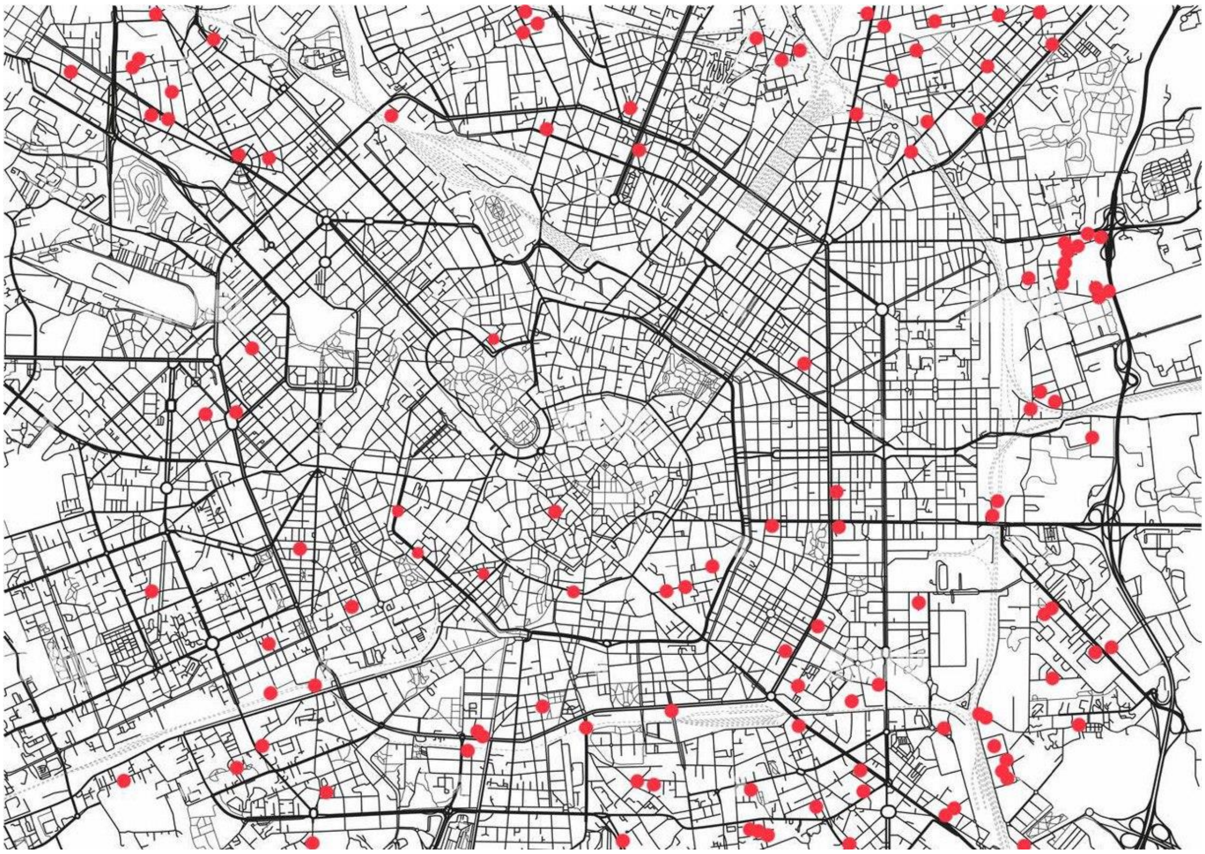
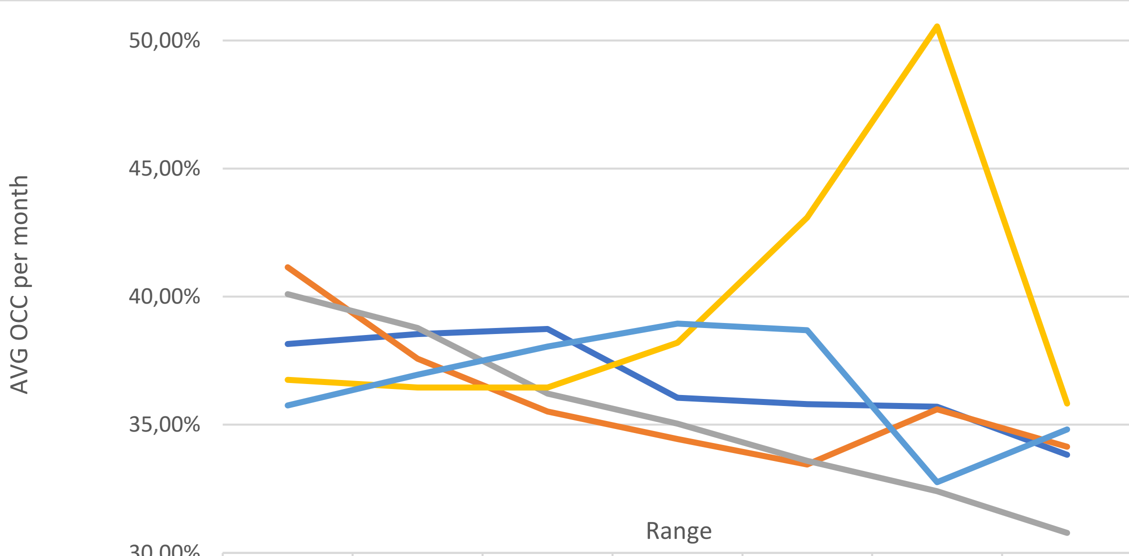


Figure 26, map of the degradation buildings and area of Milan.

OCC



Range	1	2	3	4	5	6	7
Parking	38,16%	38,54%	38,74%	36,06%	35,80%	35,71%	33,83%
ZTL	41,15%	37,58%	35,52%	34,45%	33,45%	35,60%	34,15%
Metro	40,10%	38,78%	36,22%	35,04%	33,60%	32,41%	30,78%
Sport	36,75%	36,45%	36,46%	38,21%	43,09%	50,55%	35,83%
Degradation	35,75%	36,96%	38,06%	38,95%	38,69%	32,76%	34,82%

Figure 27, average occupancy rate for each range

In *figure 27* the trend of the occupancy rate shows that going further from the closest metro station, the closest ZTL or pedestrian area and the closest public parking the reservation days drop with respect to the number of available days to be rented. In particular the metro ling evidence a drastic drop of the occupation rate from 40.10% of the range 1 (so being closer than 200m to the closest metro station) to 30.78% of range 7, having a metro station further than 2000m.

The other trends are similar to what has been explained before, with the occupancy rate of the city center to more than 50% that is reflected in the pick at range 5 and 6 of the sport facility.

Regression analysis

Regression analysis is a fundamental statistical technique used to examine relationships between variables. In the context of this thesis, regression analysis will be a crucial methodological tool to further explore and better understand the factors influencing the performance of Airbnb hosts in the city of Milan.

The regression analysis will be presented as an essential methodological tool to address research questions and test formulated hypotheses. The main objective will be to identify significant variables that impact Airbnb host performance and understand the nature of these relationships.

Specifically, we will use regression analysis to examine how qualitative variables such as the presence of subways, degraded areas, parking facilities, Restricted Traffic Zones (ZTL), and sports facilities influence Airbnb host performance. Through the application of regression models, we will explore the magnitude and direction of these relationships, identifying factors that have a significant impact on performance and providing an empirical basis for better understanding the functioning of the short-term rental market in the city of Milan.

Regression analysis will also be presented as a tool to evaluate the effectiveness of any interventions or policies that may be proposed to improve Airbnb host performance or better manage the short-term rental market in the city. Through the interpretation of estimated coefficients and analysis of results, we will provide valuable insights for stakeholders in the sector, such as hosts, tourism operators, and local authorities.

In summary, regression analysis plays a key role in the methodological approach of my thesis, offering a robust statistical framework to examine complex relationships between variables and contributing to the overall understanding of the Airbnb rental phenomenon in the city of Milan.

Univariate analysis

Univariate regression analysis aims to explore the relationship between two variables, where one variable (the independent variable) is used to predict or explain changes in another variable (the dependent variable). In the thesis, univariate regression analysis will serve as a preliminary step to investigate the relationship between each independent variable (such as proximity to the metro, parking availability, restricted traffic areas, sports facilities, and degraded areas) and the dependent variables representing Airbnb host performance metrics (such as RevPAN, Revenues, Reservation Days, ADR, and OCC). Through univariate regression analysis, we aim to assess the strength and direction of the relationship between each independent variable and performance metrics, individually. Additionally, univariate regression analysis will establish a foundation for subsequent multivariate regression analysis, where we will examine the combined influence of multiple independent variables on Airbnb host performance while controlling for any confounding factors. The equations referenced, with a linear-linear univariate regression model and a log-linear model, are as follows:

$$RevPAN_i = \alpha + \beta \, Distance_i$$

$$lnRevPAN_i = \alpha + \beta \, Distance_i$$

In this case, an example of equations with RevPAN as the dependent variable has been provided, and the equations remain the same for all other performance variables, except for OCC, where only the linear-linear model is used since it wouldn't make sense to use the logarithmic model since it is already a percentage. It is also specified that the independent variable Distance$i$ is generically presented to provide an example of these equations, but during the analysis, it is appropriately replaced with all distance variables from the various categories analyzed (metro, degradation, ZTL, sports facilities, and parking).

Revenues:

As previously anticipated, it is decided to study the relationship between the dependent variable Revenues and the independent variable Distance, repeated for all categories (metro, ZTL, degradation, sports facilities, and parking), first with a linear model (LIN-LIN), and subsequently with a logarithmic model (LOG-LIN), as summarized in the tables below:

Revenues USD	Cons	Beta	P value	R^2
Metro	1753.738	-0.4716	0.000	0.0203
Degradation	1366.144	0.2637	0.000	0.0038
ZTL	1824.065	-0.6695	0.000	**0.0405**
Parking	1761.951	-0.3343	0.000	0.0278
Sport facility	1084.988	0.6937	0.000	0.0267

Table 13, Univariate regression for the dependent variable Revenues

Ln(Revenues USD)	Cons	Beta	P value	R^2
Metro	7.0266	-0.0004	0.000	0.0162
Degradation	6.7412	0.0002	0.000	0.0027
ZTL	7.071	-0.0005	0.000	**0.0300**
Parking	7.0368	-0.0003	0.000	0.0231
Sport facility	6.5770	0.0004	0.000	0.0146

Table 14, Univariate regression for the dependent variable Ln(Revenues)

The equations referred to, with a linear-linear univariate regression model and with a log-linear model, are as follows:

$$Revenues_i = \alpha + \beta \, Distance_i$$

$$\ln (Revenues)_i = \alpha + \beta \, Distance_i$$

The results of the LIN-LIN model are as follows:

$$Revenues = 1753{,}738 - 0{,}4716 \, DistanceMetro$$

$$Revenues = 1366{,}144 + 0{,}2637 \, DistanceDegradation$$

$$Revenues = 1824{,}065 - 0{,}6695 \, DistanceZTL$$

$$Revenues = 1761{,}951 - 0{,}3343 \, DistanceParking$$

$$Revenues = 1084{,}988 + 0{,}6937 \, DistanceSport$$

Observing the first table above, we can notice that the t-value is less than -2.58 ($p < 1\%$), indicating the rejection of the null hypothesis, meaning that the coefficient β is statistically significant. Therefore, taking the example of the Metro category, a unit increase in the distance between the property and the nearest metro station implies a decrease of 0.4716 in the value of the dependent variable Revenues. This result confirms our initial hypothesis that having a metro station near the property positively influences its performance. Looking at the model's precision, the variance explained by it is relatively low (R^2 of 0.0203, i.e., 2.03%). However, by using a continuous variable, we have improved the significance compared to the same regression conducted with a dummy independent variable, which would have provided less information. To achieve a more significant model, as we will see later, it is necessary to combine multiple variables.

Additionally, it can be noted that a similar relationship to that of the metro is observed for parking lots and restricted traffic areas (ZTL), meaning that an increase in the distance for these categories leads to a reduction in the value of Revenues. Conversely, for degradation areas and sports facilities, the relationship is opposite, meaning that an increase in distance leads to an increase in Revenues. Particularly, an increase of one unit in the distance to a sports facility results in Revenues increasing by 0.6937 dollars. This latter result contradicts our initial hypotheses, so we will evaluate if this trend reoccurs in subsequent regressions.

A significant value is observed in the variance explained by the model with the distance from restricted traffic areas (ZTL), which is relatively high (R^2 of 0.0405, i.e., 4.05%).

Similarly, the results of the LOG-LIN model are as follows:

$$ln\ (Revenues) \ = \ 7{,}0266 - \ 0{,}0004\ DistanceMetro$$

$$ln\ (Revenues) \ = \ 6{,}7412 \ + \ 0{,}0002\ DistanceDegradation$$

$$ln\ (Revenues) \ = \ 7{,}0710 - \ 0{,}0005\ DistanceZTL$$

$$ln\ (Revenues) \ = \ 7{,}0368 - \ 0{,}0003\ DistanceParking$$

$$ln\ (Revenues) \ = \ 6{,}5770 \ + \ 0{,}0004\ DistanceSport$$

Observing the second table above, we can also note that the t-value is less than -2.58 ($p < 1\%$), indicating the rejection of the null hypothesis, meaning that the coefficient β is statistically significant at the 99% level. Therefore, taking the example of the Metro category, a unit increase in the distance between the property and the nearest metro station implies a decrease of 0.04% in the value of the dependent variable Revenues. The trends, even for the other categories, remain similar and consistent with those of the linear-linear model. Comparing the value of R^2, we can see that it has decreased from 2.03% to 1.62%. As observed, this trend repeats all analyzed categories. One explanation could be that the logarithmic model might not be fully capturing the relationship between the variables, leading to an information loss compared to the linear model.

Reservation days:

As the second univariate regression model, we analyzed the relationship between the dependent variable Reservation Days and the independent variable Distance, repeated for all categories (metro, ZTL, degradation, sports facilities, and parking), first with a linear model (LIN-LIN), and subsequently with a logarithmic model (LOG-LIN), as summarized in the tables below:

	Cons	Beta	P value	R^2
Metro	14.8589	-0.0014	0.000	0.0033
Degradation	13.8542	0.0005	0.000	0.0003
ZTL	14.9122	-0.0017	0.000	0.0046
Parking	14.5994	-0.0006	0.000	0.0017
Sport facility	13.0981	0.0017	0.000	0.0031

Table 15, Univariate regression for the dependent variable Reservation days

	Cons	Beta	P value	R^2
Metro	2.4808	-0.0001	0.000	0.0029
Degradation	2.3933	0.00005	0.000	0.0003
ZTL	2.4865	-0.0001	0.000	0.0042
Parking	2.4604	-0.00005	0.000	0.0016
Sport facility	2.3336	0.0001	0.000	0.0027

Table 16, Univariate regression for the dependent variable Ln (Reservation days)

The results of the LIN-LIN model are as follows:

$$Reservation\ days \ = \ 14{,}8589 - \ 0{,}0014\ DistanceMetro$$

$$Reservation\ days \ = \ 13{,}8542 \ + \ 0{,}0005\ DistanceDegradation$$

$$Reservation\ days \ = \ 14{,}9122 - \ 0{,}0017\ DistanceZTL$$

$$Reservation\ days \ = \ 14{,}5994 - \ 0{,}0006\ DistanceParking$$

$$Reservation\ days \ = \ 13{,}0981 \ + \ 0{,}0017\ DistanceSport$$

Observing the first table above, we can note that the t-value is less than -2.58 ($p < 1\%$), indicating the rejection of the null hypothesis, meaning that the coefficient β is statistically significant. Therefore, taking the example of the Metro category again, a unit increase in the distance between the property and the nearest metro station implies a decrease of 0.0014 days in the dependent variable Reservation Days. This result confirms our initial hypothesis that having a metro station near the property positively influences its performance, in this case, increasing the number of days reserved by guests. Looking at the precision of the model, the variance explained by it is very low (R^2 of 0.0033, i.e., 0.33%). This indicates that this performance variable is not strongly influenced by the independent variable studied. In summary, we can say that the distance does not greatly influence the duration of guests' stays, unlike what we saw in the previous model regarding customers' willingness to pay.

Additionally, a similar relationship to that of the metro is observed for parking lots and restricted traffic areas (ZTL), meaning that an increase in the distance for these categories leads to a reduction in the value of Reservation Days. Conversely, for degradation areas and sports facilities, the relationship is opposite, meaning that an increase in distance leads to an increase in Reservation Days. Particularly, an increase of one unit in the distance to a sports facility results in Reservation Days increasing by 0.0017 days. This latter result contradicts our initial hypotheses again. As seen for the metro, the R^2 value for all these categories is very low, never exceeding 0.42%.

Similarly, the results of the LOG-LIN model are as follows:

$$ln\ (Reservation\ days)\ =\ 2{,}4808 - \ 0{,}0001\ DistanceMetro$$

$$ln\ (Reservation\ days)\ =\ 2{,}3933\ +\ 0{,}00005\ DistanceDegradation$$

$$ln\ (Reservation\ days)\ =\ 2{,}4865 - \ 0{,}0001\ DistanceZTL$$

$$ln\ (Reservation\ days)\ =\ 2{,}4604 - \ 0{,}00005\ DistanceParking$$

$$ln\ (Reservation\ days)\ =\ 2{,}3336\ +\ 0{,}0001\ DistanceSport$$

Observing the second table above, we can also note that the t-value is less than -2.58 (p < 1%), indicating the rejection of the null hypothesis, meaning that the coefficient β is statistically significant at the 99% level. Therefore, taking the example of the Metro category again, a unit increase in the distance between the property and the nearest metro station implies a decrease of 0.01% in the value of the dependent variable Reservation Days. The trends, even for the other categories, remain similar and consistent with those of the linear-linear model. Comparing the value of R^2, we can see that it has further decreased from 0.33% to 0.29%, making this model even less significant.

ADR

Following an example of the formula used in the univariate regression analysis of the average daily rate:

$$ADR_i = \alpha + \beta\ Distance_i$$

$$\ln\ (ADR)_i = \alpha + \beta\ Distance_i$$

For the ADR, in the univariate regression, the correlation between the variable and the distance to the closest "point of interest" was studied with two criteria: the LIN-LIN and the LOG-LIN one.

The variable "Distance" reported before refers to the distance, in meters, between each listing and the closest "point of interest". Indeed, for each dependent variable five univariate regressions were made, one for each "point of interest": metro station, degradation area, public parking, ZTL/pedestrian area, and sport facility.

	Cons	beta	p value	R^2
Metro	136.0909	-0.0283	0.0000	0.0203
Degradation	136.0909	-0.0283	0.0000	0.0203
ZTL	141.0142	-0.0417	0.0000	**0.0438**
Parking	139.0923	-0.0233	0.0000	**0.0375**
Sport facility	95.4257	0.0423	0.0000	0.0283

Table 17, Univariate regression for the dependent variable ADR

Table 17 shows the correlations between the ADR and each independent variable using a LIN-LIN correlation. For each of these analyses the P-value is below 1%, so we can discard the null hypothesis, stating that the β is significative.

The β in each regression is negative, but for the correlation with the sport facility. A negative value of β means that for each meter that the listing is further from the closest "point of interest" the ADR drops for the value of β. In particular, the most negative β is the one of the ZTL, this means that the more we are closer to a ZTL and the more the ADR will be high, instead for each meter going further from the closest ZTL will reduce by 0.0417 [$ per night].

Looking at the precision of the model, the variance explained by it appears to be very low, going from an R^2 of 0.0203 to at a maximum of 0.0438, this because only one single variable at each time has been used, indeed for a better study of the model in the next chapter the variables are studied together in the multivariate regression.

	Cons	beta	p value	R^2
Metro	4.7651	-0.0002	0.0000	0.0293
Degradation	4.7651	-0.0002	0.0000	0.0293
ZTL	0.6051	-0.0000333	0.0000	**0.0611**
Parking	4.7980	-0.0002	0.0000	**0.0596**
Sport facility	4.4442	0.0003	0.0000	0.0308

Table 18, Univariate regression for the dependent variable ln(ADR)

In *table 18* the LOG-LIN regression model for the ADR is shown. This kind of model, using logarithms, focused its attention on how the dependent variable changes in % for each unitary change in the independent variable, which in our case are the distances to the closest "point of interest". This model shows again the negative correlation between going further from the first four points of interest as a percentage, explaining in a better way the model with a R^2 higher than in the LIN-LIN analysis.

The opposite trend, instead, characterizes the correlation with the sport facilities. Going further of 1 meter from the closest sport facility led to an increase in the ADR of 0.03%. This lets us understand that having a sport facility closer to your Airbnb is correlated to having a lower ADR.

OCC

Following an example of the formula used in the univariate regression analysis of the average daily rate:

$$OCC_i = \alpha + \beta \; Distance_i$$

The Occupancy rate, in the univariate regression, was studied only with the LIN-LIN model and not the LOG-LIN one, because by being a percentage value its logarithm is meaningless. The distance refers to each single distance from the closest "point of interest", which change for each of them.

	coeff	beta	p value	R^2
Metro	0.6059	-0.0001	0.0000	0.001
Degradation	0.5852	7.95e-06	0.0000	0.0001
ZTL	0.6051	-0.0000333	0.0000	0.0016
Parking	0.5992	-0.0000127	0.0000	0.0006
Sport facility	0.5677	0.0000355	0.0000	0.0011

Table 19, Univariate regression for the dependent variable OCC

Table 19 shows the correlations between the dependent variable, which in this case is the occupancy rate, and each distance from the closest "point of interest" one at each time.

The P-value for all these analyses was below 1%, so is possible to consider the β as a significative value. The β of these analyses were so low, negative for the metro stations, the ZTL and the public parking meaning that moving further from them led to a decrease in the occupancy rate of the listings. The opposite, instead, happens for the degraded areas and the sports facilities, having a positive beta show that being closer to them led to a lower occupancy rate to the listing.

Looking at the precision of the model, the variance explained by it appears to be very low, going from an R^2 of 0.0006 to at a maximum of 0.0016, this is because only one single variable at each time has been used.

RevPAN

Following an example of the formula used in the univariate regression analysis of the average daily rate:

$$RevPAN_i = \alpha + \beta \ Distance_i$$

$$\ln (RevPAN)_i = \alpha + \beta \ Distance_i$$

For the RevPAN, in the univariate regression, the correlation between the variable and the distance to the closest "point of interest" was studied with two criteria: the LIN-LIN and the LOG-LIN one. The distance refers to each single distance from the closest "point of interest", which change for each of them.

	Cons	beta	p value	R^2
Metro	72.7664	-0.0176	0.0000	0.0191
Degradation	58.3889	0.0097	0.0000	0.0035
ZTL	75.3575	-0.0249	0.0000	**0.0380**
Parking	73.4366	-0.0129	0.0000	0.0282
Sport facility	47.8902	0.0257	0.0000	0.0249

Table 20, Univariate regression for the dependent variable RevPAN

Table 20 shows the correlation between RevPAN (dependent variable) and as independent variable the distance in meters to the closest "point of interest".

The P-value is below 1%, indeed is possible to consider the β significative. Having a negative β, going further from the closest metro station, ZTL/pedestrian area and public parking show a lower RevPAN as supposed before. The exact opposite happens, instead, for the degradation area and the sport facilities.

	Cons	beta	p value	R^2
Metro	3.9430	-0.000299	0.0000	0.0146
Degradation	3.7070	0.0001491	0.0000	0.0022
ZTL	3.9809	-0.0004105	0.0000	0.0273
Parking	3.9587	-0.0002257	0.0000	0.0227
Sport facility	3.5623	0.0003687	0.0000	0.0136

Table 19, Univariate regression for the dependent variable ln(RevPAN)

In *table 19*, the same correlation has been studied using logarithms, showing a better explanation of the model (with higher R^2). The signs, obviously, remained the same, with the beta now showing the change of the RevPAN as a percentage and no more as a numerical value as before.

The P-values are all below the limit of 1% letting us consider the betas as significative values.

Multivariate Analysis

For the multivariate regression 3 different models were studied. In all three models as dependent variables were used the performance variables of the Airbnb and their natural logarithms, as control variables the ones explained before and ad independent variables the distance ones computed before. The differences between the three analysis that were done wat that in the first one has been only studied the correlation between the variables described before, instead in the second was also introduced the correlation with the long-term variable and finally in the third the implementation of the Superhost variable.

These multivariate analyses have been done to understand the correlation, if any, between the dependent variables and the independent ones, also introducing the control variables in order to study a model that can better explain the phenomenon.

Each analysis began with the first model where only the control variables were studied. In the following formula is reported an example of the M1 model with the RevPAN as dependent variable.

$$RevPAN_{i,t} = \alpha + \beta_1 MaxGuests + \beta_2 EntireHome + \beta_3 IstantBook \\ + \beta_4 SuperHost + \beta_5 LTR + \beta_6 NIL + \beta_7 Year + \beta_8 Month$$

The same happened also for the LOG-LIN study where the RevPAN as independent variable was substituted with the ln (RevPAN).

$$lnRevPAN_{i,t} = \alpha + \beta_1 MaxGuests + \beta_2 EntireHome + \beta_3 IstantBook \\ + \beta_4 SuperHost + \beta_5 LTR + \beta_6 NIL + \beta_7 Year + \beta_8 Month$$

Then the models M2, M3, M4, M5 and M6 were studied introducing, in each single one, the correlation with a distance variable computed before. In depth:

- M2: correlation with the distance to the closest metro station
- M3: correlation with the distance to the closest degradation area
- M4: correlation with the distance to the closest ZTL/pedestrian area
- M5: correlation with the distance to the closest sport facility
- M6: correlation with the distance to the closest public parking

Finally in the Model M7 all those variables were implemented arriving to the final model with all the control variables and all the distance variables as independent ones.

All these analyses described have been done for all the performances variables as dependent and for their logarithms, with the exception for the occupancy rate which, being a value expressed as a percentage, its logarithms would be meaningless.

Revenues

The model has been studied for the revenues and the natural logarithm of them, showing in M1 the correlation between the first model, having only the control variables described before. The control variable of MaxGuests, entire apartment, instantbook and superhost show a positive correlation with the dependent variables “Revenues” and “ln(Revenues)”, the opposite, instead, happen for the long term rent, associated with a decrease of the revenues.

In the following models (M2, M3, M4, M5 and M6) have been studied these variables of control plus one single variable of distance computed before. Recalling that these variables of distance computed the distance in meters from a listing to the closest “point of interest”, two of them shows a positive correlation, meaning that for each meter that a listing is further from the closest “degradation area” and “sport facility” its revenues are higher. In depth for each meter going further from the closest degradation area, the revenues increase of 0.145$ per month and for each meter going further from the closest sport facility the revenues increase of 0.053$ per month, all these values are strengthened by the P-value being <1%. The opposite trends are shown by the metro stations, the ZTL or pedestrian areas and finally by the public parking. These models show a negative correlation between going further from these “point of interests” and the revenues of the listings, indirectly is shown also that getting closer to one of them is correlated with better performances. For each meter getting closer to the closest metro stations the revenues will be 0.184$ per month higher, for each meter getting closer to the closest ZTL or pedestrian areas the revenues will be 0.154$ per month higher and finally by getting 1 meter closer to the closest public parking the revenues show an increase of 0.125$ per month. These values can be considered significative due to the fact that their P-values are below 1%.

The models referring to the logarithm of the revenues do not show in a clear way the trends because we are considering how the revenues change by moving meter by meter, so study the change in percentage of them didn’t result in the best way to study the model.

In the final model (M7) all the variables of control and all the independent variables have been grouped together in order to study a final model complete of all the variables. In the tables attached at the end of the paper all these tables are grouped to show in a better way what happened in the model.

ADR

The model was studied for the variable ADR and its natural logarithm, showing in M1 the correlation between the performance variable and general variables. In this model, a positive correlation is observed between the control variables MaxGuests, Entire Apartment, and LTR with the dependent variables ADR and ln (ADR), while the opposite occurs for Instantbook and Superhost, associated with a decrease in the value of ADR.

In the subsequent models (M2, M3, M4, M5, and M6), we added one by one the calculated distance variables (Metro, degradation, ZTL, sports facilities, and parking). Two of these distance variables show a positive correlation, meaning that for every meter an ad moves away from the nearest "degradation area" and "sports facility," the ADR value increases. In particular, for every meter moving away from the nearest degradation area, revenues increased by $0.011 per month, and for every meter moving away from the nearest sports facility, revenues increased by $0.005 per month, all of these values are supported by a p-value <1%. Conversely, opposite trends are recorded for metro stations, restricted traffic areas (ZTL) or pedestrian areas, and public parking. These models show a negative correlation between moving away from these "points of interest" and the ADR of listings, indirectly demonstrating that getting closer to one of them is correlated with better performance. For every meter approaching the nearest metro stations, revenue will increase by $0.014 per month, for every meter approaching the nearest ZTL or pedestrian areas, revenue will increase by $0.011 per month, and finally, approaching the nearest metro station by 1-meter, public parking will result in a revenue increase of $0.011 per month. These values can be considered significant since their p-values are less than 1%.

However, the Log-Lin models, still for the ADR performance variable, do not clearly show trends. This is undoubtedly since we are considering how revenues change meter by meter, so studying their percentage variation is not very meaningful, given that even the variations in the linear model are very small for this variable.

In the final model (M7), all control variables and independent variables were grouped together to study a comprehensive model with all variables.

OCC:

The model was studied for the OCC variable, showing in M1 the correlation between the performance variable and general variables. From this regression, we can observe a positive correlation between the control variables Instantbook, Entire Apartment, and Superhost with the dependent variable OCC, while the opposite occurs for MaxGuests and LTR, associated with a decrease in the value of OCC.

In the subsequent models (M2, M3, M4, M5, and M6), we added one by one the calculated distance variables (Metro, degradation, ZTL, sports facilities, and parking). These distance variables calculate the distance in meters from a listing to the nearest "point of interest." Observing the obtained data, we can see that the Beta values are equal to zero, suggesting that there is no significant relationship between this variable and the dependent variable (performance) when considering the distance from the points of interest. The fact that OCC (Occupancy Rate) is a percentage could significantly influence the relationship with the distance from the points of interest, especially when considering the variation meter by meter. Therefore, the occupancy rate may not vary significantly for each additional or removed meter of distance from the points of interest, especially considering that we are analyzing a very large area. We can notice that for all values within this regression, the p-value is <1%.

In the final model (M7), all control variables and independent variables were grouped together to study a comprehensive model with all variables, but we obtained the same data as in the previous models.

RevPAN:

The model was studied for the RevPAN variable and its natural logarithm, showing in M1 the correlation between the performance variable and general variables. In this model, we can observe a positive correlation between all control variables MaxGuests, Entire Apartment, LTR, Instantbook, and Superhost with the dependent variable RevPAN. However, considering ln(RevPAN), the opposite occurs for the control variable LTR. In fact, for the Lin-Lin model, a positive coefficient of 3.722 is found, while in the Log-Lin model, a value of -0.088 is obtained. These values are supported by a p-value < 1%.

In the subsequent models (M2, M3, M4, M5, and M6), we added one by one the calculated distance variables (Metro, degradation, ZTL, sports facilities, and parking). Two of these distance variables show a positive correlation, meaning that for every meter an ad moves away from the nearest "degradation area" and "sports facility," the value of RevPAN increases. In particular, for every meter moving away from the nearest degradation area, the value of RevPAN increases by $0.005 per month, and for every meter moving away from the nearest sports facility, revenues increase by $0.002 per month, all of these values are supported by a p-value <1%. Conversely, opposite trends are recorded for metro stations, restricted traffic areas (ZTL) or pedestrian areas, and public parking. These models show a negative correlation between moving away from these "points of interest" and the RevPAN of listings, indirectly demonstrating that getting closer to one of them is correlated with better performance. For every meter approaching the nearest metro stations, revenue will increase by $0.007 per month, for every meter approaching the nearest ZTL or pedestrian areas, revenue will increase by $0.005 per month, and finally, approaching the nearest metro station by 1 meter, public parking will result in a revenue increase of $0.004 per month. These values can be considered significant since their p-values are less than 1%.

However, the Log-Lin models, still for the RevPAN performance variable, do not clearly show trends. This is undoubtedly due to the fact that we are considering how revenues change meter by meter, so studying their percentage variation is not very meaningful, given that even the variations in the Lin-Lin model are very small for this variable.

In the final model (M7), all control variables and independent variables were grouped together to study a comprehensive model with all variables.

Multivariate regression with SuperHost

Following the previously analyzed multivariate regression models, in this paragraph, the performance variables of Airbnb listings were examined in relation to the SuperHost status to verify if this variable has an interactive effect with other variables in determining the dependent variable. Indeed, by multiplying this dummy variable by the variables representing distances from points of interest, we can verify if and how being a SuperHost has additional effects on them.

For these regressions, 6 models were identified:

M1: Classic introductory model that relates the dependent performance variable with control variables such as Max Guest, Entire Apartment, Instantbook, SuperHost, and LTR, always comparing listings for each NIL, Year and Month.

M2-M3-M4-M5-M6: In addition to the control variables used in M1, two new variables are introduced, one being the distance variables and the second being the product between the previously mentioned distance variable and the SuperHost dummy variable. This is repeated for each model from M2 to M6, always changing the type of point of interest (metro, degradation, ZTL, sports facilities, and parking).

Observing the results, we can see that regarding the OCC variable, we obtain the same results as in the previous regressions, i.e., a beta equal to 0 for all models. As seen previously, this could be related to the fact that, being a percentage value linked to a very large area, a change of one meter in the distance between the structure and the various reference points may not be significant enough to vary the occupancy rate by one percentage point.

Regarding the variable Revenues and its natural logarithm, we can observe the following:

Regarding the distance from the metro, we see a beta value of -0.178 $ per month, meaning that an increase of one meter in the distance between the listing and the point of interest decreases the Revenues by $0.178. Additionally, we see a beta value of -0.029 $ per month when considering the variable Distance*SuperHost. This suggests that being a SuperHost makes the distance between the nearest metro and the listing even more significant. Specifically, it implies that an increase of one meter in the distance between the listing and the point of interest decreases the Revenues by $0.178 + $0.029 in the case of a SuperHost. All these values are supported by a p-value <1%. This trend is also repeated for ZTL, sports facilities, and parking.

Regarding model M3, which pertains to the distance from the nearest degradation area, we see a beta value of 0.150 $ per month, indicating that an increase of one meter in the distance between the listing and the point of interest increases the Revenues by $0.150. Additionally, we see a beta value of -0.024 $ per month when considering the variable Distance*SuperHost. This suggests that being a SuperHost makes the distance from the nearest degradation area to the structure less significant. Specifically, it implies that an increase of one meter in the distance between the listing and the point of interest increases the Revenues by $0.150 - $0.024 in the case of a SuperHost. However, it is notable that for the variable Distance*SuperHost, the p-value is > 1%. This may indicate that the variable does not significantly contribute to explaining the variation in the dependent variable, or there may be additional heterogeneity or noise not captured by the variables considered in the model.

As for the natural logarithm of Revenues, we observe that the beta coefficients in models M2 to M6 are zero for the distance variables and Distance*SuperHost. This is because we are considering how revenues change meter by meter, so studying the percentage variation is not very meaningful, given that even the variations in the linear model are very small for this variable.

Moving on to analyze the trend of the ADR variable and its natural logarithm, concerning the distance from the metro, we observe a beta value of -0.014 $ per month, indicating that an increase of one meter in the distance between the listing and the point of interest decreases the ADR by $0.014. Additionally, we see a beta value of 0.000 $ per month when considering the variable Distance*SuperHost, suggesting that being a SuperHost does not further influence the trend of this performance variable. For models M4 and M5 (ZTL and sports facilities), we observe that the beta coefficients related to the regression with the distance variable and with the latter multiplied by the SuperHost dummy have the same sign. For instance, in model M4, we see a beta value of -0.011 $

per month, meaning that an increase of one meter in the distance between the listing and the point of interest decreases the ADR by $0.011, and we see a beta value of -0.001 $ per month when considering the variable Distance*SuperHost. This implies that an increase of one meter in the distance between the listing and the point of interest decreases the ADR by $0.011 + $0.001 in the case of a SuperHost. All these values are supported by a p-value <1%. However, for models M3 and M6 (Degradation and Parking), we observe that the beta coefficients related to the regression with the distance variable and with the latter multiplied by the SuperHost dummy have opposite signs. Taking model M3 as an example, we see a beta value of 0.012 $ per month, indicating that an increase of one meter in the distance between the listing and the degradation area increases the ADR by $0.012, and we see a beta value of -0.003 $ per month when considering the variable Distance*SuperHost. This implies that an increase of one meter in the distance between the listing and the degradation area increases the ADR by $0.012 - $0.003 in the case of a SuperHost. Therefore, we can notice that being a Superhost slightly attenuates the effect on the distance from the degradation area to the structure, but the trend remains the same.

Regarding the natural logarithm of ADR, we observe that the beta coefficients in models M2 to M6 are zero for the distance variables and distance * SuperHost, repeating the same situation identified for the performance variables listed previously.

Finally, the last variable analyzed is RevPAN and its natural logarithm. In this case, we can observe that there is no disagreement in signs between the coefficients obtained from the regressions with the distance variable and with the latter multiplied by the SuperHost dummy. Specifically, for models M2, M4, M5, and M6, we see that being a SuperHost further reinforces the relationship between the distance and the RevPAN performance variable. For example, in model M2, we observe a beta value of -0.006 $ per month, indicating that an increase of one meter in the distance between the listing and the point of interest (metro station) decreases the RevPAN by $0.006, and we see a beta value of -0.001 $ per month when considering the variable Distance*SuperHost. This suggests that being a SuperHost makes the distance between the nearest metro and the listing even more significant. Specifically, it implies that an increase of one meter in the distance between the listing and the point of interest decreases the RevPAN by \$0.006 + \$0.001 in the case of a SuperHost.

We notice that these variations are minimal, but these values are justified by the fact that we are analyzing a large area, and the movement is equal to one meter. All these values are supported by a p-value <1%. The only model that "differs" is M3, which pertains to degradation. In this case, we find a beta value of 0.000 \$ per month when considering the variable Distance*SuperHost. Therefore, we can deduce that being a SuperHost does not further influence the relationship between the distance from the degradation area to the structure and the RevPAN performance variable. Additionally, for the natural logarithm of RevPAN, we observe that the beta coefficients in models M2 to M6 are zero for the distance variables and distance * SuperHost.

Multivariate regression with LTR

the analysis, finally continued with the introduction of the dummy variable long term rent, described before. This variable, which can assume the values 0 or 1, has been multiplied with the distance variables computed before, in order to understand what happen to the listings in correlation with the distance and the LTR at the same time. The models studied refers to the dependent variables of the performance variables described before and their logarithms to study the LIN-LIN model and the LOG-LIN one, this last one resulting difficult to respond to questions since differences in percentage on a so little movement in meters are so small. In the tables attached at the bottom of this thesis there are presented all the values that will be described in this chapter, to have a full overview of the model and a better explanation of what has been found.

Regardless the revenues, as studied before for each meter going further from the closest metro stations, ZTL and public parking the revenues reduce. Introducing the interaction of these variables of distance with the long-term rent is possible to notice that the distance to the closest metro stations multiplied by the dummy variable of LTR is correlated with a positive value, showing that for the long-term rent listings, going further from the closest metro, public parking and ZTL increases the revenues per month. This can be explained by saying that having to booking an Airbnb for more than

1 month, guests are more willing to move away from the center where there is more concentration of ZTL, metro and parking, because a listing in the center is more think for visitors that want to enjoy the city for just a weekend or few days. For the degradation area, the long-term rent multiplier has no effects on the revenues per month of the listings. For the sports facilities the regression analysis stated that going further from them increases the revenues of the listings, the opposite happens, instead, introducing the multiplication with the LTR, indeed the listings accepting reservation only for more than 28 days, show a decrease of the revenues for each meter going further from the closest sport facility of 0.356$ per month.

Regarding the Average Daily Rate (ADR) the distance between a listing and the closest metro station follows the same trend as the proximity with the closest public parking. For each meter going further form the closest metro stations the revenues per month drop of 0.014$ per month, even more amplified for the listings offering a long-term rent, presenting a decrease in the revenue per month of ulterior 0.007$ per month for each meter going further from the closest metro station. For the listings proposing only long-term rent, is possible to say that the ADR drops of 0.014$ + 0.007$ for each meter going further from the closest metro station. For each meter going further from the closest degradation area or building, the ADR of the Airbnb increase of 0.011$ per day, even more amplified for the listing proposing only long-term rent, increasing their average daily rate by further 0.027$ per day. The opposite trend instead is shown by the proximity with the pedestrian areas, indeed, for each meter going further from the closest ZTL the ADR of the listings drops of 0.011$ per day, for the Airbnb offering only bookings of more than 28 days the ADR drops of only 0.004$ per day.

Regardless of the RevPAN, for each meter going further from the closest metro station and from the closest ZTL the RevPAN decrease, but this phenomenon is less effective for the listings proposing only long-term rent. For example, for each meter going further from the closest metro the RevPAN of the Airbnb drops 0.007$, considering only the listing proposing the long-term rent, the RevPAN drops of only 0.003$. Going away from a degradation area result in an increase in the RevPAN of the Airbnb, this phenomenon is even amplified for the LTR listings. The opposite trend is shown instead by the sport facilities. Step away from sport facilities results in better performances regarding the RevPAN, evidence not verified for the listings offering only LTR, indeed their RevPAN reduces by 0.009$ for each meter going further from the closest sport facility.

All the data described in this chapter can be considered significative, having a P-value below 1%. The analysis has been conducted also for the LOG-LIN model, but the changes of the performance variables in percentage do not show in a clear way what happens by moving only meter by meter. A better way to understand the LOG-LIN model could be done using the distinction done before of the ranges.

Conclusions

The analysis conducted in this thesis has been chosen because of the lack of research in this sector, in particular the dependence between the distance from the “points of interests” to the Airbnb have never been studied before. The research is based on the fact that, in our opinion, the distance from an Airbnb to a metro station, to a public parking, to a sport facility, to a ZTL and finally to an area in degradation condition can affect the choice of the customers, and as a consequence the performances of the Airbnb.

This analysis can be useful not only to understand the model of Airbnb and the main factors affecting the performance, but also for the hosts to develop better choices from a strategical and managerial point of view. Not only for the hosts that currently have listings on the platforms, but also from new potential hosts that want to enter in the sector or people looking to purchase a new apartment to be placed on the Airbnb platform.

At the beginning of the thesis five hypotheses have been drawn up. Our analysis has been made with the aim of verifying the validity of the predetermined hypothesis or to refute them. Afterwards are reported the five hypotheses with the results founded:

Hypothesis 1: The presence of a metro station near the property positively influences the performance of such Airbnb.

First, with the univariate regression and afterwards with the multivariate regression has been found that for each meter going further from the closest metro station the performances of the listings decrease. Both in case of the regressions, where the continuous variables of the distance in meters have been used, and in the case of the division in ranges, the performances of the Airbnb dropped by stepping away from the closest metro stations. The average of the revenues per month in the range 1 are 1292.38$ per month and in range 7 are 530.85$ per month. Also, the β in the regressions shows a negative value for every dependent variable studied. With the data collected and the analysis done has been possible to confirm our hypothesis 1.

Hypothesis 2: The presence of a degraded area near the property negatively influences the performance of such Airbnb.

The dataset regarding the degradation conditions refers to all the buildings and areas declared as is by the municipality of Milan. The regression shows a positive sign in the beta for all the performance variables studied, indeed is possible to say that for each meter going further from the closest area or building in degradation condition, the performances of the Airbnb increase.

$$Revenues \; = 1366{,}144 \, + \, 0{,}2637 \; DistanceDegradation$$

This regression states that for each meter going further from the closest area in degradation conditions, the monthly revenues of the listings drop of 0.2637$ per month. A particular trend is evidenced in the division in ranges, indeed is possible to notice that for the first 5 ranges, stepping away from the closest degradation point results in better performances. In the range 6 and 7 drop drastically, showing that listings further than 1500m from the closest degradation point have worst performance even with respect to the listings in range 1.

Hypothesis 3: The presence of parking near the property positively influences the performance of such Airbnb.

The analysis conducted in this thesis showed evidence of the confirmation of this hypothesis. Indeed, in the regression the β associated to the distance from the listings to the closest public parking show a negative sign, indicating that for each meter going further from the closest public parking the performances of the Airbnb drop. In the univariate regression with the revenues as dependent variable, the formula found has been *Revenues = 1761,951 – 0,3343*DistanceParking* showing that for each meter going further from the closest public parking the revenues of the Airbnb dropped of 0.3343$ per month. The same trend is evidenced also in the Average daily rate which decreases by 0.0233$ per night for each meter stepping away from the closest public parking. Also, in the subdivision of the distance by ranges was possible to found that the average of the RevPAN per month in range 1 is almost the double of the one in range 7.

Hypothesis 4: The presence of ZTL/Pedestrian Areas near the property positively influences the performance of such Airbnb.

Analyzing the distances of various listings from restricted areas such as pedestrian zones and ZTLs, both through an analysis for individual structures and for different ranges, yielded significant results. During the analyses conducted by dividing the listings into 7 ranges, we noticed that transitioning from Range 1, which contains structures closest to the metro, to Range 7, which contains the farthest structures, the performances decreased steeply. For example, in the case of Revenues, we observed a decrease from an average monthly value of \$1498.67 to \$587.61. Examining the city structure of Milan, we identified a high concentration of these areas in the historic center, suggesting that Range 1 presumably consists of structures located in the center of Milan, thus justifying such a high average value of Revenues. During the univariate and multivariate regressions, we further confirmed our hypotheses. Indeed, analyzing the beta coefficients in the various regressions consistently revealed negative values, indicating a decline in performance associated with increasing distance of the structure from these zones. Additionally, we observed that in all regressions, the R^2 value associated with the model containing the distance from ZTLs was consistently the highest, reaching a peak of 6.11% for the ADR performance variable, thus demonstrating a high explanation of such a model.

Hypothesis 5: The presence of sports facilities near the property positively influences the performance of such Airbnb.

Moving on to the analysis regarding the performance variables in relation to the distance from Airbnb to sports facilities, we obtained interesting results. Starting with the analysis of these performances through the division into ranges, we noticed that the trend of these variables followed an alternating and non-constant pattern. Considering the Revenues and recalling that range 1 contains structures with a sports facility very close to the property while range 7 contains structures with a greater distance from the sports facilities, we observe a linear growth in the value of the variable from range 1 to range 4 (from $948.72 to $1034.70) and subsequently an exponential growth up to range 6 reaching a peak of $2252.85, then dropping drastically to $704.70 for range 7. Through analyses of Milan's layout, we managed to explain this phenomenon by understanding that the Airbnbs in range 7 were located in the peripheral areas of Milan where the prices per night are lower. From this analysis, we began to intuit that moving away from these points of interest would lead to better performance for the property. This intuition was subsequently confirmed first through univariate regressions and later in multivariate regressions where we consistently found positive regression beta coefficients associated with p-values < 1%. Therefore, we can see how this study has reversed the initial hypothesis.

Further analysis may be conducted analyzing in a separate way the years taken in consideration in order to understand how and if the trends changed during the pandemic crisis. Moreover, may be monitored the website of the municipality of Milan to look forward for releases of new datasets or the updating of the ones used in this analysis, in order to study how the new distance variables may affect the performances of Airbnb.

Attachments

Multivariate regressions

Revenues	M1	M2	M3	M4	M5	M6	M7
Metro		-0.184***					-0.110***
		(0.008)					(0.008)
Degrado			0.145***				0.138***
			(0.008)				(0.008)
ZTL				-0.154***			-0.114***
				(0.007)			(0.007)
Sport					0.053***		0.068***
					(0.009)		(0.009)
Parcheggi						-0.125***	-0.094***
						(0.006)	(0.006)
Max Guest	186.651***	186.506***	187.063***	186.678***	186.578***	187.133***	187.243***
	(1.692)	(1.690)	(1.692)	(1.691)	(1.692)	(1.691)	(1.690)
Entire APT	248.729***	249.675***	247.861***	249.068***	248.442***	246.081***	246.370***
	(4.377)	(4.372)	(4.375)	(4.376)	(4.377)	(4.377)	(4.373)
Instantbook	395.250***	392.271***	393.158***	392.839***	395.104***	394.676***	389.069***
	(3.660)	(3.658)	(3.658)	(3.661)	(3.660)	(3.658)	(3.657)
Superhost	152.051***	151.326***	150.618***	152.448***	152.385***	151.806***	150.790***
	(4.277)	(4.273)	(4.274)	(4.274)	(4.276)	(4.275)	(4.269)
LTR	-81.141***	-81.481***	-81.193***	-82.176***	-81.014***	-80.516***	-81.531***
	(14.991)	(14.980)	(14.979)	(14.990)	(14.990)	(14.992)	(14.972)
ID NIL	yes	yes	yes	yes	yes	yes	yes
Year	yes	yes	yes	yes	yes	yes	yes
Mese	yes	yes	yes	yes	yes	yes	yes
N	4.11e+05	4.11e+05	4.11e+05	4.11e+05	4.11e+05	4.11e+05	4.11e+05
R2	0.282	0.283	0.283	0.283	0.282	0.283	0.284

Ln(Revenues)	M1	M2	M3	M4	M5	M6	M7
Metro		0.000***					0.000***
		(0.000)					(0.000)
Degrado			0.000***				0.000***
			(0.000)				(0.000)
ZTL				0.000***			0.000***
				(0.000)			(0.000)
Sport					0.000***		0.000***
					(0.000)		(0.000)
Parcheggi						0.000***	0.000***
						(0.000)	(0.000)
Max Guest	0.111***	0.111***	0.111***	0.111***	0.111***	0.111***	0.111***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Entire APT	0.372***	0.372***	0.371***	0.372***	0.372***	0.370***	0.370***
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)
Instantbook	0.366***	0.364***	0.365***	0.365***	0.366***	0.366***	0.362***
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
Superhost	0.198***	0.198***	0.198***	0.199***	0.199***	0.198***	0.198***
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)
LTR	-0.268***	-0.269***	-0.268***	-0.269***	-0.268***	-0.268***	-0.269***
	(0.014)	(0.014)	(0.014)	(0.014)	(0.014)	(0.014)	(0.014)
ID NIL	yes	yes	yes	yes	yes	yes	yes
Year	yes	yes	yes	yes	yes	yes	yes
Mese	yes	yes	yes	yes	yes	yes	yes
N	4.11e+05	4.11e+05	4.11e+05	4.11e+05	4.11e+05	4.11e+05	4.11e+05
R2	0.252	0.253	0.253	0.253	0.252	0.253	0.254

Y = RevPAN	M1	M2	M3	M4	M5	M6	M7
Metro		-0.007***					-0.004***
		(0.000)					(0.000)
Degrado			0.005***				0.005***
			(0.000)				(0.000)
ZTL				-0.005***			-0.004***
				(0.000)			(0.000)
Sport					0.002***		0.003***
					(0.000)		(0.000)
Parcheggi						-0.004***	-0.003***
						(0.000)	(0.000)
Max Guest	8.179***	8.173***	8.192***	8.180***	8.176***	8.196***	8.199***
	(0.065)	(0.065)	(0.065)	(0.065)	(0.065)	(0.065)	(0.065)
Entire APT	13.656***	13.692***	13.627***	13.669***	13.645***	13.561***	13.574***
	(0.167)	(0.167)	(0.167)	(0.167)	(0.167)	(0.167)	(0.167)
Instantbook	12.835***	12.725***	12.762***	12.749***	12.829***	12.814***	12.614***
	(0.138)	(0.138)	(0.138)	(0.138)	(0.138)	(0.138)	(0.138)
Superhost	7.157***	7.131***	7.107***	7.171***	7.170***	7.148***	7.113***
	(0.156)	(0.156)	(0.156)	(0.156)	(0.156)	(0.156)	(0.156)
LTR	3.722***	3.711***	3.717***	3.686***	3.727***	3.746***	3.710***
	(0.613)	(0.612)	(0.612)	(0.613)	(0.613)	(0.613)	(0.612)
ID NIL	yes	yes	yes	yes	yes	yes	yes
Year	yes	yes	yes	yes	yes	yes	yes
Mese	yes	yes	yes	yes	yes	yes	yes
N	4.11e+05	4.11e+05	4.11e+05	4.11e+05	4.11e+05	4.11e+05	4.11e+05
R2	0.315	0.315	0.315	0.315	0.315	0.315	0.317

Y = lnRevPAN	M1	M2	M3	M4	M5	M6	M7
Metro		-0.000***					-0.000***
		(0.000)					(0.000)
Degrado			0.000***				0.000***
			(0.000)				(0.000)
ZTL				-0.000***			-0.000***
				(0.000)			(0.000)
Sport					0.000**		0.000***
					(0.000)		(0.000)
Parcheggi						-0.000***	-0.000***
						(0.000)	(0.000)
Max Guest	0.109***	0.109***	0.109***	0.109***	0.109***	0.109***	0.109***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Entire APT	0.439***	0.440***	0.439***	0.440***	0.439***	0.438***	0.438***
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)
Instantbook	0.305***	0.303***	0.304***	0.304***	0.305***	0.305***	0.302***
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
Superhost	0.237***	0.236***	0.236***	0.237***	0.237***	0.236***	0.236***
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
LTR	-0.088***	-0.088***	-0.088***	-0.088***	-0.088***	-0.087***	-0.088***
	(0.012)	(0.012)	(0.012)	(0.012)	(0.012)	(0.012)	(0.012)
ID NIL	yes	yes	yes	yes	yes	yes	yes
Year	yes	yes	yes	yes	yes	yes	yes
Mese	yes	yes	yes	yes	yes	yes	yes
N	4.11e+05	4.11e+05	4.11e+05	4.11e+05	4.11e+05	4.11e+05	4.11e+05
R2	0.299	0.299	0.299	0.299	0.299	0.299	0.300

Y = ADR	M1	M2	M3	M4	M5	M6	M7
Metro		-0.014***					-0.008***
		(0.000)					(0.000)
Degrado			0.011***				0.011***
			(0.000)				(0.000)
ZTL				-0.011***			-0.008***
				(0.000)			(0.000)
Sport					0.005***		0.006***
					(0.000)		(0.000)
Parcheggi						-0.011***	-0.009***
						(0.000)	(0.000)
Max Guest	17.653***	17.640***	17.689***	17.653***	17.645***	17.699***	17.707***
	(0.104)	(0.104)	(0.104)	(0.104)	(0.104)	(0.104)	(0.104)
Entire APT	18.311***	18.385***	18.237***	18.337***	18.282***	18.073***	18.079***
	(0.273)	(0.272)	(0.272)	(0.273)	(0.273)	(0.272)	(0.272)
Instantbook	-9.926***	-10.161***	-10.099***	-10.104***	-9.941***	-9.985***	-10.421***
	(0.223)	(0.223)	(0.223)	(0.223)	(0.223)	(0.223)	(0.223)
Superhost	-13.792***	-13.845***	-13.911***	-13.764***	-13.758***	-13.812***	-13.892***
	(0.226)	(0.226)	(0.226)	(0.226)	(0.226)	(0.226)	(0.226)
LTR	12.492***	12.467***	12.491***	12.412***	12.506***	12.533***	12.472***
	(0.930)	(0.928)	(0.927)	(0.929)	(0.930)	(0.927)	(0.925)
ID NIL	yes	yes	yes	yes	yes	yes	yes
Year	yes	yes	yes	yes	yes	yes	yes
Mese	yes	yes	yes	yes	yes	yes	yes
N	4.09e+05	4.09e+05	4.09e+05	4.09e+05	4.09e+05	4.09e+05	4.09e+05
R2	0.304	0.306	0.305	0.305	0.304	0.306	0.308

Y = lnADR	M1	M2	M3	M4	M5	M6	M7
Metro		-0.000***					-0.000***
		(0.000)					(0.000)
Degrado			0.000***				0.000***
			(0.000)				(0.000)
ZTL				-0.000***			-0.000***
				(0.000)			(0.000)
Sport					0.000***		0.000***
					(0.000)		(0.000)
Parcheggi						-0.000***	-0.000***
						(0.000)	(0.000)
Max Guest	0.118***	0.118***	0.118***	0.118***	0.118***	0.119***	0.119***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Entire APT	0.295***	0.296***	0.295***	0.296***	0.295***	0.293***	0.294***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Instantbook	-0.056***	-0.058***	-0.057***	-0.057***	-0.056***	-0.056***	-0.060***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Superhost	-0.093***	-0.093***	-0.094***	-0.093***	-0.093***	-0.093***	-0.094***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
LTR	0.018***	0.018***	0.018***	0.017**	0.018***	0.018***	0.018***
	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)
ID NIL	yes	yes	yes	yes	yes	yes	yes
Year	yes	yes	yes	yes	yes	yes	yes
Mese	yes	yes	yes	yes	yes	yes	yes
N	4.09e+05	4.09e+05	4.09e+05	4.09e+05	4.09e+05	4.09e+05	4.09e+05
R2	0.398	0.400	0.399	0.399	0.398	0.400	0.403

OCC	**M1**	**M2**	**M3**	**M4**	**M5**	**M6**	**M7**
Metro		0.000***					0.000***
		(0.000)					(0.000)
Degrado			0.000***				0.000***
			(0.000)				(0.000)
ZTL				0.000***			0.000***
				(0.000)			(0.000)
Sport					0.000***		0.000***
					(0.000)		(0.000)
Parcheggi						0.000***	0.000***
						(0.000)	(0.000)
Max Guest	-0.002***	-0.002***	-0.002***	-0.002***	-0.002***	-0.002***	-0.002***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Entire APT	0.053***	0.053***	0.053***	0.053***	0.053***	0.053***	0.053***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Instantbook	0.121***	0.121***	0.121***	0.121***	0.121***	0.121***	0.121***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Superhost	0.115***	0.115***	0.115***	0.115***	0.115***	0.115***	0.115***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
LTR	-0.019***	-0.019***	-0.019***	-0.019***	-0.019***	-0.019***	-0.019***
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)
ID NIL	yes	yes	yes	yes	yes	yes	yes
Year	yes	yes	yes	yes	yes	yes	yes
Mese	yes	yes	yes	yes	yes	yes	yes
N	4.14e+05	4.14e+05	4.14e+05	4.14e+05	4.14e+05	4.14e+05	4.14e+05
R2	0.126	0.126	0.126	0.126	0.126	0.126	0.126

Multivariate regressions with SuperHost

Revenues	M1	M2	M3	M4	M5	M6
Metro		-0.178***				
		(0.008)				
Metro * Superhost		-0.029***				
		(0.009)				
Degrado			0.150***			
			(0.008)			
Degrado * Superhost			-0.024*			
			(0.013)			
ZTL				-0.142***		
				(0.007)		
ZTL * Superhost				-0.058***		
				(0.010)		
Sport					0.047***	
					(0.009)	
Sport * Superhost					0.036**	
					(0.017)	
Parcheggi						-0.123***
						(0.006)
Parcheggi * Superhost						-0.005
						(0.005)
Max Guest	186.651***	186.541***	187.026***	186.722***	186.582***	187.124***
	(1.692)	(1.690)	(1.692)	(1.691)	(1.692)	(1.691)
Entire APT	248.729***	249.578***	248.025***	248.814***	248.263***	246.098***
	(4.377)	(4.372)	(4.375)	(4.376)	(4.375)	(4.377)
Instantbook	395.250***	392.209***	393.073***	392.936***	395.209***	394.720***
	(3.660)	(3.658)	(3.659)	(3.660)	(3.659)	(3.659)
Superhost	152.051***	166.010***	163.787***	179.771***	130.683***	155.966***
	(4.277)	(6.942)	(8.206)	(7.049)	(10.142)	(6.646)
LTR	-81.141***	-81.721***	-81.243***	-82.358***	-81.275***	-80.524***
	(14.991)	(14.980)	(14.978)	(14.987)	(14.989)	(14.992)
ID NIL	yes	yes	yes	yes	yes	yes
Year	yes	yes	yes	yes	yes	yes
Mese	yes	yes	yes	yes	yes	yes
N	4.11e+05	4.11e+05	4.11e+05	4.11e+05	4.11e+05	4.11e+05
R2	0.282	0.283	0.283	0.283	0.282	0.283

Ln(Revenues)	M1	M2	M3	M4	M5	M6
Metro		0.000***				
		(0.000)				
Metro * Superhost		0.000***				
		(0.000)				
Degrado			0.000***			
			(0.000)			
Degrado * Superhost			0.000***			
			(0.000)			
ZTL				0.000***		
				(0.000)		
ZTL * Superhost				0.000***		
				(0.000)		
Sport					0.000***	
					(0.000)	
Sport * Superhost					0.000***	
					(0.000)	
Parcheggi						0.000***
						(0.000)
Parcheggi * Superhost						0.000***
						(0.000)
Max Guest	0.111***	0.111***	0.111***	0.111***	0.111***	0.111***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Entire APT	0.372***	0.373***	0.372***	0.372***	0.372***	0.370***
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)
Instantbook	0.366***	0.364***	0.365***	0.365***	0.366***	0.366***
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
Superhost	0.198***	0.188***	0.229***	0.189***	0.220***	0.184***
	(0.004)	(0.006)	(0.007)	(0.006)	(0.008)	(0.006)
LTR	-0.268***	-0.269***	-0.269***	-0.269***	-0.268***	-0.268***
	(0.014)	(0.014)	(0.014)	(0.014)	(0.014)	(0.014)
ID NIL	yes	yes	yes	yes	yes	yes
Year	yes	yes	yes	yes	yes	yes
Mese	yes	yes	yes	yes	yes	yes
N	4.11e+05	4.11e+05	4.11e+05	4.11e+05	4.11e+05	4.11e+05
R2	0.252	0.253	0.253	0.253	0.252	0.253

RevPAN	M1	M2	M3	M4	M5	M6
Metro		-0.006***				
		(0.000)				
Metro * Superhost		-0.001***				
		(0.001)				
Degrado			0.005***			
			(0.000)			
Degrado * Superhost			-0.000			
			(0.000)			
ZTL				-0.005***		
				(0.000)		
ZTL * Superhost				-0.003***		
				(0.000)		
Sport					0.002***	
					(0.000)	
Sport * Superhost					0.002***	
					(0.001)	
Parcheggi						-0.004***
						(0.000)
Parcheggi * Superhost						-0.001***
						(0.000)
Max Guest	8.179***	8.175***	8.192***	8.182***	8.176***	8.195***
	(0.065)	(0.065)	(0.065)	(0.065)	(0.065)	(0.065)
Entire APT	13.656***	13.686***	13.629***	13.657***	13.635***	13.563***
	(0.167)	(0.167)	(0.167)	(0.167)	(0.167)	(0.167)
Instantbook	12.835***	12.722***	12.761***	12.754***	12.835***	12.818***
	(0.138)	(0.138)	(0.138)	(0.138)	(0.138)	(0.138)
Superhost	7.157***	7.889***	7.249***	8.498***	5.938***	7.565***
	(0.156)	(0.255)	(0.300)	(0.256)	(0.369)	(0.243)
LTR	3.722***	3.698***	3.717***	3.676***	3.712***	3.745***
	(0.613)	(0.612)	(0.612)	(0.612)	(0.612)	(0.612)
ID NIL	yes	yes	yes	yes	yes	yes
Year	yes	yes	yes	yes	yes	yes
Mese	yes	yes	yes	yes	yes	yes
N	4.11e+05	4.11e+05	4.11e+05	4.11e+05	4.11e+05	4.11e+05
R2	0.315	0.315	0.315	0.315	0.315	0.315

Ln(RevPAN)	M1	M2	M3	M4	M5	M6
Metro		0.000***				
		(0.000)				
Metro * Superhost		0.000***				
		(0.000)				
Degrado			0.000***			
			(0.000)			
Degrado * Superhost			0.000***			
			(0.000)			
ZTL				0.000***		
				(0.000)		
ZTL * Superhost				0.000***		
				(0.000)		
Sport					0.000***	
					(0.000)	
Sport * Superhost					0.000***	
					(0.000)	
Parcheggi						0.000***
						(0.000)
Parcheggi * Superhost						0.000***
						(0.000)
Max Guest	0.109***	0.109***	0.109***	0.109***	0.109***	0.109***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Entire APT	0.439***	0.440***	0.439***	0.440***	0.439***	0.438***
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)
Instantbook	0.305***	0.303***	0.304***	0.304***	0.305***	0.305***
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
Superhost	0.237***	0.231***	0.259***	0.231***	0.255***	0.230***
	(0.003)	(0.005)	(0.006)	(0.005)	(0.007)	(0.005)
LTR	-0.088***	-0.088***	-0.088***	-0.088***	-0.088***	-0.087***
	(0.012)	(0.012)	(0.012)	(0.012)	(0.012)	(0.012)
ID NIL	yes	yes	yes	yes	yes	yes
Year	yes	yes	yes	yes	yes	yes
Mese	yes	yes	yes	yes	yes	yes
N	4.11e+05	4.11e+05	4.11e+05	4.11e+05	4.11e+05	4.11e+05
R2	0.299	0.299	0.299	0.299	0.299	0.299

ADR	M1	M2	M3	M4	M5	M6
Metro		-0.014***				
		(0.000)				
Metro * Superhost		-0.000				
		(0.001)				
Degrado			0.012***			
			(0.001)			
Degrado * Superhost			-0.003***			
			(0.001)			
ZTL				-0.011***		
				(0.000)		
ZTL * Superhost				-0.001**		
				(0.001)		
Sport					0.005***	
					(0.001)	
Sport * Superhost					0.004***	
					(0.001)	
Parcheggi						-0.012***
						(0.000)
Parcheggi * Superhost						0.003***
						(0.000)
Max Guest	17.653***	17.641***	17.684***	17.654***	17.646***	17.703***
	(0.104)	(0.104)	(0.104)	(0.104)	(0.104)	(0.104)
Entire APT	18.311***	18.383***	18.257***	18.332***	18.264***	18.064***
	(0.273)	(0.272)	(0.272)	(0.273)	(0.272)	(0.272)
Instantbook	-9.926***	-10.162***	-10.109***	-10.102***	-9.929***	-10.012***
	(0.223)	(0.223)	(0.223)	(0.223)	(0.223)	(0.223)
Superhost	-13.792***	-13.608***	-12.311***	-13.197***	-15.884***	-16.155***
	(0.226)	(0.372)	(0.444)	(0.365)	(0.516)	(0.351)
LTR	12.492***	12.463***	12.483***	12.408***	12.481***	12.536***
	(0.930)	(0.928)	(0.927)	(0.929)	(0.930)	(0.927)
ID NIL	yes	yes	yes	yes	yes	yes
Year	yes	yes	yes	yes	yes	yes
Mese	yes	yes	yes	yes	yes	yes
N	4.09e+05	4.09e+05	4.09e+05	4.09e+05	4.09e+05	4.09e+05
R2	0.304	0.306	0.305	0.305	0.304	0.306

Ln(ADR)	M1	M2	M3	M4	M5	M6
Metro		0.000***				
		(0.000)				
Metro * Superhost		0.000***				
		(0.000)				
Degrado			0.000***			
			(0.000)			
Degrado * Superhost			0.000***			
			(0.000)			
ZTL				0.000***		
				(0.000)		
ZTL * Superhost				0.000***		
				(0.000)		
Sport					0.000***	
					(0.000)	
Sport * Superhost					0.000***	
					(0.000)	
Parcheggi						0.000***
						(0.000)
Parcheggi * Superhost						0.000***
						(0.000)
Max Guest	0.118***	0.118***	0.118***	0.118***	0.118***	0.118***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Entire APT	0.295***	0.296***	0.295***	0.295***	0.295***	0.293***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Instantbook	-0.056***	-0.058***	-0.057***	-0.057***	-0.056***	-0.056***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Superhost	-0.093***	-0.081***	-0.087***	-0.073***	-0.116***	-0.097***
	(0.002)	(0.003)	(0.003)	(0.002)	(0.003)	(0.002)
LTR	0.018***	0.017***	0.018***	0.017***	0.018***	0.018***
	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)
ID NIL	yes	yes	yes	yes	yes	yes
Year	yes	yes	yes	yes	yes	yes
Mese	yes	yes	yes	yes	yes	yes
N	4.09e+05	4.09e+05	4.09e+05	4.09e+05	4.09e+05	4.09e+05
R2	0.398	0.400	0.399	0.399	0.398	0.400

OCC	M1	M2	M3	M4	M5	M6
Metro		0.000***				
		(0.000)				
Metro * Superhost		0.000***				
		(0.000)				
Degrado			0.000***			
			(0.000)			
Degrado * Superhost			0.000***			
			(0.000)			
ZTL				0.000***		
				(0.000)		
ZTL * Superhost				0.000***		
				(0.000)		
Sport					0.000***	
					(0.000)	
Sport * Superhost					0.000***	
					(0.000)	
Parcheggi						0.000***
						(0.000)
Parcheggi * Superhost						0.000***
						(0.000)
Max Guest	-0.002***	-0.002***	-0.002***	-0.002***	-0.002***	-0.002***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Entire APT	0.053***	0.053***	0.053***	0.053***	0.053***	0.053***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Instantbook	0.121***	0.121***	0.121***	0.121***	0.121***	0.121***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Superhost	0.115***	0.108***	0.122***	0.107***	0.130***	0.111***
	(0.001)	(0.002)	(0.002)	(0.002)	(0.003)	(0.002)
LTR	-0.019***	-0.018***	-0.019***	-0.018***	-0.018***	-0.019***
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)
ID NIL	yes	yes	yes	yes	yes	yes
Year	yes	yes	yes	yes	yes	yes
Mese	yes	yes	yes	yes	yes	yes
N	4.14e+05	4.14e+05	4.14e+05	4.14e+05	4.14e+05	4.14e+05
R2	0.126	0.126	0.126	0.126	0.126	0.126

Multivariate regressions with long term rent

Y = Revenues	M1	M2	M3	M4	M5	M6
Metro		-0.188***				
		(0.008)				
Metro * LTR		0.191***				
		(0.030)				
Degrado			0.145***			
			(0.008)			
Degrado * LTR			0.000			
			(0.048)			
ZTL				-0.160***		
				(0.007)		
ZTL * LTR				0.336***		
				(0.037)		
Sport					0.059***	
					(0.009)	
Sport * LTR					-0.356***	
					(0.061)	
Parcheggi						-0.126***
						(0.006)
Parcheggi * LTR						0.066***
						(0.017)
Max Guest	186.651***	186.545***	187.063***	186.798***	186.609***	187.151***
	(1.692)	(1.690)	(1.692)	(1.691)	(1.692)	(1.691)
Entire APT	248.729***	249.913***	247.861***	248.994***	248.506***	246.300***
	(4.377)	(4.373)	(4.376)	(4.376)	(4.377)	(4.378)
Instantbook	395.250***	392.044***	393.158***	392.030***	394.956***	394.425***
	(3.660)	(3.658)	(3.658)	(3.661)	(3.660)	(3.659)
Superhost	152.051***	151.369***	150.618***	152.391***	152.627***	151.783***
	(4.277)	(4.273)	(4.274)	(4.274)	(4.277)	(4.275)
LTR	-81.141***	-189.288***	-81.267***	-252.082***	123.017***	-133.635***
	(14.991)	(24.533)	(28.356)	(24.900)	(37.418)	(22.970)
ID NIL	yes	yes	yes	yes	yes	yes
Year	yes	yes	yes	yes	yes	yes
Mese	yes	yes	yes	yes	yes	yes
N	4.11e+05	4.11e+05	4.11e+05	4.11e+05	4.11e+05	4.11e+05
R2	0.282	0.283	0.283	0.283	0.282	0.283

Y = lnRevenues	M1	M2	M3	M4	M5	M6
Metro		-0.000***				
		(0.000)				
Metro * LTR		0.000***				
		(0.000)				
Degrado			0.000***			
			(0.000)			
Degrado * LTR			0.000*			
			(0.000)			
ZTL				-0.000***		
				(0.000)		
ZTL * LTR				0.000***		
				(0.000)		
Sport					0.000***	
					(0.000)	
Sport * LTR					-0.000*	
					(0.000)	
Parcheggi						-0.000***
						(0.000)
Parcheggi * LTR						-0.000
						(0.000)
Max Guest	0.111***	0.111***	0.111***	0.111***	0.111***	0.111***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Entire APT	0.372***	0.373***	0.371***	0.372***	0.372***	0.370***
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)
Instantbook	0.366***	0.364***	0.365***	0.364***	0.366***	0.366***
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
Superhost	0.198***	0.198***	0.198***	0.199***	0.199***	0.198***
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)
LTR	-0.268***	-0.323***	-0.314***	-0.319***	-0.217***	-0.255***
	(0.014)	(0.022)	(0.028)	(0.022)	(0.033)	(0.021)
ID NIL	yes	yes	yes	yes	yes	yes
Year	yes	yes	yes	yes	yes	yes
Mese	yes	yes	yes	yes	yes	yes
N	4.11e+05	4.11e+05	4.11e+05	4.11e+05	4.11e+05	4.11e+05
R2	0.252	0.253	0.253	0.253	0.252	0.253

Y = RevPAN	M1	M2	M3	M4	M5	M6
Metro		-0.007***				
		(0.000)				
Metro * LTR		0.004***				
		(0.001)				
Degrado			0.005***			
			(0.000)			
Degrado * LTR			0.004*			
			(0.002)			
ZTL				-0.006***		
				(0.000)		
ZTL * LTR				0.009***		
				(0.001)		
Sport					0.002***	
					(0.000)	
Sport * LTR					-0.011***	
					(0.003)	
Parcheggi						-0.005***
						(0.000)
Parcheggi * LTR						0.000
						(0.001)
Max Guest	8.179***	8.174***	8.191***	8.183***	8.177***	8.196***
	(0.065)	(0.065)	(0.065)	(0.065)	(0.065)	(0.065)
Entire APT	13.656***	13.696***	13.632***	13.667***	13.648***	13.562***
	(0.167)	(0.167)	(0.167)	(0.167)	(0.167)	(0.167)
Instantbook	12.835***	12.720***	12.763***	12.728***	12.824***	12.812***
	(0.138)	(0.138)	(0.138)	(0.139)	(0.138)	(0.138)
Superhost	7.157***	7.132***	7.108***	7.169***	7.177***	7.148***
	(0.156)	(0.156)	(0.156)	(0.156)	(0.156)	(0.156)
LTR	3.722***	1.552	1.677	-0.773	9.893***	3.438***
	(0.613)	(1.008)	(1.171)	(1.016)	(1.540)	(0.931)
ID NIL	yes	yes	yes	yes	yes	yes
Year	yes	yes	yes	yes	yes	yes
Mese	yes	yes	yes	yes	yes	yes
N	4.11e+05	4.11e+05	4.11e+05	4.11e+05	4.11e+05	4.11e+05
R2	0.315	0.315	0.315	0.315	0.315	0.315

Y = lnRevPAN	M1	M2	M3	M4	M5	M6
Metro		-0.000***				
		(0.000)				
Metro * LTR		0.000**				
		(0.000)				
Degrado			0.000***			
			(0.000)			
Degrado * LTR			0.000*			
			(0.000)			
ZTL				-0.000***		
				(0.000)		
ZTL * LTR				0.000**		
				(0.000)		
Sport					0.000***	
					(0.000)	
Sport * LTR					-0.000*	
					(0.000)	
Parcheggi						-0.000***
						(0.000)
Parcheggi * LTR						-0.000**
						(0.000)
Max Guest	0.109***	0.109***	0.109***	0.109***	0.109***	0.109***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Entire APT	0.439***	0.440***	0.439***	0.440***	0.439***	0.438***
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)
Instantbook	0.305***	0.303***	0.304***	0.304***	0.305***	0.305***
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
Superhost	0.237***	0.236***	0.236***	0.237***	0.237***	0.236***
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
LTR	-0.088***	-0.121***	-0.125***	-0.120***	-0.044	-0.058***
	(0.012)	(0.019)	(0.024)	(0.019)	(0.029)	(0.019)
ID NIL	yes	yes	yes	yes	yes	yes
Year	yes	yes	yes	yes	yes	yes
Mese	yes	yes	yes	yes	yes	yes
N	4.11e+05	4.11e+05	4.11e+05	4.11e+05	4.11e+05	4.11e+05
R2	0.299	0.299	0.299	0.299	0.299	0.299

Y = ADR	M1	M2	M3	M4	M5	M6
Metro		-0.014***				
		(0.000)				
Metro * LTR		-0.007***				
		(0.002)				
Degrado			0.011***			
			(0.000)			
Degrado * LTR			0.027***			
			(0.003)			
ZTL				-0.011***		
				(0.000)		
ZTL * LTR				0.007***		
				(0.002)		
Sport					0.006***	
					(0.001)	
Sport * LTR					-0.014***	
					(0.004)	
Parcheggi						-0.011***
						(0.000)
Parcheggi * LTR						-0.005***
						(0.001)
Max Guest	17.653***	17.639***	17.678***	17.655***	17.646***	17.697***
	(0.104)	(0.104)	(0.104)	(0.104)	(0.104)	(0.104)
Entire APT	18.311***	18.375***	18.271***	18.336***	18.284***	18.056***
	(0.273)	(0.272)	(0.272)	(0.273)	(0.273)	(0.272)
Instantbook	-9.926***	-10.153***	-10.093***	-10.122***	-9.947***	-9.964***
	(0.223)	(0.223)	(0.223)	(0.223)	(0.223)	(0.223)
Superhost	-13.792***	-13.847***	-13.904***	-13.766***	-13.749***	-13.809***
	(0.226)	(0.226)	(0.226)	(0.226)	(0.226)	(0.226)
LTR	12.492***	16.557***	-2.134	8.871***	20.822***	16.905***
	(0.930)	(1.518)	(1.885)	(1.520)	(2.258)	(1.420)
ID NIL	yes	yes	yes	yes	yes	yes
Year	yes	yes	yes	yes	yes	yes
Mese	yes	yes	yes	yes	yes	yes
N	4.09e+05	4.09e+05	4.09e+05	4.09e+05	4.09e+05	4.09e+05
R2	0.304	0.306	0.305	0.305	0.304	0.306

Y = lnADR	M1	M2	M3	M4	M5	M6
Metro		-0.000***				
		(0.000)				
Metro * LTR		-0.000**				
		(0.000)				
Degrado			0.000***			
			(0.000)			
Degrado * LTR			0.000***			
			(0.000)			
ZTL				-0.000***		
				(0.000)		
ZTL * LTR				0.000***		
				(0.000)		
Sport					0.000***	
					(0.000)	
Sport * LTR					-0.000***	
					(0.000)	
Parcheggi						-0.000***
						(0.000)
Parcheggi * LTR						-0.000***
						(0.000)
Max Guest	0.118***	0.118***	0.118***	0.118***	0.118***	0.119***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Entire APT	0.295***	0.296***	0.295***	0.296***	0.295***	0.293***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Instantbook	-0.056***	-0.058***	-0.057***	-0.057***	-0.056***	-0.056***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Superhost	-0.093***	-0.093***	-0.094***	-0.093***	-0.093***	-0.093***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
LTR	0.018***	0.035***	-0.078***	-0.020*	0.090***	0.061***
	(0.007)	(0.011)	(0.014)	(0.010)	(0.016)	(0.010)
ID NIL	yes	yes	yes	yes	yes	yes
Year	yes	yes	yes	yes	yes	yes
Mese	yes	yes	yes	yes	yes	yes
N	4.09e+05	4.09e+05	4.09e+05	4.09e+05	4.09e+05	4.09e+05
R2	0.398	0.400	0.399	0.399	0.398	0.400

Y = OCC	M1	M2	M3	M4	M5	M6
Metro		-0.000				
		(0.000)				
Metro * LTR		0.000***				
		(0.000)				
Degrado			-0.000***			
			(0.000)			
Degrado * LTR			-0.000*			
			(0.000)			
ZTL				-0.000		
				(0.000)		
ZTL * LTR				-0.000		
				(0.000)		
Sport					-0.000	
					(0.000)	
Sport * LTR					0.000	
					(0.000)	
Parcheggi						0.000***
						(0.000)
Parcheggi * LTR						-0.000
						(0.000)
Max Guest	-0.002***	-0.002***	-0.002***	-0.002***	-0.002***	-0.002***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Entire APT	0.053***	0.053***	0.053***	0.053***	0.053***	0.053***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Instantbook	0.121***	0.121***	0.121***	0.121***	0.121***	0.121***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Superhost	0.115***	0.115***	0.115***	0.115***	0.115***	0.115***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
LTR	-0.019***	-0.033***	-0.005	-0.013**	-0.024***	-0.017***
	(0.004)	(0.006)	(0.008)	(0.006)	(0.009)	(0.006)
ID NIL	yes	yes	yes	yes	yes	yes
Year	yes	yes	yes	yes	yes	yes
Mese	yes	yes	yes	yes	yes	yes
N	4.14e+05	4.14e+05	4.14e+05	4.14e+05	4.14e+05	4.14e+05
R2	0.126	0.126	0.126	0.126	0.126	0.126

Acknowledgments

First of all, I would like to thank the Professors Elisabetta Raguseo and Francesco Luigi Milone, supervisors of this thesis, for the great availability, kindness and patience shown to me during the drafting of this project and for the precious advice and suggestions which have aroused in me a strong interest in the topic covered.

I am immensely thankful to my family for all they have done during these years for me, without them I'm sure I would not have arrived at this point, they have supported me economically and mentally for all these years.

I would like to say thanks also to the friends that never left me and gave me support during these years.

Bibliography

Adamiak, Czesław. "Changes in the global Airbnb offer during the COVID-19 pandemic." (2021).

Aguilera, T., Artioli, F., & Colomb, C. (2019). Explaining the diversity of policy responses to platform-mediated short-term rentals in European cities: A comparison of Barcelona, Paris and Milan. Environment and Planning A: Economy and Space, https://doi.org/10.1177/0308518X19862286

Airbnb (2020). A Message from Co-Founder and CEO Brian Chesky. Retrieved October 18, 2020, from https://news.airbnb.com/a-message-from-co-founder-and-ceo-brian-chesky/

Anselmi, Guido, and Veronica Conte. "Airbnb a Milano: geografie, concentrazione ed attori." (2021): 274-283.

Bakker, Matthijs. Airbnb listing prices in Amsterdam and distance to tourist spots. Diss. 2021.

Balaram, B. (2016, January 12), "Fair Share: Reclaiming power in the sharing economy"

Barron, Kyle, Edward Kung, and Davide Proserpio. "When Airbnb listings in a city increase, so do rent prices." Harvard Business Review 17 (2019): 1-62.

Baxter, E. and Bowen, D. 2004. Anatomy of tourism crisis: explaining the effects on tourism of the UK foot and mouth disease epidemics of 1967–68 and 2001 with special reference to media portrayal. International Journal of Tourism Research 6.

Benoit, S., Baker, T.L., Bolton, R.N., Gruber, T. and Kandampully, J. (2017), "A triadic framework for collaborative consumption (CC): motives, activities and resources & capabilities of actors", Journal of Business Research, Vol. 79, pp. 219-227.

Blal, I., Singal, M. and Templin, J. (2018), "Airbnb's effect on hotel sales growth", International Journal of Hospitality Management, Vol. 73, pp. 85-92.

Böckmann, M. (2013), The shared economy: It is time to start caring about sharing: value creating factors in the shared economy. Retrieved from http://thelovettcenter.com/wp-content/uploads/2014/11/bockmannshared-economy.pdf

Boto-García, David, and Veronica Leoni. "The hedonic value of coastal amenities in peer-to-peer markets." (2022).

Botsman, R., Rogers, R., Foley, K., Tantor Media. (2010), "What's mine is yours: The rise of collaborative consumption", Old Saybrook, Conn.: Tantor Media.

Bresciani, S., Ferraris, A., Santoro, G., Premazzi, K., Quaglia, R., Yahiaoui, D. and Viglia, G. (2021), "The seven lives of Airbnb. The role of accommodation types", Annals of Tourism Research, Vol. 88, p. 103170.

Bugalski, Łukasz. "The undisrupted growth of the Airbnb phenomenon between 2014–2020. The touristification of European Cities before the COVID-19 outbreak." Sustainability 12.23 (2020): 9841.

Caldieraro, F., Zhang, J.Z., Cunha, M., Jr. and Shulman, J.D. (2018), "Strategic information transmission in peer-to-peer lending markets", Journal of Marketing, Vol. 82 No. 2, pp. 42-63.

Christensen, C. and Raynor, M. E. (2003), The innovator's solution: Creating and sustaining successful growth, Harvard Business School Press, Boston.

Chua, Evelyn L., Jason L. Chiu, and Nelson C. Bool. "Sharing economy: An analysis of Airbnb business model and the factors that influence consumer adoption." Review of Integrative Business and Economics Research 8 (2019): 19.

Clauss, T. (2016). Measuring business model innovation: conceptualization, scale development, and proof of performance. R&D Management, 47(3), 385–403. https://doi.org/10.1111/radm.12186

Dolnicar, S. and Zare, S. (2020), "COVID19 and Airbnb–disrupting the disruptor", Annals of Tourism Research, Vol. 83, p. 102961.

DREDGE, Dianne and GYIMÓTHY, Szilvia. The collaborative economy and tourism: Critical perspectives, questionable claims and silenced voices. In: Tourism Recreation Research. 2015. Vol. 40, pages 286 302.

Edelman, B. and Luca, M. (2014), Digital discrimination: The case of Airbnb.com. Harvard Business School NOM Unit Working paper no. 14–054, Harvard Business School, Cambridge MA.

Frenken, K., Schor, J. (2017), "Putting the sharing economy into perspective", Environmental Innovation and Societal Transitions, 23, 3-10.

Gu, H. and Wall, G. 2006. Sars in China: Tourism impacts and market rejuvenation. Tourism Analysis 11.

Guttentag, D. (2015), "Airbnb: disruptive innovation and the rise of an informal tourism accommodation sector", Current issues in Tourism, Vol. 18 No. 12, pp. 1192-1217.

Guttentag, D. (2015), "Airbnb: disruptive innovation and the rise of an informal tourism accommodation sector", Current issues in Tourism, Vol. 18 No. 12, pp. 1192-1217.

Guttentag, D. and Smith, S. L. (2017), "Assessing Airbnb as a disruptive innovation relative to hotels: Substitution and comparative performance expectations", International Journal of Hospitality Management, Vol. 64, pp. 1-10.

Guttentag, D., Smith, S., Potwarka, L. and Havitz, M. (2017), "Why Tourists Choose Airbnb: A Motivation-Based Segmentation Study", Journal of Travel Research, pp. 1-18.

Guttentag, D., Smith, S., Potwarka, L. and Havitz, M. (2018), "Why tourists choose Airbnb: a motivation-based segmentation study", Journal of Travel Research, Vol. 57 No. 3, pp. 342-359.

Hall, C.M. 2010. Crisis events in tourism: subjects of crisis in tourism. Current Issues in Tourism 13.

Hamari, J., Sjöklint, M. and Ukkonen, A. (2016), “The sharing economy: why people participate in collaborative consumption”, Journal of the Association for Information Science and Technology, Vol. 67 No. 9, pp. 2047-2059.

Hamari, J., Sjöklint, M., Ukkonen, A. (2015), “The sharing economy: Why people participate in collaborative consumption”, Journal of the Association for Information Science and Technology, 67(9), 2047-2059.

Heinrichs, H. (2013), “Sharing Economy: A Potential New Pathway to Sustainability”, GAIA - Ecological Perspectives for Science and Society, 22(4), 228-231.

Henderson, J.C. and Ng, A. 2004. Responding to crisis: severe acute respiratory syndrome (SARS) and hotels in Singapore. International Journal of Tourism Research 6

Hossain, M. (2021), “The effect of the Covid-19 on sharing economy activities”, Journal of Cleaner Production, Vol. 280, p. 124782

Kirkos, Efstathios. "Airbnb listings’ performance: Determinants and predictive models." European Journal of Tourism Research 30 (2022): 3012-3012.

Lessig, L. (2008), “Remix: making art and commerce thrive in the hybrid economy”

Lim, J. and Won, D. 2020. How Las Vegas’ tourism could survive an economic crisis? Cities 100.

Matofska, B. (2016, September 1), “Welcome to the discovery site for the Sharing Economy and founders of Global Sharing Week”

May, S., Königsson, M. and Holmström, J. (2017). Unlocking the sharing economy: Investigating the barriers for the sharing economy in a city context. First Monday, Vol. 22, No. 2, pp. 1-17.

McAleer, M., Huang, B.-W., Kuo, H.-I., Chen, C.-C. and Chang, C.-L. 2010. An econometric analysis of SARS and Avian Flu on international tourist arrivals to Asia. Environmental Modelling & Software 25.

Möhlmann, M. (2015), “Collaborative consumption: determinants of satisfaction and the likelihood of using a sharing economy option again”, Journal of Consumer Behaviour, 14(3), 193-207.

Perez-Sanchez, V. Raul, et al. "The what, where, and why of Airbnb price determinants." Sustainability 10.12 (2018): 4596.

Petropoulos, Georgios. (2017), “An economic review of the collaborative economy (5)”, Retrieved from Bruegel Policy Contribution; No.2017/5 website: http://bruegel.org/2017/02/an-economic-review-of-t..

Pizam, A. (2014), "Peer-to-peer travel: Blessing or blight?", International Journal of Hospitality Management, Vol. 38, pp. 118-119.

PricewaterhouseCoopers. (2015), “The sharing economy”, Retrieved from https://www.pwc.com/us/en/services/consulting/library/consumer-intelligence-series/sharing-economy.html

Rassy, D. and Smith, R.D. 2013. The economic impact of H1N1 on Mexico’s tourist and pork sectors. Health Economics 22.

Richardson, L. (2015), “Performing the sharing economy”, Geoforum, 67, 121-129.

Roh, T. H. (2016), “The Sharing Economy: Business Cases of Social Enterprises Using Collaborative Networks”, Procedia Computer Science, 91, 502-511.

Sainaghi, Ruggero, and Rodolfo Baggio. "Airbnb Host Scaling, Seasonal Patterns, and Competition." Engineering Proceedings 5.1 (2021): 4.

Schor, J. B., Fitzmaurice, C. J. (2014), “Collaborating and connecting: the emergence of the sharing economy”, Handbook of Research on Sustainable Consumption, 410-425.

Sharpley, R. and Craven, B. 2001. The 2001 foot and mouth crisis – Rural economy and tourism policy implications: A comment. Current Issues in Tourism 4.

So, K. K. F., Oh, H. and Min, S. (2018), "Motivations and constraints of Airbnb consumers: Findings from a mixed-methods approach", Tourism Management, Vol. 67, pp. 224-236.

Statista (2022), “Revenue of platform providers in the global sharing economy 2017-2022”, available at: www.statista.com/statistics/878844/global-sharing-economy-revenue-platform-providers/

Stephany, A. (2015), “The business of sharing: Making it in the new sharing economy”, Houndmills, Basingstoke: Palgrave Macmillan.

STR, 2020. STR: Europe hotel performance for April 2020. Available online at https://str.com/press-release/ str-europe-hotel-performance-april-2020

Täuscher, K. and Laudien, S. M. (2017), "Understanding platform business models: A mixed methods study of marketplaces", European Management Journal, pp. 1-11, in press.

Tussyadiah, I. P. (2016), "Factors of satisfaction and intention to use peer-to-peer accommodation", International Journal of Hospitality Management, Vol. 55, pp. 70-80.

Tussyadiah, I. P. and Pesonen, J. (2016), "Drivers and barriers of peer-to-peer accommodation stay– an exploratory study with American and Finnish travellers", Current Issues in Tourism, pp. 1- 18.

Tussyadiah, I.P. and Pesonen, J. (2016), “Impacts of peer-to-peer accommodation use on travel patterns”, Journal of Travel Research, Vol. 55 No. 8, pp. 1022-1040.

Ugur, N. G. and Akbiyik, A. (2020). Impacts of COVID-19 on global tourism industry: A cross-regional comparison. Tourism Management Perspectives, Vol. 36, pp. 1-13.

UNWTO (2020). Tourism and COVID-19 – Unprecedented Economic Impacts. Retrieved November 15, 2020, from https://www.unwto.org/tourism-and-covid-19-unprecedented-economic-impacts

Varma, A., Jukic, N., Pestek, A., Shultz, C. J. and Nestorov, S. (2016), "Airbnb: Exciting innovation or passing fad?", Tourism Management Perspectives, Vol. 20, pp. 228-237.

Wang, Dan, and Juan L. Nicolau. "Price determinants of sharing economy based accommodation rental: A study of listings from 33 cities on Airbnb. com." International Journal of Hospitality Management 62 (2017): 120-131.

Wirtz, J., So, K.K.F., Mody, M.A., Liu, S.Q. and Chun, H.H. (2019), “Platforms in the peer-to-peer sharing economy”, Journal of Service Management, Vol. 30 No. 4, pp. 452-483

Wirtz, Jochen, et al. "Platforms in the peer-to-peer sharing economy." Journal of Service Management 30.4 (2019): 452-483.

Wu, E.H.C., Law, R. and Jiang, B. 2010. The impact of infectious diseases on hotel occupancy rate based on independent component analysis. International Journal of Hospitality Management 29.

Yao, B., Qiu, R. T., Fan, D. X., Liu, A., & Buhalis, D. (2019). Standing out from the crowd–an exploration of signal attributes of Airbnb listings. International Journal of Contemporary Hospitality Management, 31(12), 4520-4542.

Youssef, Adel Ben, Srdjan Redzepagic, and Adelina Zeqiri. "The key changes to the hospitality business model under COVID-19." Strategic Management-International Journal of Strategic Management and Decision Support Systems in Strategic Management 27.4 (2022).

Zhang, Zhihua, et al. "Key factors affecting the price of Airbnb listings: A geographically weighted approach." Sustainability 9.9 (2017): 1635.

Zervas, G., Proserpio, D., Byers, J.W., 2017. The rise of the sharing economy: Estimating the impact of airbnb on the hotel industry. Journal of Marketing Research doi:10.1509/jmr.15.0204.