



**Politecnico
di Torino**

Politecnico di Torino

Ingegneria Gestionale per l'Innovazione

A.a. 2021/2022

Sessione di Laurea Luglio 2022

L'impatto delle amenities sui prezzi in Airbnb

Analisi dei fattori che influiscono sui prezzi dei rentals sulla
piattaforma Airbnb con particolare interesse per le
amenities

Relatore:

Abrardi Laura

Candidato:

Combina Gabriele

*Un ringraziamento alla mia famiglia e alle persone
che mi hanno sostenuto in questi anni.*

Indice

Sommario

1. Introduzione	4
2. Analisi della Letteratura	6
2.1 <i>La Sharing Economy</i>	6
2.2 <i>Airbnb</i>	7
2.3 <i>Determinanti del prezzo nel settore turistico</i>	8
2.4 <i>Differenze tra Airbnb e hotels</i>	11
2.5 <i>Il ruolo delle amenities nel settore turistico</i>	11
2.6 <i>Il ruolo delle amenities in Airbnb e sull'influenza del prezzo</i>	13
3. Database di partenza	16
3.1 <i>Fonte dati</i>	16
3.2 <i>Composizione Database di partenza</i>	17
3.3 <i>Analisi approfondita dei dati usati</i>	18
4. Analisi dei dati	22
4.1 <i>Analisi del Database</i>	22
4.1.1 <i>Pulizia del Database preliminare</i>	23
4.2 <i>Analisi delle amenities</i>	24
4.2.1 <i>Amenities per proprietà</i>	24
4.3 <i>Variabili di controllo</i>	29
4.3.1 <i>Soddisfazione clienti</i>	29
4.3.2 <i>Posizione geografica</i>	30
4.3.3 <i>Conformazione della proprietà</i>	30
4.3.4 <i>Reputazione dell'host</i>	31
4.3.5 <i>Tipologie di stanza</i>	31
4.4 <i>Amenities</i>	32
4.4.1 <i>Singole amenities</i>	32
4.4.2 <i>Amenities raggruppate</i>	33
5 Regressione	34
5.1 <i>Introduzione</i>	34

<i>5.2 Correlazione tra prezzo/notte e rating e superhost</i>	<i>35</i>
<i>5.3 Correlazione tra prezzo/notte e caratteristiche della proprietà.....</i>	<i>37</i>
<i>5.4 Correlazione tra prezzo/notte e posizione geografica</i>	<i>39</i>
<i>5.5 Correlazione tra prezzo/notte e amenities</i>	<i>43</i>
<i>5.6 Regressione finale</i>	<i>45</i>
<i>5.6.1 Introduzione.....</i>	<i>45</i>
<i>5.6.2 Affitto di un'intera proprietà</i>	<i>45</i>
<i>5.6.3 Affitto di una singola stanza</i>	<i>47</i>
<i>5.7 Conclusione</i>	<i>49</i>
Bibliografia	50

1. Introduzione

Il fenomeno della Sharing Economy sta diventando sempre più rilevante nello scenario economico globale andando ad impattare in vari settori, dal turismo ai trasporti. Questo modello prevede l'acquisto, la fornitura o la condivisione dell'accesso a beni e servizi, andando così a creare un consumo collaborativo. L'economia della condivisione si è sviluppata grazie all'uso di piattaforme network in grado di connettere in maniera diretta, attraverso un sistema definito peer-to-peer, diverse categorie di utenti e permettendo di ottenere un guadagno dalle risorse sottoutilizzate, come ad esempio una seconda casa o i posti liberi nella propria auto durante una trasferta. Gli scambi avvengono attraverso piattaforme economiche intermediarie che permettono agli utenti, interagendo, di ottenere benefici di rete, si parla in questa situazione di mercato a due parti o rete a due parti (two-sided market).

Una delle principali piattaforme di sharing, nonché la più conosciuta nel settore turistico, è Airbnb. Nata nel 2008 a San Francisco ha in poco tempo rivoluzionato il mercato degli short term rentals, finendo per diventare uno dei principali competitor del settore in cui opera.

Con questa nuova modalità di affitto si sta osservando una trasformazione del metodo di scelta della location in cui alloggiare da parte dei turisti. Una delle tante variabili su cui si basano gli utenti per la decisione del luogo in cui pernottare e quella delle amenities, dotazioni e servizi messi a disposizione dagli ospitanti, che sono specifiche per questa tipologia di affitto (Abrate e Viglia, 2017) e differiscono da quanto viene solitamente offerto e richiesto nelle strutture turistiche tradizionali.

L'obiettivo della tesi è quello di analizzare come le amenities offerte, in unione con altre variabili di controllo, e le loro diverse tipologie influenzino i prezzi degli affitti presenti su Airbnb. Per farlo si analizzeranno i dati forniti dalla piattaforma e liberamente disponibili attraverso il sito [insideairbnb.com](https://www.airbnb.com) riguardanti la città di Milano.

Lo studio si svilupperà in tre parti:

- La prima è uno studio della letteratura già esistente, partendo dall'argomento turismo in generale per poi concentrarsi principalmente su quali siano i fattori che più influenzano i prezzi e come il fattore dotazioni impatti il settore del turismo e di Airbnb

- Nella seconda si fornirà una descrizione del dataset usato e delle scelte fatte per lo svolgimento delle analisi
- L'ultima sezione sarà dedicata alla presentazione del modello di regressione lineare multipla in cui viene valutato come le diverse variabili impattino sul prezzo delle proprietà presenti su Airbnb

2. Analisi della Letteratura

2.1 La Sharing Economy

Con il termine Sharing Economy si fa riferimento all'uso collaborativo dell'inventario sottoutilizzato attraverso la condivisione a pagamento di esso. La suddetta economia si è sviluppata fortemente nell'ultimo decennio grazie alle innovazioni tecnologiche nella comunicazione, quali la diffusione degli smartphones e dei social media, che hanno portato a quella che viene definita Digital Sharing Economy. Questa è fondata su un modello decentralizzato di scambio in cui gli individui interagiscono in maniera paritaria, senza che vi sia l'intermediazione di terze parti, chiamato Peer-to-Peer. Il P2P dà la possibilità ai consumatori di recensire pubblicamente i fornitori attraverso un sistema di valutazione, solitamente a "stelline" o con un valore numerico su una scala di misura predefinita e con un commento descrittivo di ciò che è stato messo a disposizione così da permettere di capire se le aspettative sono state mantenute. L'apparizione dell'economia collaborativa ha portato a risultati unici rispetto alla crescita del numero di utenti, potenziando aree innovative di interazione socioeconomica (Sundararajan, 2016).

I business basati su questa economia non risultano essere sempre profittevoli; infatti, devono scontrarsi con lo status quo del mercato che può presentare delle forti barriere all'entrata, un esempio può essere il caso di Uber che è stato bloccato direttamente dallo Stato in Italia per proteggere i taxisti e sta affrontando notevoli problemi legislativi in molti paesi.

Questo nuovo modello sta invece riscontrando un grande successo nel settore degli affitti a breve termine, facendo emergere l'interesse di molti per nuove tipologie di esperienze. Si sta assistendo sempre più ad una crescente ricerca di novità e di consumo collaborativo su piattaforme di sharing rental (Guttentag, 2015). È stata osservata una rivoluzione nella pratica del peer-to-peer lodging arrivando ad impattare alcune strutture turistiche tradizionali (Zervas et al., 2017).

2.2 Airbnb

Nel settore turistico la piattaforma di sharing che si è imposta sul mercato con maggiore efficacia è Airbnb, questa permette un'interazione diretta tra host che mettono a disposizione foto e dati della proprietà che sono intenzionati ad affittare e potenziali clienti, che possono anche osservare, oltre ai dati forniti dal ospitante, le valutazioni e i commenti dei precedenti turisti che hanno pernottato nella proprietà, ciò rende difficile per un proprietario inserire dati falsi senza venir scoperto.

Fondata nel 2008 negli Stati Uniti, ha iniziato a svilupparsi fortemente a partire dal 2011 per poi raggiungere la popolarità anche oltre oceano dal 2013 grazie alla creazione di un sistema facile da usare ed intuitivo anche per le persone tecnologicamente meno esperte, finendo per diventare ad oggi il maggior concorrente sul mercato. Grazie alle sue speciali caratteristiche, quali internet business model, enorme numero di hosts e guests, bassi costi, fornitura di servizi domestici e il potenziale per un'esperienza locale più autentica (Guttentag, 2015) Airbnb sta godendo di una crescita esponenziale, recentemente bloccata a causa dell'epidemia di Coronavirus. Altre caratteristiche che hanno contribuito al suo rapido aumento di popolarità hanno a che vedere con gli aspetti esperienziali unici che può offrire un viaggio fatto sfruttando questo tipo di piattaforma, come il senso di comunità e il contatto umano (Yannopoulou et al., 2013), questi risultano estremamente difficili da trovare decidendo di pernottare in un classico hotel.

Degli studi hanno osservato che la presenza di piattaforme come Airbnb hanno un impatto molto positivo non solo perché portano un cambiamento dello scenario competitivo nell'industria degli hotel, ma anche nell'industria del turismo più in generale riuscendo ad attirare nuove tipologie di clienti e favorendo l'innovazione nel mercato (Heo et al., 2019).

Airbnb è diventata nell'arco di poco tempo oggetto di interesse per i ricercatori, oltre che per potenziali imprenditori. La sua popolarità è dovuta alla possibilità di sfruttare ambienti domestici inutilizzati, ottenendo sia un guadagno economico per gli ospitanti che esperienziale per i turisti. La rapida ascesa di Airbnb ha innescato un flusso di ricerche che osservano diversi aspetti di questo fenomeno, come i prezzi e i dati demografici degli utenti, nonché le conseguenze sulle prestazioni del settore

alberghiero, sulla disoccupazione, sul valore delle proprietà e persino sulla pianificazione urbana (Kirkos, 2021); l'interesse dei prossimi paragrafi si concentra principalmente sulle ricerche che descrivono le determinanti del prezzo, tra cui ci sono le amenities, utili per un confronto con i risultati presentati nei capitoli successivi.

2.3 Determinanti del prezzo nel settore turistico

Come da regola di mercato i clienti mettono in relazione i prezzi con i servizi messi a disposizione (Wang e Nicolau, 2017), quindi esistono moltissime analisi che cercano di comprendere al meglio come diversi fattori incidano su questo nelle diverse strutture turistiche.

Per ciò che riguarda gli hotel esistono molteplici studi che osservano le determinanti del prezzo di una stanza dal punto di vista della domanda (Becerra et al., 2013; Lee e Jang, 2012) e dal punto di vista dell'offerta (Heo e Hyun, 2015; Masiero et al., 2015). Negli ultimi anni sono in forte aumento studi simili che riguardano le piattaforme di sharing, ci sono analisi più generali che tengono in considerazione molteplici variabili (Benítez-Aurioles, 2018; Qiu et al., 2018) e altri che si concentrano su aspetti più specifici come, ad esempio, l'importanza del Superhost (Gunter, 2018); host che hanno un badge di qualità, fornito in base a dei parametri decisi da Airbnb, visibile dai potenziali ospiti durante le loro ricerche.

In linea generale si è osservato che i prezzi su Airbnb tendono ad essere più bassi rispetto a quelli che si trovano negli hotels a parità di esperienza (Festila e Muller, 2017) e questo si è rilevato essere uno dei maggiori punti di forza della piattaforma; infatti, secondo tutte le analisi finora prodotte questa è una delle principali motivazioni, se non la principale, per cui i clienti decidono di passare dalle strutture tradizionali a questa nuova tipologia di turismo.

Le variabili che sono state considerate quando si è cercato di creare un modello di previsione del prezzo nei diversi papers sono:

- Rapporto con l'host: possibilità di comunicarci direttamente, superhost, tempi di risposta
- Recensioni e punteggi
- Luogo geografico rispetto ai punti di interesse della città

- Servizi e dotazioni offerti all'interno della proprietà
- Tipologia di sede di pernottamento: numero di persone ospitabili, necessità di condividere degli spazi con sconosciuti, numero di stanze, tipologia di struttura di ricezione

Nella seguente tabella è possibile osservare una descrizione di quali fattori e metodi siano stati presi in considerazione per la previsione del prezzo in Airbnb in alcuni dei più importanti studi riguardo questo topic.

Autore	Metodo	Determinanti del prezzo
Amaro et al., 2018	Regressione PLS	Intenzioni di prenotare su Airbnb, atteggiamento verso Airbnb, aorma soggettiva, rischio percepito, benefici economici, desiderio di alloggi unici e varietà
Benítez-Aurioles, 2018	Regressione OLS	casa intera, stanza privata, camere da letto, bagni, numero persone, Superhost, numero di recensioni, punteggio, cancellazione flessibile
Gunter, 2018	Modelli logit e probit	Criteri del Super host, determinanti della domanda
Liang et al., 2018	Structural equation modelling (SEM)	Soddisfazione basata sulle transazioni, soddisfazione basata sull'esperienza, fiducia verso Airbnb, fiducia negli host, intenzione di cambio, intenzione di riacquisto
Mauri et al., 2018	Regressione Shapley Value	Variabili di popolarità insieme agli attributi reputazionali personali e alla descrizione dell'offerta
Qui et al., 2018	Binomial logistic model	31 variabili: inclusi il numero di recensioni, posizione geografica della struttura in affitto, amenities
Abrate e Viglia, 2017	Regressione di frontiera	Variabili relative alla reputazione
Gunter e Önter, 2017	Cluster-Robusto minimi quadrati ordinari	Numero e contenuto recensioni, location, tempo di risposta
Horn e Merante, 2017	Regressione	Attributi presi dai database di Airbnb
Liang et al., 2017	Regressione di Poisson, modello di Tobit	Numero recensioni, punteggi, superhost badge, capacità, stanze, letti, possibilità di cancellare, foto, descrizione

Zhang et al., 2017	General Linear Model, Geographically Weighted	Prezzo, anni di presenza su Airbnb, distanza dai luoghi di interesse, punteggi, recensioni
Ert et al., 2016	Regression	Attributi della struttura interessata, comprese le amenities, e attributi reputazionali dell'host
Varma et al., 2016	ANOVA	Domande di un sondaggio sui fattori per la selezione di una struttura ricettiva in Airbnb, caratteristiche delle locations

Tab.1 Lista di papers che studiano la correlazione del prezzo dei rentals sulla piattaforma Airbnb a diverse variabili indipendenti, creata da Chattopadhyay et al. 2019

2.4 Differenze tra Airbnb e hotels

Gli annunci di Airbnb sono correlati a proprietà inutilizzate di host non professionisti; quindi, è necessario rivedere le determinanti che influiscono sui prezzi rispetto a come si analizzano gli alberghi tradizionali (Botsman e Rogers, 2011).

Per valutare correttamente le piattaforme P2P bisogna anche cambiare il metodo di analisi; infatti, concetti come l'affiliazione al marchio e il prestigio di una struttura ricettiva non sono adatti alle offerte di affitto poiché la maggior parte delle proprietà di Airbnb sono residenze personali (Guttentag, 2015) e ci sono variazioni maggiori di prezzo, questo rende meno immediati i confronti.

Sono presenti delle analisi che comparano come diversi attributi incidano sulla soddisfazione dei clienti negli hotels e nelle piattaforme P2P, ad esempio è stato riportato che gli ospiti in Airbnb valutano come più rilevante il rapporto host-guest e il quartiere in cui decidono di stare, invece i servizi della stanza, cibo e bevande influenzano maggiormente i clienti degli hotel (Belarmino, 2017). Un'altra ricerca condotta su 630 persone che hanno usato entrambi i metodi di short-term rental ha concluso che Airbnb è superiore agli hotel per ciò che concerne le interazioni host-guest, amenities nella struttura ricettiva, tipicità (localness) e personalizzazione mentre gli altri attributi sono tendenzialmente simili (Mody et al., 2017).

2.5 Il ruolo delle amenities nel settore turistico

Le amenities risultano essere una delle variabili principali nella scelta del luogo in cui pernottare per i turisti, questa affermazione è valida per qualunque tipo di struttura

venga scelta. Il ruolo che hanno è difficile da comparare nelle diverse tipologie in quanto negli hotel vengono valorizzati servizi che magari non hanno senso di esistere in proprietà di sharing rental, ad esempio il minibar, e allo stesso tempo chi decide di usare Airbnb potrebbe essere più interessato ad avere amenities che ricordino la propria casa, come la cucina o la lavatrice.

Da una ricerca che compara ciò che influisce maggiormente sulla soddisfazione dei clienti tra strutture tradizionali e rentals presenti sulla piattaforma Airbnb, basandosi sui commenti da questi rilasciati, risulta che chi pernotta negli hotels sia più interessato alle dotazioni offerte rispetto a chi si affida allo sharing rental, sia per ciò che concerne le general amenities sia per le room amenities (Belarmino et al. 2017).

Per comprendere l'impatto dei servizi offerti sulla soddisfazione dei clienti di hotels sono presenti vari studi che li considerano nei loro modelli di determinazione del prezzo (Becerra et al., 2013; Chen and Rothschild, 2010; Salò et al., 2014). L'offerta di amenities quali express checkout, mini-bars, televisione, colazione, cassette di sicurezza, prenotazione anticipata, asciugacapelli e l'alto rapporto tra governanti e ospiti hanno un'influenza molto elevata sul prezzo della camera (Yang et al., 2016; Schamel, 2012; Masiero et al., 2015; Thrane, 2007; Lee and Jang, 2012). Altri servizi, quali ad esempio internet, si sono rivelati inconsistenti nell'influenza del prezzo (Yang et al., 2016). La letteratura esistente ha anche riportato che le caratteristiche fisiche dell'hotel come il numero di camere, l'età dell'hotel, la disponibilità di un business center, bar e piscina sono fattori che incidono sul prezzo (Yang et al., 2016; Thrane, 2007; Lee e Jang, 2012; Becerra et al., 2013; Chen e Rothschild, 2010). Allo stesso tempo, il parcheggio e la palestra si sono rivelati essere costantemente determinanti per i prezzi più alti (Chen e Rothschild, 2010; Yang et al., 2016).

Un gran numero di clienti sceglie di usare Airbnb a causa dei servizi che sono un importante determinante del prezzo e di soddisfazione dell'ospite (Wang e Nicolau, 2017; Gibbs et al., 2018). Questi servizi sono in parte diversi da quelli resi disponibili negli hotel in quanto i clienti di Airbnb collegano servizi come lavatrice, cucina, asciugatrice, frigorifero e simili a ciò che hanno normalmente a disposizione nella loro casa (Chattopadhyay, 2019).

Attualmente non esistono studi che prendono in considerazione le amenities come variabile principale per il modello di determinazione del prezzo in Airbnb, uno degli obiettivi di questa tesi sarà proprio questo tipo di analisi.

2.6 Il ruolo delle amenities in Airbnb e sull'influenza del prezzo

Airbnb offre la possibilità di alloggiare in strutture estremamente differenti tra loro e che si adattano ad ogni tipologia di esperienza ricercata. Questa grande varietà si ripercuote anche sulle dotazioni e i servizi messi a disposizione, che possono essere di diverse tipologie e più o meno rari, ad esempio la piscina è molto più rara rispetto alla TV ed è considerata un servizio di lusso, ma allo stesso tempo la TV è un bene che molti vogliono avere. Ci sono anche servizi che molti utenti considerano essenziali, come ad esempio il WiFi; la mancanza di questi potrebbe portare ad una notevole diminuzione dei potenziali clienti.

Gli host che usano questa piattaforma devono sempre più affrontare nuove sfide; infatti, come Gibbs et al. (2018) sottolineano: <<A causa della crescita esponenziale della piattaforma, gli host stanno affrontando una maggiore concorrenza e quindi le loro pratiche di marketing richiedono approcci più informati.>>. Essendo non professionisti per gli ospiti risulta difficile capire come agire in un mercato competitivo, che nella maggior parte dei casi conoscono solo in maniera superficiale, riuscendo a massimizzare i propri benefici.

La varietà di scelta unita al fatto che gli host non possano contare su una brand identity del locale che decidono di affittare, a differenza di buona parte degli hotels, comporta che debbano scegliere con accuratezza le amenities che intendono offrire così da poter attirare il maggior numero di clienti possibile e allo stesso tempo riuscire a differenziarsi dagli altri host. Questa scelta è così rilevante che la stessa Airbnb rilascia spesso su un blog dedicato consigli su quali dotazioni offrire, quali vanno più di moda, quali sono ritenuti indispensabili e così via. Per gli ospiti è particolarmente difficile capire quali dotazioni mettere a disposizione in quanto non essendo esperti del settore, ma solitamente si tratta persone che decidono di mettere in affitto una seconda casa per non lasciarla inutilizzata, non sanno quanto valore aggiunto può offrire un bene e se i ritorni dell'eventuale investimento siano maggiori dei costi. Per scegliere quali amenities

offrire e se fare un eventuale investimento in una o più di queste la soluzione per molti risulta essere quella di fare un confronto con i concorrenti, cioè gli altri appartamenti della zona offerti sulla piattaforma, per poi adeguarsi. Questo metodo presenta dei potenziali problemi:

- c'è il rischio che tutti finiscano per imitarsi
- il campione di confronto è piccolo e questo potrebbe portare a decisioni sbagliate
- non è detto che gli affittuari della zona prendano decisioni corrette sulla scelta delle amenities da offrire è più in generale su come impostare l'annuncio
- con il tempo cambiano le richieste dei clienti e con questo metodo risulta particolarmente difficile tenersi aggiornati

Uno degli obiettivi di questa ricerca è anche quella di dare dei possibili consigli agli host basandosi su analisi più strutturate, cercando di capire se determinate categorie di dotazioni o singole amenities abbiano un impatto positivo sui prezzi.

Ci sono alcuni studi che si concentrano solo su alcune amenities (Gibbs et al., 2018) e altri che cercano di capire più in generale se queste abbiano un qualche tipo di influenza sulla soddisfazione dei clienti e sui prezzi (Wang and Nicolau, 2017). Oltre ai servizi ovviamente vengono considerate anche altre variabili, altrimenti i modelli ottenuti non sarebbero attendibili.

Per ciò che concerne le amenities nello studio di Gibbs ne sono state valutate tre in particolare: parcheggio, piscina e palestra, arrivando a concludere che queste creavano un valore aggiunto per i clienti. Gli effetti sul prezzo però sono diversi: la palestra è quella che incide maggiormente sull'aumento del costo che i clienti sono disposti ad accettare, la piscina ha un effetto positivo in tutte le città che sono state analizzate anche se in maniera inferiore rispetto alla palestra, lo stesso invece non si può dire per il parcheggio, questo ha portato benefici in termini economici in solo due delle cinque città valutate.

Il paper di Wang e Nicolau invece analizza la correlazione tra 5 categorie di variabili, una di queste sono le dotazioni e i servizi offerti, e il prezzo per 180.533 alloggi messi in affitto su Airbnb attraverso due regressioni: OLS e QR. Secondo questo studio le amenities sono buoni perditori di prezzo e hanno un'influenza positiva su questo anche se l'effetto varia a seconda del range di prezzo considerato.

Un ulteriore esempio di paper che trova una correlazione positiva tra servizi domestici e prezzo è “Why tourists choose Airbnb: a motivation-based segmentation study” di Guttentag et al. del 2018 in cui vengono intervistati 800 turisti che hanno usato il servizio offerto dalla piattaforma. Si osserva che i rispondenti concordano con il fatto che le amenities hanno un ruolo importante nell’esperienza avuta, tanto da essere considerate una delle due più importanti variabili, insieme al risparmio, per cui preferire Airbnb agli hotel più tradizionali.

Diversamente da alcuni dei paper appena citati in un recente studio (Kirkos, 2021) in cui vengono osservate le prestazioni degli annunci Airbnb in termini di tasso di occupazione, numero di prenotazioni e entrate, utilizzando metodologie di data mining giunge alla conclusione che è maggiormente rilevante il numero di amenities rispetto a quali effettivamente siano presenti, si legge: <<Per ciò che riguarda gli attributi relativi all’alloggio, solo la variabile ‘totale servizi’ ha un forte impatto positivo, suggerendo che i clienti non si concentrano su servizi particolari ma apprezzano le disposizioni complessive.>> e ancora: <<Rimarchevolmente, sia la fornitura di servizi che gli alti rating aumentano il valore percepito dai potenziali clienti>>. Se le due precedenti frasi risultassero vere significherebbe che esiste rapporto positivo tra servizi offerti e prezzo, ma considerando determinate categorie di amenities o singole dotazioni non dovrebbero notarsi significative differenze tra queste in rapporto ai prezzi.

Ci si è anche posti il quesito su quali tipologie di dotazioni e servizi siano di maggiore interesse per i clienti: gli attributi pratici, quali le dotazioni e il risparmio rispetto alle strutture tradizionali, o quelli esperienziali, come la condivisione della camera di pernottamento con altri viaggiatori e il rapporto con l’host. Nell’analisi di Guttentag del 2018 dove sono stati intervistati oltre 800 utenti Airbnb, principalmente in Canada e Stati Uniti; si è concluso che vantaggi come prezzo, posizione e servizi domestici avessero più rilevanza in confronto a fattori quali interazione sociale, autenticità e novità, anche se questi sono comunque importanti per alcuni ospiti. Andando ulteriormente nel particolare, la ricercatrice Iis Tussyadiah (Tussyadiah, 2016) ha rilevato che l’attributo ‘location’ (camera condivisa, camera singola, appartamento privato), non risulta essere statisticamente significativo per le persone che decidono di usare Airbnb, al contrario, amenities e cost saving appaiono avere un’influenza positiva.

L'unico lavoro che sembra andare contro una visione generale di correlazione positiva tra prezzo e amenities e che si contrappone anche alle conclusioni dei due ultimi papers citati, secondo i quali gli attributi pratici risultano più importanti di quelli esperienziali per i clienti, è "An Analysis of Social Features Associated with Room Sales of Airbnb" di Donghun Lee del 2015 in cui sono stati analizzati i dati 4178 stanze dove si conclude che le dotazioni offerte non siano tra i fattori principali che incidono sulle scelte dei clienti, ma questi vengono principalmente influenzati da fattori sociali quali la reattività dell'host, il numero delle recensioni e l'anzianità di iscrizione.

Esistono diversi papers che correlano prezzo e amenities, come visto nei paragrafi precedenti, e in quasi tutti viene osservata una correlazione positiva, però non esiste alcuno studio che consideri le dotazioni offerte o alcune sottocategorie di queste come fattore principale del loro modello. Probabilmente questo è dovuto al fatto che le altre determinanti hanno un impatto troppo importante per essere considerate delle semplici variabili di controllo e si corre il rischio di avere una visione troppo limitata finendo per commettere degli errori. Nei seguenti capitoli si proverà a fare questo tipo di analisi, cercando in particolare di osservare se determinate tipologie di servizi offerti dall'host incidano più di altre sulla soddisfazione dei clienti (rating) e sui prezzi.

3. Database di partenza

3.1 Fonte dati

Il set di dati che adesso verrà presentato è quello usato nei prossimi capitoli per svolgere le analisi di interesse, in particolare si osserveranno i fogli di lavoro e le variabili a disposizione, comprese quelle che si è deciso di non usare nel modello statistico così da dare una visione più generale.

La sorgente dei dati è il sito Inside Airbnb (www.insideairbnb.com) in cui vengono pubblicate e sono visionabili gratuitamente le informazioni riguardanti tutti gli host e delle relative strutture che sono presenti sulla piattaforma Airbnb nelle più grandi città del mondo.

Una volta entrati nel sito si può accedere a svariati set di dati delle città di interesse, questi dati vengono forniti in formato Microsoft Excel Comma Separated Values File, comunemente conosciuti come file “.csv”. Come detto i dati vengono separati geograficamente, quindi per lo svolgimento delle analisi è stato necessario di scegliere una città, questa scelta è ricaduta su Milano. Tra le città italiane Milano è certamente una di quelle con più turisti e questi si caratterizzano per una generale eterogeneità di interessi, da coloro che la esplorano come vacanzieri a chi la deve visitare la città per motivi lavorativi, e di origini, sono presenti persone che arrivano da tutto il mondo; ciò permette di avere un campione dove eventuali bias tendono ad essere eliminati. Inside Airbnb permette inoltre di visualizzare gli annunci all'interno di una mappa della città della quale si è interessati.

Per ogni città è presente uno storico dei dati, per questa analisi si è deciso di lavorare usando solo quelli più recenti prendendo in considerazione solo le strutture che presentano almeno una recensione tra il 2019 e l'ottobre 2021.

3.2 Composizione Database di partenza

Una volta scelto il luogo per l'analisi di interesse si ha accesso ai seguenti file di informazioni (alcune risultano ripetute nei vari fogli di analisi, ma in realtà questo non è un fattore negativo in quanto rende più veloce il processo di pulizia del database e l'utilizzo di molteplici fogli dati in contemporanea):

- *Detailed Listings for:*
 - Nella matrice presente in questo documento è possibile reperire tutte le informazioni disponibili riguardanti gli annunci, gli host, prezzi e recensioni con delle relative sottocategorie. I dati presenti in questo file e alcune loro rilavorazioni saranno usati per la creazione del modello di regressione, per questo motivo si è deciso di farne una descrizione più approfondita nei seguenti paragrafi
- *Detailed Calendar Data for listings in:*
 - In questo documento sono presenti informazioni riguardanti la disponibilità per ogni proprietà, il prezzo e il numero di notti massime e minime concesse dall'host
- *Detailed Review Data for listings in:*
 - Contiene tutte le recensioni lasciate dagli ospiti associate alla struttura ricettiva da loro usata con la relativa data
- *Summary information and metrics for listings in:*
 - In questo documento sono visibili ogni annuncio per le informazioni generali dell'host, quelle relative alla posizione geografica (latitudine e longitudine) e metriche riguardanti le recensioni ricevute (totali, mensili, ecc...)
- *Summary Review data and Listing ID:*
 - Viene dato l'ID di ogni appartamento e la data di iscrizione alla piattaforma Airbnb, può essere usato per svolgere delle analisi temporali
- *Neighbourhood list for geo filter:*
 - Per ogni struttura è presente il quartiere e il sotto quartiere, nel caso in cui questo dato fosse presente, il foglio può essere usato per svolgere analisi geografiche. Il quartiere è fornito anche in altri documenti, ma come detto in precedenza sono presenti delle ridondanze che semplificano l'uso contemporaneo di più documenti
- *GeoJSON file of neighbourhood of the city:*
 - Questo documento contiene gli stessi dati del file precedentemente presentato ma in formato GeoJSON

3.3 Analisi approfondita dei dati usati

Il foglio di lavoro che sarà al centro del modello statistico è “Detailed listing for Milan”, per questo ne verrà data una descrizione dettagliata dei parametri che si trovano al suo interno, compresi quelli che non sono stati usati, così da permettere di comprendere come si è svolta la fase di pulizia del database.

In ogni riga della matrice presente nel documento vengono forniti tutti i dati relativi ad una proprietà messa in affitto su Airbnb.

A partire dalla colonna a sinistra i primi dati riguardano la proprietà, identificata attraverso un ID e collegata all’annuncio da un URL associato. Poi si passa ai dati inseriti dall’host riguardo la struttura, infatti questo deve darle un nome, una descrizione, indicare in che zona si trova e inserire delle foto, è presente una colonna con l’URL di quest’ultime. Gli ospitanti sono particolarmente attenti a questi dati in quanto risultano importanti per attirare più clienti possibile o farsi notare da determinate categorie di utenti (es. i giovani o chi arriva in città per lavoro), nel contempo non possono permettersi di mentire, abbellendo troppo l’annuncio, altrimenti si troveranno molte recensioni di persone insoddisfatte e bassi voti finendo per ottenere l’effetto opposto a quello cercato.

Successivamente vengono presentati i dati del host, quali: il codice identificativo assegnato dalla piattaforma, URL associato al suo profilo, il nome, da quando si è iscritto al servizio, dove vive e una breve descrizione di sé, anche in questo caso l’obiettivo è quello di presentarsi al meglio. Sapere se l’host vive nella stessa città della proprietà che affitta potrebbe essere un punto di interesse per alcuni utenti in quanto permette potenzialmente di avere un rapporto umano superiore anche a quello che si trova negli hotel; inoltre, risulta più facile reperire informazioni come, ad esempio, dove andare a mangiare e che luoghi visitare al di fuori dei classici posti per turisti. Alcuni ospitanti decidono di non inserire nulla nella propria descrizione, ciò potrebbe diminuire la fiducia da parte dei potenziali guest nei loro confronti.

Una volta presentati i dati della struttura e al host si passa alle informazioni riguardo il comportamento di quest’ultimo, quali il tasso di accettazione, il tasso e il tempo di risposta, per molte strutture in realtà questi dati non compaiono, e se è un Superhost.

I Superhost sono gli ospitanti che risultano più meritevoli secondo Airbnb, a questi viene concesso un badge visibile affianco ai loro annunci, ciò può attirare potenziali clienti, in particolare quelli più preoccupati per possibili truffe o che vogliono essere più sicuri di avere un'esperienza positiva, magari anche accettando di pagare un prezzo leggermente più alto. Per ottenere questo riconoscimento vanno rispettate delle caratteristiche da parte di un host:

- Aver completato almeno 10 soggiorni e 3 prenotazioni, per un totale di almeno 100 notti
- Aver mantenuto un tasso di risposta minimo del 90%
- Aver conservato un tasso di cancellazione massimo dell'1%, fatta eccezione per i casi che rientrano nei termini delle circostanze attenuanti di Airbnb
- Aver mantenuto una valutazione complessiva di 4,8 su 5 negli ultimi 365 giorni, prendendo in riferimento le date in cui gli ospiti hanno scritto la recensione e non a quella in cui hanno effettuato il check-out

All'ospitante viene poi data la possibilità di mostrare una propria foto, nel caso compare il relativo URL nella matrice.

Finite le colonne con i dati di identificazione della proprietà e relative al padrone di casa si passa a quelle che probabilmente sono più importanti per un cliente, riguardanti la posizione geografica e come è composta la struttura di interesse. Per ciò che riguarda la posizione vengono forniti il quartiere, la latitudine e la longitudine.

Per descrivere la location sono presenti come variabili: tipologia di proprietà (villa, appartamento, ecc...), tipologia di stanza (condivisa o privata), numero di persone accettabili, di bagni, di stanze e di letti.

A questo punto si passa alla variabile centrale delle analisi effettuate, le amenities. Queste vengono fornite direttamente dagli host, i quali creano una lista con tutte le dotazioni messe a disposizione. Come si vedrà più approfonditamente nel capitolo successivo molti inseriscono dotazioni essenzialmente irrilevanti, questo ha reso alcune analisi potenzialmente meno efficaci, ad esempio si è deciso di considerare come variabile del modello presentato successivamente il numero di amenities presenti in ogni struttura in quanto questo valore viene visualizzato nell'annuncio e può influire

sulla scelta dei guest, ma in realtà non rappresenta il reale numero di dotazioni e servizi che interessano i clienti.

Le successive informazioni date sono: prezzo per notte, che sarà la variabile dipendente nel modello, numero minimo e massimo di notti in cui si può pernottare e durata media dei pernottamenti.

Viene poi indicato se è possibile se la proprietà è immediatamente disponibile alla prenotazione (nella data di estrazione del database), quando è stata effettuata l'ultima prenotazione per ogni struttura ed è data una panoramica dell'ultimo anno, fornendo le date in cui l'host ha reso accessibile l'abitazione, andando nel particolare ad isolare i dati relativi agli ultimi 30, 60 e 90 giorni dall'ultimo aggiornamento del report.

Delle diciannove colonne rimanenti la maggior parte sono dedicate alle reviews e i relativi rating. La prima informazione disponibile riguarda il numero di recensioni totale, andando poi più nel particolare a definire quante di queste siano state fornite nell'ultimo anno e negli ultimi trenta giorni, poi viene fornito un loro collocamento temporale attraverso le date della prima e dell'ultima recensione per ogni struttura; nell'ultima colonna è anche presente il numero di recensioni mediamente rilasciate al mese da quando la proprietà è stata inserita sulla piattaforma.

Successivamente vengono esposti i rating assegnati, si parte con il "review_scores_rating" che indica il punteggio complessivo delle strutture; questo è ricavato dall'unione, fatta attraverso un algoritmo gestito direttamente da Airbnb, dei punteggi che forniscono i clienti. Alla fine di un'esperienza viene richiesto al guest di valutare separatamente: accuratezza della descrizione dell'alloggio presente nell'annuncio, pulizia al momento dell'entrata nella proprietà, fase di check-in, comunicazione con il padrone di casa o con chi fa le sue veci, posizione e rapporto qualità-prezzo.

Dopo i dati riguardanti i rating viene specificata la policy di cancellazione scelta dall'ospitante; infatti, Airbnb concede la possibilità a chi inserisce una proprietà sulla piattaforma di scegliere una delle otto policy standard, queste differiscono tra loro per tolleranza. L'informazione viene fornita ai clienti all'interno dell'annuncio con la possibilità di vedere le relative clausole, ciò risulta particolarmente comodo per persone che rischiano di dover annullare il viaggio.

Le ultime informazioni nella matrice specificano per l'host della struttura considerata quante altre proprietà ha quest'ultimo in rental su Airbnb, prima dando il totale e poi suddividendolo in case intere indipendenti, stanze singole all'interno di una struttura condivisa e locali affittabili solo in condivisione.

4. Analisi dei dati

4.1 *Analisi del Database*

Come detto precedentemente il database di partenza di trova su InsiderAirbnb.com, un sito open source che consente di visionare tutti i dati delle proprietà divise per città. Si è deciso di prendere in esame Milano, in quanto sono presenti un grande numero di rentals di diverse tipologie e ogni anno molti turisti, con interessi e necessità diverse, visitano la città.

I dati forniti dal sito vengono presentati attraverso vari fogli di lavoro in formato Microsoft Excel Comma Separated File, di quelli a disposizione si è usato solo il “Detailed listing for Milan” in quanto contiene tutte le informazioni principali riguardo la proprietà, l’host, la soddisfazione dei clienti che hanno pernottato nelle diverse strutture ricettive e infine le amenities, che saranno al centro delle analisi.

I tool usati per effettuare lo studio sono due: Microsoft Excel e Stata. Il primo è risultato particolarmente comodo visto il formato con cui venivano forniti i dati e ha permesso di effettuare la pulizia del database e la creazione di variabili usate nel modello di regressione. Per la creazione di una regressione lineare multipla con decine di variabili si è optato per il software Stata, questo è usato da ricercatori in vari ambiti per la manipolazione dei dati e la creazione di report automatici.

4.1.1 *Pulizia del Database preliminare*

Il foglio di lavoro, nel momento in cui è stato scaricato dal sito, presenta tutte le proprietà in rental dal 19/04/2010 al 19/10/2021, comprese quelle che non sono più in affitto da molto tempo. Si è deciso di escludere quest’ultime prendendo in considerazione esclusivamente le strutture che fossero state messe a disposizione negli ultimi tre anni, dal 01/01/2019 al 19/10/2021. In aggiunta alle locations con una data considerata troppo “vecchia” rispetto a quella che si è deciso di prendere come limite si sono anche eliminate le strutture in cui la data dell’ultima recensione non era segnalata. Il documento di partenza conteneva le informazioni di 17230 locations, conseguentemente a questo prima scrematura si è passati ad averne 9401. Questo passaggio è stato eseguito per avere solo dati ottenuti in situazioni simili; infatti, è

difficile e inaccurato paragonare degli alloggi che non sono più sul mercato da molti anni con gli attuali in quanto nell'arco di un decennio i prezzi nella città sono aumentati notevolmente, così come sono aumentate anche le aspettative da parte dei clienti.

Successivamente alla pulizia temporale del database si è deciso quali variabili fossero potenzialmente utili e quali sicuramente eliminabili, così da effettuare un'ulteriore scrematura dei dati andando a creare un sotto-database più facile da analizzare, passando da 74 a 27 variabili. In realtà, si è poi deciso di considerare una parte di queste.

Le variabili che si sono tenute in considerazione riguardano:

- Identificazione della proprietà (id e nome)
- Identificazione dell'host (id e se è superhost)
- Locazione della proprietà (quartiere e coordinate)
- Caratteristiche della proprietà (numero persone ospitabili, camere, letti, bagni e amenities)
- Prezzo per notte
- Rating con le relative sottocategorie

L'ultimo passaggio di questo processo è consistito nell'eliminare le proprietà che consentivano solamente l'affitto con degli sconosciuti nella stessa stanza (Shared Room), perché sono in numero troppo basso rendendo impossibile svolgere un'analisi statistica significativa, e gli hotels, anche questi in numero troppo basso per una regressione e non inerenti allo studio di rentals effettuato da non professionisti. In conclusione, il numero di proprietà usabili per il modello è 9350.

4.2 Analisi delle amenities

La base per le analisi delle dotazioni e dei servizi offerti nelle varie strutture è la colonna della matrice del documento riferimento chiamata "Amenities". Questa contiene la lista delle suddette per ogni proprietà all'interno di parentesi quadre, con delle "" e la virgola a separarle, ad esempio la proprietà con id 23765630 contiene le seguenti amenities:

```
["TV", "Host greets you", "Carbon monoxide alarm", "Washer", "Bed linens", "Iron", "Long term stays allowed", "Essentials", "Hot water", "Kitchen", "Air conditioning", "Heating", "Hangers", "Free parking on premises", "Wifi", "Smoke alarm", "Hair dryer", "Dedicated workspace"]
```


4.2.1 Amenities per proprietà

Per ogni proprietà si è osservato che mediamente nell'annuncio vengono indicate 24,86 amenities con una deviazione standard di 10,17, la variabilità risulta essere quindi molto ampia tra le diverse strutture ricettive. In particolare, si nota che è presente un gruppo ristretto di rentals, una trentina circa, che ha solo una o due amenities e in buona parte di questi casi l'unica amenity è: possibilità affittare per un lungo periodo. Probabilmente questi host decidono di affidarsi semplicemente alle foto e alla descrizione generale della proprietà per attirare clienti. Com'è possibile osservare nelle seguenti tabelle la tendenza generale invece è opposta, cercare di mettere più dotazioni e servizi possibili con la speranza di attirare i guest, anche aggiungendo dotazioni inutili solo "per far numero".

Il processo di ottenimento della lista completa delle singole amenities ha presentato delle complessità in quanto oltre alle dotazioni standard gli host possono aggiungere delle amenities a loro scelta, ad esempio al posto di "coffee machine" alcuni hanno inserito "Nespresso machine", ciò ha portato ad avere delle ripetizioni da eliminare o delle amenities da sommare. Inoltre, si sono trovati dei nomi a cui mancavano le " " è altre che contenevano il simbolo " " al loro interno, a questo si aggiunge che non tutte le amenities iniziano con la lettera maiuscola o ci sono stati dei casi in cui mancava la virgola tra due nomi, ciò ha reso impossibile trovare un'occorrenza su cui basarsi per creare un algoritmo che permettesse di estrarre automaticamente tutte le amenities. Quindi arrivare ad avere un elenco più completo possibile si è partiti da liste già presenti in altri studi o trovate sul sito di Airbnb, anche se non ne esiste una ufficiale completa, quindi si è dovuto controllare che nel database ci fossero tutte quelle indicate, per poi svolgere una ricerca per lettera andando a ricercare le amenities mancanti fino ad ottenere la lista finale. Le dotazioni trovate, con il relativo numero di volte che si ripetono all'interno del campione di proprietà prese in considerazione sono:

Single amenities	Ripetizioni
Heating	9136
Essentials	9057
Wifi	8876
Dryer	8803
Hair dryer	8742
Kitchen	8508
Hangers	8050
Washer	7861
Long term stays allowed	7738
TV	7738
Iron	7561
Hot water	7400
Shampo	6943
Air conditioning	6925
Dedicated workspace	6884
Dishes and silverware	6684
Refrigerator	6520
Cooking basics	6230
Bed linens	5427
Stove	5121
Coffee maker	5112
Elevator	4748
Oven	4483
Microwave	4078
Microwave	4078
Dishwasher	3557
Extra pillows and blankets	3441
Host greets you	3203
First aid kit	3106
Patio or balcony	2906
Private entrance	2783
Free street parking	2688
Fire extinguisher	2423
Smoke alarm	2304
Carbon monoxide alarm	2146
Luggage dropoff allowed	1910
Shower gel	1897
Single level home	1731
Bidet	1598
Breakfast	1585
Freezer	1583
Body soap	1503
Clothing storage	1442
Dining table	1374
Cleaning products	1367
Crib	1330
Room-darkening shades	1319
Backyard	1270
Cable TV	1192
TV with standard cable	1144
Paid parking on premises	1133
Baking sheet	1120
Drying rack for clothing	1060
Hot water kettle	1035
Lock on bedroom door	1018
Wine glasses	948
Bathtub	875
Free parking on premises	806
Ethernet connection	803
High chair	657
Lockbox	650
Nespresso machine	621
Clothing storage: closet	615
Pack \u2019n play/Travel crib	580
Laundromat nearby	541
Conditioner	537

Single amenities	Ripetizioni
Toaster	473
Children\u2019s books and toys	454
Central heating	449
Portable fans	424
Building staff	422
Private patio or balcony	416
Cleaning before checkout	395
Security cameras on property	356
Free washer	339
Mosquito net	335
Pocket wifi	326
Free washer \u2013 In unit	313
Keypad	313
Outdoor furniture	295
HDTV	285
Babysitter recommendations	263
Children\u2019s dinnerware	260
Induction stove	246
Indoor fireplace	236
Safe	223
Hot tub	219
Central air conditioning	209
Outdoor dining area	205
Window guards	202
Clothing storage: dresser and closet	174
EV charger	162
Changing table	138
Ceiling fan	132
Baby bath	131
Dedicated workspace: desk	131
Pour-over coffee	131
Gym	119
Stainless steel gas stove	116
Outlet covers	104
Clothing storage: wardrobe	103
Board games	94
BBQ grill	89
Table corner guards	69
Baby safety gates	66
Game console	64
Bikes	61
Radiant heating	60
Wifi \u2013 1000 Mbps	56
Private fenced garden or backyard	54
32" HDTV	53
Barbecue utensils	47
Fireplace guards	47
Baby monitor	43
Bread maker	36
Rice maker	36
Wifi \u2013 300 Mbps	36
Record player	28
Private hot tub	22
Keurig coffee machine	17
32" HDTV with Netflix	15
Dedicated workspace: table and desk	15
Dryer \u2013 In unit	15
Sauna	7
Dedicated workspace: office chair and table	6
Private pool	4
40" HDTV with Netflix, standard cable, Amazon Prime Video	1
Amazon Echo sound system	1
Induction stove	1
Clothing storage: wardrobe, dresser, walk-in closet, and closet	1
top quality stainless steel electric stove	1
Variable body soap	1

Tab.2 Lista delle amenities presenti nel database di riferimento, con il numero di proprietà in cui sono presenti

Come si può osservare sono state trovate in totale 133 amenities, che variano da servizi offerti dal host, come “Long term stay allowed”, a utensili che tipicamente si trovano in casa, ad esempio “Washer”, a beni considerati comunemente di lusso, come la “Sauna”, a servizi che non sono forniti dall’host ma che sono dovuti alla posizione della struttura interessata, quale “Free parking on premises”.

Alcune dotazioni al posto di essere scritte con il loro nome normale si presentano con una specie di codice, ciò non crea problemi in quanto è comunque facilmente comprensibile a quali forniture ci si riferisca, come ad esempio “Pack \u2019n play/Travel crib” indica chiaramente le culle per bambini.

Mediamente un’amenity è presente in 1851,69 strutture ricettive, ma è facilmente osservabile che questo dato è estremamente variabile, infatti la deviazione standard è uguale a 2662,88.

Un altro aspetto, già segnalato in precedenza, è quello della ripetizione di alcune dotazioni che però vengono presentate con nomi diversi, queste sono:

- Tv
- Dedicated workspace
- Coffee maker
- Dryer
- Cloting storage
- Patio or balcony
- Clothing storage
- Wifi,
- Stainless steel gas stove
- Washer
- Stove
- Hot hub

In alcuni casi è stato necessario eliminare le ripetizioni, come nel caso di “TV” e “32/” HDTV with Netflix”, e in altri casi si sono dovute sommare il numero di rilevazioni, ad esempio per “Washer” e “Free Washer”.

La lista di amenities dopo l’eliminazione dei doppioni risulta essere:

Singole amenities	Ripetizioni
Heating	9136
Essentials	9057
Wifi	8920
Dryer	8803
Hair dryer	8742
Kitchen	8508
Washer	8476
Hangers	8050
Long term stays allowed	7738
TV	7738
Iron	7561
Hot water	7400
Shampo	6943
Air conditioning	6925
Dedicated workspace	6884
Dishes and silverware	6684
Refrigerator	6520
Cooking basics	6230
Stove	5477
Bed linens	5427
Coffee maker	5112
Elevator	4748
Oven	4483
Microwave	4078
Dishwasher	3557
Extra pillows and blankets	3441
Host greets you	3203
First aid kit	3106
Patio or balcony	2972
Private entrance	2783
Free street parking	2688
Fire extinguisher	2423
Smoke alarm	2304
Carbon monoxide alarm	2146
Luggage dropoff allowed	1910
Shower gel	1897
Single level home	1731
Bidet	1598
Breakfast	1585
Freezer	1583
Body soap	1509
Clothing storage	1442
Dining table	1374
Cleaning products	1367
Crib	1330
Room-darkening shades	1319
Backyard	1270
Paid parking on premises	1133
Baking sheet	1120
Drying rack for clothing	1060
Hot water kettle	1035

Singole amenities	Ripetizioni
Lock on bedroom door	1018
Wine glasses	948
Bathtub	875
Free parking on premises	806
High chair	657
Lockbox	650
Pack \u2019n play/Travel crib	580
Laundromat nearby	541
Conditioner	537
Toaster	473
Children\u2019s books and toys	454
Central heating	449
Portable fans	424
Building staff	422
Cleaning before checkout	395
Security cameras on property	356
Mosquito net	335
Keypad	313
Outdoor furniture	295
Hot tub	288
Babysitter recommendations	263
Children\u2019s dinnerware	260
Indoor fireplace	236
Safe	223
Central air conditioning	209
Outdoor dining area	205
Window guards	202
Mini fridge	172
EV charger	162
Changing table	138
Ceiling fan	132
Baby bath	131
Gym	119
Outlet covers	104
Board games	94
BBQ grill	89
Table corner guards	69
Baby safety gates	66
Game console	64
Bikes	61
Radiant heating	60
Private fenced garden or backyard	54
Barbecue utensils	47
Fireplace guards	47
Baby monitor	43
Bread maker	36
Rice maker	36
Record player	28
Sauna	7
Private pool	4
Amazon Echo sound system	1

Tab. 3 Lista amenities senza ripetizioni con relativo numero di volte in cui queste si presentano all'interno delle strutture ricettive

Si è passati ad avere 102 amenities, quindi 29 di quelle trovate precedentemente erano fittizie. Dopo l'eliminazione delle ripetizioni si osserva che mediamente un'amenity è presente in 2320,6 proprietà con una deviazione standard di 2869,9.

Alcune delle dotazioni inserite dagli host nei loro annunci risultano essere totalmente irrilevanti per i potenziali clienti e probabilmente molti li segnalano solo per far sembrare l'offerta da loro presentata più completa, in quanto il numero di amenities offerte è visibile nell'annuncio; ad esempio nessun turista interessa la presenza del bagno schiuma per decidere dove pernottare ma questo è segnalato in oltre 1500 strutture. Ciò rende la variabile "Numero di amenities" sporca, come si vedrà si è deciso di tenerla comunque di tenerla in considerazione.

4.3 Variabili di controllo

La prima variabile che è stata scelta è il prezzo per notte, questo valore viene scelto dall'host in base a quello che ritiene sia il valore della struttura che ha deciso di mettere in listing, si è optato per usarla come unica variabile dipendente. Il motivo per cui per cui non è stato considerato il rating come variabile dipendente è dovuto al fatto che in Airbnb mediamente i punteggi dei clienti risultano essere tutti molto alti e quindi omogenei, ad esempio nel campione preso in considerazione su 9350 strutture 7577 hanno un voto medio maggiore o uguale a 4,5, ciò rende molto difficile trovare delle differenze tra le proprietà; inoltre, per alcune questo dato non viene fornito, quindi si sarebbe dovuto ridurre ulteriormente il database effettivamente usato.

Le variabili indipendenti prese in considerazione, altre alla amenities, che svolgono il ruolo di controllo e servono a ridurre la variabilità nel modello di regressione sono divisibili in quattro categorie, più una che verrà usata per sdoppiarlo:

- Soddisfazione dei clienti
- Posizione geografica
- Conformazione della proprietà
- Reputazione dell'host
- Tipologia di stanza

4.3.1 Soddisfazione clienti

Per tenere in conto la soddisfazione dei clienti si è deciso di usare due variabili: la “review_scores_value” e il “reviews_per_month”.

La prima rappresenta il punteggio dato dai clienti per ciò che riguarda il rapporto qualità/prezzo, questo voto deve essere assegnato entro i 14 giorni successivi al loro pernottamento. I voti che vengono assegnati sulla piattaforma Airbnb, come già sottolineato in altri studi, hanno valori molto alti; per questa specifica categoria in questo database si ha una media di 4,6 su 5.

La seconda variabile mostra quante recensioni al mese mediamente vengono date, fa capire ogni quanto viene usata da dei clienti la struttura. Si suppone che strutture vendute a prezzi più alti abbiano meno clienti ma ricavi per ogni pernottamento più alti. Si vuole verificare la veridicità di questa ipotesi.

4.3.2 Posizione geografica

In che zona della città si trova una proprietà incide sull’esperienza che si può avere, è facile ipotizzare che le zone con più attrazioni e migliori trasporti tendano a far aumentare i prezzi, come avviene anche per i normali affitti.

Per tenere sotto controllo la posizione vengono fornite diverse variabili, ma quella che si è deciso di tenere in considerazione è “neighbourhood_cleansed”, in cui viene presentato il quartiere in cui è situata la struttura di interesse.

In totale la città di Milano viene suddivisa in 82 quartieri. Questi quartieri verranno rappresentati attraverso una variabile chiamata “Effetti fissi quartiere” che li rappresenterà in contemporanea tutti quanti.

4.3.3 Conformazione della proprietà

Per ottenere un modello con capacità previsionali si devono considerare gli spazi e la conformazione che sono disponibili all’interno della struttura, per far ciò si sono presi in considerazione:

- “accomodities”, cioè quante persone al massimo possono pernottare contemporaneamente nella proprietà
- “bedroom”, numero di stanze disponibili
- “bath”, numero di bagni
- “Number of amenities”, quante amenities sono segnalate nell’annuncio dall’host

Si è deciso di prendere in considerazione il numero di amenities perché essendo questo valore è presente nell’annuncio c’è la possibilità che influenzi la decisione di un cliente su dove alloggiare. Inoltre, si ipotizza che maggiore è il numero di dotazioni e servizi offerti più il prezzo sarà alto anche se come visto in precedenza questa potrebbe essere una variabile sporca.

4.3.4 Reputazione dell’host

L’affidabilità degli affittuari può essere di interesse per i clienti poiché non andando in strutture ricettive controllate da non professionisti il rischio di incorrere in truffe o semplicemente esperienze negative è maggiore, per questo Airbnb concede agli host che vengono considerati più meritevoli un badge di garanzia definito “Superhost”, questa sarà la variabile presa in considerazione per tenere sotto controllo gli effetti della reputazione sul prezzo.

4.3.5 Tipologie di stanza

Ci sono tre tipologie di stanza possibili:

- “Entirehomeapt”, il cliente prenota l’intero appartamento; nel database il numero di proprietà che offrono questa tipologia di affitto sono 7579, cioè l’81,06% del totale
- “Privateroom”, il potenziale cliente può affittare una stanza all’interno della struttura, questo implica che dovrà condividere i luoghi comuni e le relative amenities con degli sconosciuti; sono presenti 1689 proprietà che prenotabili con questa modalità, pari al 18,06%

- “sharedroom”, in questo caso viene condivisa anche la stanza, questa scelta viene solitamente effettuata da chi è interessato a risparmiare o ad avere interazioni con altre persone; sono presenti solamente 82 proprietà in cui l’host ha deciso di adottare questo metodo di affitto, pari al 0,88%
- “Hotelroom”, c’è un numero estremamente limitato di hotel che ha deciso di inserire le proprie camere all’interno della piattaforma, per un totale di 12 stanze. Si è deciso di escludere subito questi in quanto non rientrano nelle casistiche che si è intenzionati ad analizzare, inoltre l’inserimento di annunci di hotel è un fatto estremamente raro e costituisce un’eccezione.

Questa non è semplicemente una variabile, infatti si creeranno dei modelli separati per le prime due tipologie, perché l’uso delle amenities cambia fortemente in base al livello di condivisione scelto. Si è deciso di non prendere in considerazione “sharedroom” in quanto il numero di osservazioni è troppo basso per ottenere un modello significativo.

4.4 Amenities

Le variabili rappresentanti le amenities scelte per svolgere le analisi sono di due tipologie: singole o raggruppate, in questo secondo caso vengono unite più amenities per creare un’unica variabile che rappresenti una categoria.

4.4.1 Singole amenities

Sono stati presi in considerazione tre singoli servizi: “TV”, “WiFi” e “HostGreetYou”; quest’ultimo si traduce come l’ospitante ti accoglie. Le prime due sono dotazioni che di per sé non offrono un grande valore aggiunto ma i clienti si aspettano che ci siano, la loro mancanza potrebbe quindi costituire un fastidio per molti. Si è deciso di considerare anche una variabile che tiene in considerazione il rapporto diretto con l’host perché molti degli studi finora effettuati sono giunti alla conclusione che la relazione tra ospitante e ospitato e tra i fattori che incidono positivamente sulla soddisfazione dei clienti, si vuole quindi verificare se ciò è vero e se questo ha un effetto anche sul prezzo.

Tutte queste vengono presentate come variabili dummy, quindi possono presentare solo due valori: 0 e 1, con il valore 1 che viene dato quando si verifica la condizione

scelta. La presentazione delle variabili in questa forma serve a catturare l'effetto sul modello della condizione a cui si riferisce la variabile stessa. Queste si presentano con le seguenti numerosità (quante volte la dummy risulta uguale a 1):

- TV 7569 volte
- Wifi 8719 volte
- HostGreetYou 3172 volte

4.4.2 Amenities raggruppate

Sono stati creati tre raggruppamenti che potrebbero interessare diverse tipologie di clienti offrendo un grande valore aggiunto per questi, si presuppone quindi che siano disposti a pagare cifre più alte solo grazie alla presenza di queste dotazioni.

Si è creata una variabile "Luxory" a cui viene assegnato il valore 1 a tutte le proprietà che hanno una o più di queste amenities: sauna, piscina, palestra e vasca idromassaggio, che nella matrice di partenza vengono chiamate:

- "Sauna"
- "Private pool"
- "Gym"
- "Hot tub"

Tutte le dotazioni citate si trovano generalmente in appartamenti estremamente lussuosi ed è quindi facile presupporre che la loro presenza porti ad una maggiorazione del prezzo. In totale sono presenti nel database 288 strutture che rientrano in questa categoria, un numero molto basso rispetto al totale com'era prevedibile in quanto si tratta di amenities costose da installare per l'host e quindi estremamente rare.

Si è deciso di osservare anche che impatto potesse avere la possibilità per i clienti di sfruttare degli spazi all'aria aperta, si è quindi creata una variabile "Outdoor" composta da:

- "Outdoor dining area"
- "Private fenced garden or backyard"
- "BBQ grill"

Anche in questo caso, come per “Lusso” la categoria creata è una variabile binaria 0-1 a cui viene assegnato il valore 1 a tutte le proprietà che presentano almeno una delle variabili precedentemente elencate. In totale sono presenti 286 strutture che rispettano i requisiti richiesti.

L’ultimo raggruppamento è stato dominato “Babies” ed è usato per capire se le persone con figli sono disposte a pagare di più per avere servizi collegati ad essi e se questo ha un impatto rilevante sul prezzo di affitto. Questa nuova variabile è composta da:

- “Crib”, indica la culla per bambini
- “High chair”, seggiolone
- “Children’s books and toys”, giochi e libri per bambini
- “Babysitter recommendations”, viene fornito un consiglio su a chi lasciare il bambino in caso si volesse passare del tempo senza lui/lei
- “Children’s dinnerware”, stoviglie per bambini
- “Baby bath”, bagnetto
- “Table corner guards”, protezioni per evitare in caso di caduta del bambino che questo si faccia male sbattendo sugli spigoli
- “Baby safety gates”, cancelli per evitare che il bambino vada in certe parti dell’appartamento come ad esempio le scale
- “Baby monitor”, per tenerlo sotto controllo quando dorme
- “Changing table”, fasciatoio

In totale sono presenti 1544 strutture che presentano almeno una di queste variabili.

5 Regressione

5.1 Introduzione

In questo capitolo verrà spiegato com'è stato costruito il modello di regressione lineare multipla utilizzando le variabili precedentemente presentate.

Il tool scelto per la creazione del modello è Stata, per agevolare il funzionamento di questo si è creato un foglio di calcolo Excel in formato “.xls” dove all'interno si sono riportate solamente le colonne utilizzate nell'analisi statistica. Partendo dalla colonna che indica l'ID della proprietà e da quella del prezzo/notte, successivamente sono state inserite le altre comprese quelle costruite dalla combinazione di più colonne precedentemente create in un altro foglio. A questo punto è bastato importare la matrice creata su Stata indicando che la prima riga corrispondeva al nome della variabile per poter iniziare la creazione del modello. Si ricorda che sono presenti sia variabili di tipo discreto continuo, che possono dunque assumere qualsiasi tipo di valore, sia variabili dummy, che quindi possono essere rappresentate solo da due valori [0;1]. L'1 indica che la proprietà presenta quella determinata caratteristica e lo 0 invece ne indica la mancanza.

Prima di presentare la regressione finale si osserveranno quattro sotto-modelli, tutti questi hanno come variabile dipendente il prezzo per notte e come variabili indipendenti una parte di quelle che verranno usate nell'ultima regressione. Questi sotto-modelli, sono a loro volta divisi in base al tipo di affitto: intero appartamento o singola stanza; si è deciso di fare questo distinguo in quanto le amenities vengono sfruttate in maniera considerevolmente diversa in base alla possibile presenza di estranei all'interno della proprietà. Infine, verrà presentato il modello finale in cui si terranno in considerazione tutte le variabili precedentemente analizzate e si potrà osservare quali risultano significative e quanto impatto hanno, come per i casi precedenti la variabile dipendente sarà il prezzo per notte e ci sarà la divisione in base alla tipologia di affitto.

Per l'analisi delle regressioni ci si è concentrati su quattro risultati visibili nell'output di Stata:

- P-Value: test di significatività del modello

- Adjusted R-squared: coefficiente di determinazione del modello, si usa la versione Adjusted per tenere in considerazione il numero di variabili considerate; può assumere valori compresi tra 0 e 1, più il risultato è vicino a quest'ultimo migliore è il modello
- T-test: significatività delle variabili, limite impostato al 5%
- Coefficienti: proporzionalità rispetto alla variabile dipendente

5.2 Correlazione tra prezzo/notte e rating e superhost

In questi sotto-modelli si sono analizzate le relazioni del prezzo con il rating e con il superhost. I primi risultati si riferiscono al caso in cui ci sia la possibilità di affittare l'intera struttura, poi si passa al caso di affitto delle singole stanze, questo schema si ripeterà anche nei prossimi paragrafi.

```
. regress pricenight review_scores_value reviews_per_month superhost if entirehomeapt == 1
```

Source	SS	df	MS	Number of obs	=	7,503
Model	1184458.7	3	394819.567	F(3, 7499)	=	25.73
Residual	115059823	7,499	15343.3555	Prob > F	=	0.0000
				R-squared	=	0.0102
				Adj R-squared	=	0.0098
Total	116244282	7,502	15495.1055	Root MSE	=	123.87

pricenight	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
review_scores_value	-7.399029	3.083158	-2.40	0.016	-13.44288	-1.355176
reviews_per_month	-5.659665	.90063	-6.28	0.000	-7.425152	-3.894178
superhost	-11.91239	3.332176	-3.57	0.000	-18.44439	-5.380391
_cons	155.6155	14.04444	11.08	0.000	128.0845	183.1466

Figura 1: Modello Prezzo/Notte e Rating Value, Voti al mese e Superhost con entirehomeapt==1

```
. regress pricenight review_scores_value reviews_per_month superhost if privateroom == 1
```

Source	SS	df	MS	Number of obs	=	1,658
Model	174339.38	3	58113.1267	F(3, 1654)	=	16.65
Residual	5771203.8	1,654	3489.24051	Prob > F	=	0.0000
				R-squared	=	0.0293
				Adj R-squared	=	0.0276
Total	5945543.18	1,657	3588.1371	Root MSE	=	59.07

pricenight	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
review_scores_value	-13.95953	3.505812	-3.98	0.000	-20.83583	-7.083235
reviews_per_month	-3.731949	.9247915	-4.04	0.000	-5.545834	-1.918063
superhost	-8.591889	3.422623	-2.51	0.012	-15.30502	-1.878758
_cons	136.2047	16.33807	8.34	0.000	104.1593	168.2502

Figura 2: Modello Prezzo/Notte e Rating Value, Voti al mese e Superhost con entirehomeapt==1

La prima regressione presenta 7.503 osservazioni, questo numero indica quanti appartamenti rispettano la condizione “if entirehomeapt == 1” e hanno un valore per tutte le variabili, infatti Stata automaticamente esclude le proprietà che non presentano il dato per un determinato fattore. Per la condizione “if privateroom ==1” sono invece presenti 1658 osservazioni.

Per entrambe le regressioni i risultati sono simili, il P-value è 0,00% quindi siamo in presenza di modelli significativi; questo dato non è mai uguale esattamente a 0 ma servirebbe un numero maggiore di cifre decimali per notarlo. L’Adjusted R-squared è basso in entrambi i casi quindi le variabili sono poco rappresentative dei due modelli, prevedibile visto che sono stati tenuti in considerazione pochi fattori.

Tutte le variabili superano il t-test al 5% per ogni tipologia di affitto e sono quindi significative, se si considerasse il t-test all’1% il “review_score_value” nel caso di “entirehomeapt” == 1 e “superhost” con “privateroom == 1” non passerebbero il test. I coefficienti delle variabili sono tutti negativi in ogni caso.

Il “review_score_value” si ricorda essere il rapporto qualità/prezzo, la sua negatività indica che secondo i clienti all’aumentare del prezzo la qualità non aumenta corrispondentemente, probabilmente c’è un miglioramento ma questo è inferiore alle aspettative. Anche il “reviews_per_month” ha un coefficiente minore di 0, solitamente le locations con prezzi più alti hanno meno clienti interessati e puntano a fare grandi margini di guadagno su piccoli volumi mentre le strutture con prezzi bassi hanno più clienti e quindi vengono affittati più spesso ma il guadagno per ogni pernottamento è significativamente inferiore. L’unico coefficiente inaspettato è il superhost, infatti con una migliore fama si dovrebbero avere più clienti interessati alla proprietà e all’aumento della domanda dovrebbe conseguire un aumento del prezzo, come appreso anche in altri studi che si sono occupati di questa relazione. Analizzando meglio come viene dato questo badge di merito si può notare che molti dei requisiti richiesti, quali il numero di soggiorni minimi per essere presi in considerazione, il tasso di risposta ad almeno il 90% e quello di cancellazione minore o uguale all’1%, non hanno una correlazione con l’esperienza dei clienti, l’unico requisito utile da questo punto di vista è quello di aver mantenuto un rating totale di almeno 4,8 nell’ultimo anno. Quest’ultimo requisito però

contiene un errore endogeno in quanto i voti medi dati dai clienti alle strutture di Airbnb a Milano è 4,7, perciò praticamente identico a quello richiesto per diventare superhost, rendendo quest'ultimo troppo facile da acquisire e poco significativo.

5.3 Correlazione tra prezzo/notte e caratteristiche della proprietà

```
. regress pricenight accommodates bedrooms bath numberofamenities if entirehomeapt == 1
```

Source	SS	df	MS	Number of obs	=	6,856
Model	13857105.5	4	3464276.38	F(4, 6851)	=	240.97
Residual	98493434	6,851	14376.5047	Prob > F	=	0.0000
				R-squared	=	0.1233
				Adj R-squared	=	0.1228
Total	112350540	6,855	16389.5754	Root MSE	=	119.9

pricenight	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
accommodates	5.116027	1.336477	3.83	0.000	2.496118	7.735936
bedrooms	20.2897	3.773255	5.38	0.000	12.89295	27.68645
bath	69.33819	4.521092	15.34	0.000	60.47544	78.20093
numberofamenities	-.8578057	.1456761	-5.89	0.000	-1.143376	-.5722352
_cons	10.39314	5.486044	1.89	0.058	-.3612062	21.14749

```
. regress pricenight accommodates bedrooms bath numberofamenities if privateroom == 1
```

Source	SS	df	MS	Number of obs	=	1,658
Model	282496.093	4	70624.0233	F(4, 1653)	=	20.24
Residual	5767191.45	1,653	3488.92405	Prob > F	=	0.0000
				R-squared	=	0.0467
				Adj R-squared	=	0.0444
Total	6049687.54	1,657	3650.98826	Root MSE	=	59.067

pricenight	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
accommodates	11.75406	1.960169	6.00	0.000	7.909382	15.59873
bedrooms	14.80459	6.381213	2.32	0.020	2.288474	27.3207
bath	-8.569239	4.421157	-1.94	0.053	-17.2409	.1024193
numberofamenities	-.6697246	.1458245	-4.59	0.000	-.9557448	-.3837045
_cons	51.20311	7.044491	7.27	0.000	37.38604	65.02017

Figura 3: Modelli prezzo/notte e caratteristiche proprietà per intero appartamento o singola stanza prenotati

Questi sotto-modelli riguardano la correlazione tra il prezzo per notte e com'è conformata ogni proprietà, prendendo in considerazione il numero di persone ospitabili, di stanze, di bagni e di amenities. Si è deciso di non usare tra le possibili variabili il numero di letti in quanto risulta interdipendente con il numero di stanze e quante persone sono ospitabili.

Per caso in cui al cliente venga offerta la possibilità di affittare l'intero appartamento, "entirehomeapt == 1", si hanno 6.856 rilevazioni in cui tutte le variabili considerate sono

presenti. Il modello risulta essere fortemente significativo, ciò viene mostrato dal P-Value pari a 0,00%. Adjusted R-squared all'12,28% indica che il modello è abbastanza ben rappresentato dalle variabili, soprattutto considerando che se ne sono tenute in considerazione solo quattro.

Tutte le variabili superano il t-test al 5%, in questo caso specifico addirittura all'1%, quindi sono significative. I posti disponibili, il numero di bagni e di letti hanno tutti coefficienti positivi, quindi al loro aumentare aumenta anche il prezzo/notte. Inaspettatamente invece il numero di amenities offerte dall'host risulta avere un coefficiente leggermente negativo, perciò è inversamente proporzionale al prezzo; il risultato potrebbe avere due possibili cause: molti host inseriscono una serie di dotazioni essenzialmente inutili per aumentare questo valore visibile nell'annuncio con l'obiettivo di renderlo più interessante e quindi ci si trova davanti ad una variabile sporca, si è anche osservato che sono presenti come possibili rental delle sale per feste che vengono affittate tendenzialmente per una sola notte che hanno pochissime amenities, in quanto non necessarie, ma prezzi molto alti.

Per le strutture in cui è concesso di affittare solo una stanza, "privateroom" == 1, sono presenti 1658 osservazioni. In questo caso si ha un modello significativo, il P-value è pari a 0,00%, ma poco rappresentato dalle variabili prese in considerazione, con un Adjusted R-Squared di 4,44%. Contrariamente al modello precedente la variabile "bath" non supera il t-test e presenta anche un coefficiente negativo, probabilmente essendo i bagni condivisi con degli sconosciuti risultano impattare meno sul prezzo rispetto agli spazi privati.

5.4 Correlazione tra prezzo/notte e posizione geografica

Di seguito sono riportati i modelli statistici dove viene messa in relazione la posizione geografica, definita per ogni proprietà da uno degli 82 possibili quartieri in cui viene divisa la mappa di Milano, con il prezzo/notte. Come per gli altri paragrafi prima viene esposto il caso in cui il guest possa affittare l'intera proprietà e poi quello in cui può affittare le singole stanze e condividere le zone comuni della struttura con sconosciuti.

Source	SS	df	MS	Number of obs	=	7,579
Model	7000989.71	79	88620.1229	F(79, 7499)	=	6.04
Residual	110047475	7,499	14674.9533	Prob > F	=	0.0000
				R-squared	=	0.0598
				Adj R-squared	=	0.0499
Total	117048464	7,578	15445.8253	Root MSE	=	121.14

pricenight	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
ticinese	48.72671	121.3282	0.40	0.688	-189.1105	286.564
bovisa	33.94545	122.2365	0.28	0.781	-205.6724	273.5633
sarpi	67.63342	121.3034	0.56	0.577	-170.1552	305.422
dergano	16.77027	121.956	0.14	0.891	-222.2977	255.8382
ripamonti	22.56098	122.6086	0.18	0.854	-217.7863	262.9083
loredo	33.41438	121.3475	0.28	0.783	-204.4607	271.2895
cittastudi	38.54545	121.5068	0.32	0.751	-199.6419	276.7328
parcolambrocimiano	17.97917	122.3956	0.15	0.883	-221.9505	257.9088
lodicorvetto	23.38261	121.6658	0.19	0.848	-215.1164	261.8816
padova	17.72308	121.6053	0.15	0.884	-220.6573	256.1035
farini	43.18421	122.7238	0.35	0.725	-197.3889	283.7573
ghisolfa	32.57843	121.7326	0.27	0.789	-206.0516	271.2085
centrale	68.82216	121.3167	0.57	0.571	-168.9926	306.6369
bruzzano	12.8	125.1131	0.10	0.919	-232.4567	258.0567
bandenere	26.58716	121.6946	0.22	0.827	-211.9685	265.1428
mecenate	23.83333	123.1427	0.19	0.847	-217.5608	265.2275
cantalupa	-9.17e-10	171.3181	-0.00	1.000	-335.8316	335.8316
baggio	3.2	125.1131	0.03	0.980	-242.0567	248.4567
vialemonza	15.4186	121.8425	0.13	0.899	-223.4268	254.264
scaloromana	40.24528	122.2777	0.33	0.742	-199.4533	279.9439
stadera	12.30952	121.8592	0.10	0.920	-226.5686	251.1876
tibaldi	27.95294	121.8507	0.23	0.819	-210.9086	266.8145
buenosairesvenezia	49.35201	121.2463	0.41	0.684	-188.3246	287.0287
portaromana	41.19186	121.4919	0.34	0.735	-196.9663	279.35
deangelimonterosa	50.61429	121.5721	0.42	0.677	-187.7011	288.9297
lambrate	41.6129	123.0786	0.34	0.735	-199.6556	282.8814
corsica	26.01351	121.956	0.21	0.831	-213.0544	265.0814
isola	53.59783	121.3595	0.44	0.659	-184.3008	291.4964
gratosoglioticinello	5.571429	129.5043	0.04	0.966	-248.2934	259.4363
navigli	55.63587	121.3047	0.46	0.647	-182.1554	293.4271
scristoforo	31.79787	121.7829	0.26	0.794	-206.9307	270.5265
xxiimarzo	56.06699	121.4297	0.46	0.644	-181.9692	294.1032
gallaratese	30.90625	123.0185	0.25	0.802	-210.2445	272.057
comasina	6.6	132.7025	0.05	0.960	-253.534	266.734
forzearmate	10.76923	125.7131	0.09	0.932	-235.6638	257.2023
selinunte	36.77586	122.1801	0.30	0.763	-202.7313	276.2831
ssiro	66.74468	122.4222	0.55	0.586	-173.2371	306.7264
umbriamolise	21.40741	121.6998	0.18	0.860	-217.1583	259.9731
adriano	55.5	124.8685	0.44	0.657	-189.2772	300.2772
affori	28.77778	122.2567	0.24	0.814	-210.8797	268.4353
tortona	57.77193	121.4939	0.48	0.634	-180.3902	295.9341
vigentina	46.17742	121.6277	0.38	0.704	-192.247	284.6018
triulzosuperiore	-6.285714	129.5043	-0.05	0.961	-260.1505	247.5791
barona	11.4	127.0529	0.09	0.929	-237.6594	260.4594
villapizzone	24.1405	121.6398	0.20	0.843	-214.3076	262.5886

exommorivione	46.75	123.0185	0.38	0.704	-194.4007	287.9007
greco	7.383333	122.1456	0.06	0.952	-232.0562	246.8229
maciachinimaggiolina	55.15686	121.7326	0.45	0.650	-183.4732	293.7869
giambellino	51.66972	121.6946	0.42	0.671	-186.8859	290.2253
washington	44.44	121.5434	0.37	0.715	-193.819	282.699
maggioremusocco	14.46667	123.1427	0.12	0.906	-226.9275	255.8608
rogoredo	26.19048	123.991	0.21	0.833	-216.8666	269.2475
parcoforlaniniortica	56.44444	127.693	0.44	0.658	-193.8696	306.7585
qt8	24.27273	126.5268	0.19	0.848	-223.7553	272.3008
bicocca	11.91667	126.0867	0.09	0.925	-235.2487	259.082
guastalla	94.39548	121.4819	0.78	0.437	-143.7432	332.5342
niguardacagrande	4.64	122.3456	0.04	0.970	-235.1917	244.4717
ronchettosulnaviglio	23.72093	122.5407	0.19	0.847	-216.4933	263.9351
trenno	-4.333333	139.8807	-0.03	0.975	-278.5387	269.872
quartocagnino	22.875	128.4886	0.18	0.859	-228.9987	274.7487
quintoromano	-6.8	132.7025	-0.05	0.959	-266.934	253.334
garibaldirepubblica	86.07759	121.6613	0.71	0.479	-152.4126	324.5678
parconord	0	(omitted)				
lorenteggio	14.16	123.5393	0.11	0.909	-228.0116	256.3316
duomo	123.2598	121.243	1.02	0.309	-114.4105	360.9301
magentasvittore	91.12644	121.4878	0.75	0.453	-147.0238	329.2766
portello	43.17073	122.6086	0.35	0.725	-197.1766	283.518
bovisasca	0	(omitted)				
parcomonluepontelambro	7.7	127.0529	0.06	0.952	-241.3594	256.7594
ortomercato	6.133333	125.1131	0.05	0.961	-239.1234	251.3901
brera	101.6701	121.2974	0.84	0.402	-136.1069	339.4471
PARCOBOSCO	-10.5	148.3659	-0.07	0.944	-301.3387	280.3387
pagano	82.83077	122.0685	0.68	0.497	-156.4577	322.1193
quartooggio	10.0625	124.8685	0.08	0.936	-234.7147	254.8397
figino	-11.5	148.3659	-0.08	0.938	-302.3387	279.3387
parcosempione	95.8	125.1131	0.77	0.444	-149.4567	341.0567
tretorri	137.2	132.7025	1.03	0.301	-122.934	397.334
parcodelleabbazie	-9.17e-10	171.3181	-0.00	1.000	-335.8316	335.8316
parcodeinavigli	-3.5	148.3659	-0.02	0.981	-294.3387	287.3387
giardiniportavenezia	56.11111	127.693	0.44	0.660	-194.203	306.4252
muggiano	0	(omitted)				
chiaravalle	17.5	148.3659	0.12	0.906	-273.3387	308.3387
_cons	56	121.1402	0.46	0.644	-181.4688	293.4688

Figura 4: Modello di correlazione tra prezzo per notte e posizione geografica in caso di affitto di un'intera struttura

Con “entirehomeapt==1” sono presenti 7.579 osservazioni che coinvolgono tutti i quartieri tranne tre, che risultano essere collineari (presentano una forte dipendenza lineare tra loro): Parco Nord, Bovisasca, Parco delle Abbazie. Il modello è significativo, avendo ottenuto lo 0,00% come P-Value, ma il prezzo per notte è spiegato solo in basse percentuale dai quartieri essendo l’Adjusted R-squared di 4,49%, va però tenuto in considerazione che non sono state considerate altre variabili. Nessun quartiere supera il t-test perché se presi singolarmente hanno una significatività irrisoria rispetto al modello totale. Ci sono quartieri che hanno coefficienti positivi, tra questi spiccano Duomo e Brera che sono centrali nella mappa di Milano e con molte attrazioni nei

d'intorni e quindi i clienti sono disposti a spendere di più per stare in quelle zone, e coefficienti negativi come per Figino e Parco Bosco che sono zone che si trovano alla periferia della città e non ben collegate al centro.

Source	SS	df	MS	Number of obs	=	1,689
Model	1239254.84	78	15887.8826	F(78, 1610)	=	5.09
Residual	5029648.58	1,610	3124.00533	Prob > F	=	0.0000
				R-squared	=	0.1977
				Adj R-squared	=	0.1588
Total	6268903.42	1,688	3713.80535	Root MSE	=	55.893

pricenight	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
ticinese	-12.16327	56.46026	-0.22	0.829	-122.9066	98.58006
bovisa	-38.6	56.68565	-0.68	0.496	-149.7854	72.58541
sarpi	3.101695	56.36448	0.06	0.956	-107.4538	113.6572
dergano	-35.28571	57.20811	-0.62	0.537	-147.4959	76.92447
ripamonti	-18.23077	58.0027	-0.31	0.753	-131.9995	95.53797
cittastudi	-23.16949	56.36448	-0.41	0.681	-133.725	87.38597
loreto	-21.71622	56.26919	-0.39	0.700	-132.0848	88.65234
parcolambrocimiano	-25.16667	57.4244	-0.44	0.661	-137.8011	87.46776
lodiceorvetto	-31.14583	56.47202	-0.55	0.581	-141.9122	79.62056
padova	-21.23333	56.81671	-0.37	0.709	-132.6758	90.20916
farini	-28.77778	58.91619	-0.49	0.625	-144.3383	86.7827
ghisolfi	-11.40541	56.64307	-0.20	0.840	-122.5073	99.6965
centrale	26.0411	56.27433	0.46	0.644	-84.33754	136.4197
bruzzano	-39	58.62087	-0.67	0.506	-153.9812	75.98122
bandenere	-28.90323	56.78714	-0.51	0.611	-140.2877	82.48127
mecenate	-31.84615	58.0027	-0.55	0.583	-145.6149	81.92258
cantalupa	-55	68.45442	-0.80	0.422	-189.2691	79.26914
baggio	-37	58.91619	-0.63	0.530	-152.5605	78.56048
vialemonza	-30.12	56.9997	-0.53	0.597	-141.9214	81.6814
scoloromana	-34.66667	58.17507	-0.60	0.551	-148.7735	79.44016
stadera	-24.77778	56.91844	-0.44	0.663	-136.4198	86.86425
tibaldi	-17.65385	56.95752	-0.31	0.757	-129.3725	94.06483
buenosairesvenezia	-5.533784	56.08131	-0.10	0.921	-115.5338	104.4663
portaromana	4.566667	56.81671	0.08	0.936	-106.8758	116.0092
deangelimonterosa	-11.36	56.9997	-0.20	0.842	-123.1614	100.4414
lambrate	-25.90909	58.37813	-0.44	0.657	-140.4142	88.59602
corsica	2.269231	56.95752	0.04	0.968	-109.4494	113.9879
isola	-20.5814	56.53898	-0.36	0.716	-131.4791	90.31634
gratosoglioticinello	-34.77778	58.91619	-0.59	0.555	-150.3383	80.7827
navigli	-2.372549	56.43811	-0.04	0.966	-113.0724	108.3273
scristoforo	-21.30435	57.09494	-0.37	0.709	-133.2926	90.68386
xxiimarzo	.775	56.58715	0.01	0.989	-110.2172	111.7672
gallaratese	-37.57692	56.95752	-0.66	0.510	-149.2956	74.14176
forzearmate	-39.25	59.28327	-0.66	0.508	-155.5305	77.03049
comasina	110	61.2275	1.80	0.073	-10.09397	230.094
selinunte	-25.38095	57.20811	-0.44	0.657	-137.5911	86.82923
ssiro	-24.86667	57.72584	-0.43	0.667	-138.0923	88.35902
umbriamolise	-11.65	57.27308	-0.20	0.839	-123.9876	100.6876
adriano	-39.16667	60.37113	-0.65	0.517	-157.5809	79.24759
tortona	.725	56.58715	0.01	0.990	-110.2672	111.7172
affori	53.66667	60.37113	0.89	0.374	-64.74759	172.0809
vigentina	-4.846154	58.0027	-0.08	0.933	-118.6149	108.9226
triulzosuperiore	-51	79.04436	-0.65	0.519	-206.0407	104.0407
barona	-39.85714	59.75192	-0.67	0.505	-157.0569	77.34257
villapizzone	-20.7027	56.64307	-0.37	0.715	-131.8046	90.39921
exommarivione	-33	60.37113	-0.55	0.585	-151.4143	85.41425
greco	-33.4375	57.61298	-0.58	0.562	-146.4418	79.56682
maciachinimaggiolina	-11.85106	56.48428	-0.21	0.834	-122.6415	98.93937
giambellino	-17.14286	56.68565	-0.30	0.762	-128.3283	94.04256
washington	-6.137931	56.8483	-0.11	0.914	-117.6424	105.3665
maggioremusocco	-39	61.2275	-0.64	0.524	-159.094	81.09397
rogoredo	-50	79.04436	-0.63	0.527	-205.0407	105.0407
parcoforlaniniortica	-35.71429	59.75192	-0.60	0.550	-152.914	81.48543

qt8	-32.58333	58.17507	-0.56	0.575	-146.6902	81.5235
bicocca	-34.66667	60.37113	-0.57	0.566	-153.0809	83.74759
guastalla	3.111111	57.4244	0.05	0.957	-109.5233	115.7455
niguardacagrande	-40	58.0027	-0.69	0.491	-153.7687	73.76873
ronchettosulnaviglio	-30.57143	59.75192	-0.51	0.609	-147.7711	86.62829
trenno	-23.33333	64.53945	-0.36	0.718	-149.9235	103.2568
quartocagnino	-42.5	68.45442	-0.62	0.535	-176.7691	91.76914
quintoromano	0	(omitted)				
garibaldirepubblica	-20.16667	58.17507	-0.35	0.729	-134.2735	93.94016
parconord	-41.5	68.45442	-0.61	0.544	-175.7691	92.76914
lorenteggio	-28.33333	58.91619	-0.48	0.631	-143.8938	87.22715
duomo	71.1	56.35665	1.26	0.207	-39.44011	181.6401
magentasvittore	1.354839	56.78714	0.02	0.981	-110.0297	112.7393
portello	122	62.49005	1.95	0.051	-57.0397	244.5704
bovisasca	-40	79.04436	-0.51	0.613	-195.0407	115.0407
parcomonluelpontelambro	-30	68.45442	-0.44	0.661	-164.2691	104.2691
ortomercato	271.6667	64.53945	4.21	0.000	145.0765	398.2568
brera	11.03571	56.88213	0.19	0.846	-100.5351	122.6065
PARCOBOSCOINCITT	10	79.04436	0.13	0.899	-145.0407	165.0407
pagano	-5.045455	57.14898	-0.09	0.930	-117.1397	107.0488
quartooggiano	-12.2	61.2275	-0.20	0.842	-132.294	107.894
figino	0	(omitted)				
parcosempione	-11.66667	64.53945	-0.18	0.857	-138.2568	114.9235
tretorri	-12.75	62.49005	-0.20	0.838	-135.3204	109.8204
parcodelleabbazie	-25	79.04436	-0.32	0.752	-180.0407	130.0407
parcodeinavigli	0	(omitted)				
giardiniportavenezia	51	79.04436	0.65	0.519	-104.0407	206.0407
muggiano	-10	79.04436	-0.13	0.899	-165.0407	145.0407
chiaravalle	0	(omitted)				
_cons	75	55.8928	1.34	0.180	-34.6303	184.6303

Figura 5: Modello di correlazione tra prezzo per notte e posizione geografica in caso di affitto di singole stanze

Se viene concessa la possibilità di affittare una singola stanza si hanno 1689 osservazioni ed essendo il P-Value del 0,00% il modello è significativo; inoltre, le variabili rappresentano bene il modello con un Adjusted R-squared di 15,88%. Come nella situazione precedente sono presenti quartieri che presentano collinearità: Quinto Romano, Figino Parco dei Navigli e Chiaravalle.

I quartieri che presentano i peggiori coefficienti sono Cantalupa e Rogoredo, mentre nella situazione opposta si trovano Ortomercato e Portello. Anche per questa tipologia di affitto è confermato che le zone più centrali, con più attrazioni nelle vicinanze e meglio collegate tendono ad avere prezzi più alti, come previsto.

5.5 Correlazione tra prezzo/notte e amenities

```
. regress pricenight wifi tv outdoor babies luxury hostgreetyou if entirehomeapt == 1
```

Source	SS	df	MS	Number of obs	=	7,575
Model	2172806.77	6	362134.462	F(6, 7568)	=	41.45
Residual	66117690.3	7,568	8736.48127	Prob > F	=	0.0000
				R-squared	=	0.0318
				Adj R-squared	=	0.0310
Total	68290497.1	7,574	9016.43742	Root MSE	=	93.469

pricenight	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
wifi	23.47791	4.469335	5.25	0.000	14.71677	32.23905
tv	16.64758	3.365663	4.95	0.000	10.04995	23.24521
outdoor	5.052473	6.424318	0.79	0.432	-7.540973	17.64592
babies	21.49461	2.767833	7.77	0.000	16.06889	26.92033
luxury	61.2089	6.565421	9.32	0.000	48.33885	74.07894
hostgreetyou	-9.12987	2.279052	-4.01	0.000	-13.59744	-4.662296
_cons	69.94759	5.030693	13.90	0.000	60.08604	79.80915

```
. regress pricenight wifi tv outdoor babies luxury hostgreetyou if privateroom == 1
```

Source	SS	df	MS	Number of obs	=	1,689
Model	161314.949	6	26885.8248	F(6, 1682)	=	7.40
Residual	6107588.48	1,682	3631.14654	Prob > F	=	0.0000
				R-squared	=	0.0257
				Adj R-squared	=	0.0223
Total	6268903.42	1,688	3713.80535	Root MSE	=	60.259

pricenight	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
wifi	9.774586	7.219488	1.35	0.176	-4.38554	23.93471
tv	15.18716	2.981854	5.09	0.000	9.338623	21.03569
outdoor	-14.95058	7.967505	-1.88	0.061	-30.57785	6.676845
babies	12.59722	6.300828	2.00	0.046	2.389323	24.95551
luxury	9.240495	7.21043	1.28	0.200	-4.901865	23.38285
hostgreetyou	-7.54151	3.030054	-2.49	0.013	-13.48458	-1.598436
_cons	49.47412	7.276198	6.80	0.000	35.20276	63.74547

Figura 6: Correlazione tra amenities e prezzo per notte nelle due tipologie di affitto interessate

Il modello rappresenta la correlazione tra prezzo/notte e le amenities, o raggruppamenti di esse, precedentemente scelte. Ci si aspetta che tutti i fattori presi in considerazioni influiscano positivamente sul prezzo.

Per entrambe le tipologie di affitto si ha una regressione significativa con un P-Value uguale a 0,00%, ma il prezzo per notte risulta essere solo in piccola parte spiegato della amenities essendo gli Adjusted R-squared di 3,10% e 2,23%.

Ci sono 7.575 rilevazioni per strutture totalmente affittabili e in questa situazione l'unica variabile che non supera il t-test e quindi non risulta significativa è "outdoor", si ricorda

che questa è stata creata dall'unione di fattori che riguardavano la presenza di spazi all'aperto. Invece, i coefficienti sono tutti fortemente positivi tranne per `hostgreetyou`, ciò potrebbe indicare che il rapporto con l'host non crea valore aggiunto contrariamente a quanto osservato in altri studi oppure che la mancanza di variabili di controllo rende il modello ancora troppo impreciso.

Con `"privateroom == 1"` si hanno 1689 osservazioni e le variabili risultano meno significative come dimostrato dal fatto che `"wifi"`, `"outdoor"` e `"luxury"` non passano il t-test al 5% e anche `"babies"` è estremamente vicino a questa soglia. I coefficienti risultano tutti con una minore positività e nel caso di `"outdoor"` il valore diventa addirittura negativo, questo può indicare che quando le amenities vengono condivise risultano avere un impatto notevolmente inferiore.

5.6 Regressione finale

5.6.1 Introduzione

Dopo aver studiato i singoli sotto-modelli possiamo passare a creare una regressione lineare multivariata che racchiuda tutte le variabili prese in considerazione precedentemente, quindi si terrà in considerazione la conformazione della proprietà, la posizione geografica, delle variabili osservabili sull'annuncio e che dovrebbero fornire garanzie di qualità (`rating` e `superhost`) e le amenities.

I quartieri verranno presentati in un'unica variabile chiamata `"Effetti fissi di quartiere"` in quanto non risulta utile visualizzare l'impatto che questi hanno singolarmente.

5.6.2 Affitto di un'intera proprietà

Regressione lineare multivariabile con `"entirehomeapt" == 1`:

Source	SS	df	MS	Number of obs	=	6,782
Model	19135918.5	92	207999.114	F(92, 6689)	=	31.80
Residual	43755802.2	6,689	6541.45644	Prob > F	=	0.0000
				R-squared	=	0.3043
				Adj R-squared	=	0.2947
Total	62891720.6	6,781	9274.69704	Root MSE	=	80.879

pricenight	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
review_scores_value	-5.672022	2.121763	-2.67	0.008	-9.831354	-1.51269
reviews_per_month	-7.074638	.6497502	-10.89	0.000	-8.348355	-5.80092
accommodates	3.310361	.9273562	3.57	0.000	1.492448	5.128275
bedrooms	23.09359	2.590043	8.92	0.000	18.01628	28.1709
bath	54.13148	3.142686	17.22	0.000	47.97082	60.29215
numberofamenities	-.7023731	.1210365	-5.80	0.000	-.9396433	-.465103
wifi	14.30525	4.294924	3.33	0.001	5.885829	22.72467
tv	12.51323	3.165929	3.95	0.000	6.307	18.71946
outdoor	26.73207	6.032756	4.43	0.000	14.90595	38.5582
babies	8.949262	2.826191	3.17	0.002	3.409028	14.4895
luxury	31.61146	5.987735	5.28	0.000	19.87359	43.34933
hostgreetsyou	-5.900924	2.13748	-2.76	0.006	-10.09107	-1.710783
superhost	-6.495825	2.370995	-2.74	0.006	-11.14373	-1.847918
Effetti fissi di quartiere	Yes			Yes		

Figura 7: Correlazione tra tutte le variabili indipendenti e prezzo per notte nel caso di "entirehomeapt == 1"

Tenendo in considerazione tutte le variabili, il numero di proprietà a cui manca almeno una di queste aumenta e quindi le osservazioni sono in quantità leggermente inferiore rispetto ai casi precedenti, attestandosi a 6.782. Si tratta di una regressione significativa, P-Value = 0,00%, e il modello è ben spiegato dalle variabili prese in considerazione essendo l'Adjusted R-squared di 29,47%.

Tutte le variabili sono significative, non solo entro il limite deciso precedentemente, con t-test minore o uguale al 5% ma anche considerando una soglia più stringente dell'1%. Per ciò che riguarda i coefficienti nessuno ha modificato il proprio segno in questo nuovo modello; in particolare per le variabili di controllo il numero di stanze e di letti sono fortemente positivi, il numero di persone ospitabili è anch'esso positivo mentre tutte le altre sono negative. Tra queste gli unici due valori che vanno contro ciò che ci si aspettava di ottenere prima dell'inizio delle analisi sono i valori del numero di amenities e del superhost, i possibili motivi sono stati spiegati nei paragrafi precedenti.

Le amenities hanno tutte coefficienti fortemente positivi tranne una. Com'era prevedibile le dotazioni che hanno un impatto maggiore sul prezzo sono quelle che richiedono un maggiore investimento per essere implementate da parte dell'host, si tratta delle dotazioni di lusso e degli spazi aperti. Anche "babies" è direttamente proporzionale con il prezzo ma con un coefficiente più basso rispetto alle altre dotazioni,

probabilmente perché interessa solo un gruppo di clienti, ma non solo le persone con figli che affittano quelle proprietà, per tutti gli altri queste amenities risultano totalmente inutili. Come detto precedentemente il fatto che la variabile “hostgreetyou” abbia coefficiente potrebbe indicare che il rapporto con l’host non sia in realtà particolarmente significativo e che una fase di check-in più distaccata e professionale come quella degli hotel sia anche generalmente più apprezzata.

5.6.3 Affitto di una singola stanza

Regressione lineare multivariata con “privateroom” == 1:

Source	SS	df	MS	Number of obs	=	1,628
Model	1640634.04	90	18229.2672	F(90, 1537)	=	6.86
Residual	4084980.11	1,537	2657.76195	Prob > F	=	0.0000
				R-squared	=	0.2865
				Adj R-squared	=	0.2448
Total	5725614.15	1,627	3519.12364	Root MSE	=	51.553

pricenight	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
review_scores_value	-13.86658	3.277241	-4.23	0.000	-20.29491	-7.438241
reviews_per_month	-3.813581	.8596678	-4.44	0.000	-5.499827	-2.127335
accommodates	9.904182	1.862784	5.32	0.000	6.250314	13.55805
bedrooms	19.47627	6.190024	3.15	0.002	7.334481	31.61805
bath	-6.73994	4.018472	-1.68	0.094	-14.62221	1.142328
numberofamenities	-.661245	.1622258	-4.08	0.000	-.9794524	-.3430377
wifi	10.65944	6.660828	1.60	0.110	-2.405832	23.72471
tv	15.71821	2.731519	5.75	0.000	10.36032	21.07611
outdoor	3.581711	7.450689	0.48	0.631	-11.03288	18.1963
babies	14.37748	5.942223	2.42	0.016	2.721757	26.0332
luxory	2.742083	6.636613	0.41	0.680	-10.27569	15.75986
hostgreetyou	.0237565	2.82375	0.01	0.993	-5.515053	5.562566
superhost	-8.76898	3.184833	-2.75	0.006	-15.01606	-2.521902
Effetti fissi di quartiere	Yes			Yes		

Figura 7: Correlazione tra tutte le variabili indipendenti e prezzo per notte nel caso di “privateroom” == 1”

Le proprietà in cui sono presenti tutte le variabili e che offrono la possibilità di affittare singole stanze è 1.628.

La regressione, come in tutti i precedenti casi, è significativo come mostrato dal P-Value di 0,00% e le variabili spiegano efficacemente il modello essendo l’Adjusted R-squared di 24,43%.

I fattori di controllo risultano tutti significativi eccetto “bath” il cui t-test ha un valore di 9,5%, quindi maggiore della soglia del 5%; inoltre, solo il numero di persone ospitabili e il numero di letti risultano avere un impatto positivo sul prezzo. Anche la maggior parte

delle amenities risultano essere poco significative, le uniche che fanno eccezione sono “TV” e “babies”, Quest’ultima però ha un valore al t-test abbastanza alto anche se nella soglia massima scelta.

I coefficienti delle variabili di controllo sono simili alla situazione precedente con l’unica variazione dei coefficienti dei bagni che diventano negativi; invece, le amenities presentano un generale abbassamento di tutti i valori, quindi l’impatto che hanno sul prezzo continua ad essere positivo ma meno rilevante. L’unica eccezione è “hostgreetsyou” che è diventato leggermente positivo ma non passando il t-test questo cambio di segno non risulta significativo.

5.7 Conclusion

Guardando i risultati ottenuti dal modello di regressione precedentemente presentato e confrontandoli con gli studi esistenti si può dire che in linea generale si siano ottenuti risultati simili; infatti struttura, luogo geografico, amenities e rating hanno gli effetti previsti. L'unica eccezione riguarda il rapporto guest-host, si è analizzata la reputazione di quest'ultimo attraverso la variabile "superhost" e la prima interazione tra gli interessati con "hostgreetsyou" e in entrambi i casi questi sono risultati inversamente proporzionali al prezzo. Potrebbero esserci degli errori nel campione preso in considerazione, come supposto nel paragrafo 5.2 riguardo al Superhost, oppure i clienti non sono effettivamente interessati a questo tipo di rapporto umano e i motivi che spingono a scegliere questa tipologia di affitto e da ricercarsi esclusivamente nel possibile risparmio economico è nel poter alloggiare in una struttura "casalinga" o comunque diversa da quelle trovabile nei classici hotels.

L'interesse di questo studio è rivolto alle dotazioni e servizi offerti da host e si è cercato in particolare di comprendere che effetti avessero sul prezzo. Si osserva che le dotazioni hanno sempre un effetto positivo sul prezzo, anche se con notevoli differenze tra loro. Si è anche scoperto che però le amenities non hanno lo stesso effetto sui diversi tipi di affitti, infatti quando viene prenotata un'intera struttura l'importanza delle amenities risulta maggiore rispetto a quando si affitta una singola stanza. Questa differenza è probabilmente dovuta al fatto che, in caso di affitto di singole stanze, queste devono essere condivise con degli sconosciuti e quindi le possibilità di usufruirne a pieno diminuisce. Perciò, i proprietari che devono tendere a tenere maggiormente in considerazione le amenities da rendere disponibili e eventualmente investirvi sono coloro che sono intenzionati a rendere affittabile la loro proprietà a singoli gruppi di turisti.

Bibliografia

Abrate, G., Viglia, G., 2017. Personal or product reputation? Optimizing revenues in the sharing economy. *J. Travel. Res.*

Amaro, S., Andreu, L., Huang, S., 2018. Millenials' intentions to book on Airbnb. *Curr. Issues Tour.* 1–15.

Becerra, M., Santaló, J., Silva, R., 2013. Being better vs. Being different: differentiation, competition, and pricing strategies in the Spanish hotel industry. *Tour. Manag.* 34, 71–79.

Belarmino, A., Koh, Y., Whalen, E., Bowen, J., 2017. Comparing guests' key attributes of peer-to-peer accommodations and hotels: mixed-methods approach. *Current Issues in Tourism.*

Benítez-Aurioles, B., 2018. Why are flexible booking policies priced negatively? *Tour. Manag.* 67, 312–325.

Chattopadhyay, M., Kumar Mitra, S., 2019. Do Airbnb host listing attributes influence room pricing homogenously? *Int. J. Hosp. Manag.* 81, 54-64.

Chen, C.F., Rothschild, R., 2010. An application of hedonic pricing analysis to the case of hotel rooms in Taipei. *Tour. Econ.* 16 (3), 685–694.

Ert, E., Fleischer, A., Magen, N., 2016. Trust and reputation in the sharing economy: the role of personal photos in Airbnb. *Tour. Manag.* 55, 62–73

Gibbs, C., Guttentag, D., Gretzel, U., Morton, J., Goodwill, A., 2018. Pricing in the sharing economy: a hedonic pricing model applied to Airbnb listings. *J. Travel Tour. Mark.* 35 (1), 46–56.

Gunter, U., 2018. What makes an Airbnb host a superhost? Empirical evidence from San Francisco and the Bay Area. *Tour. Manag.* 66, 26–37.

Gunter, U., Önder, I., 2017. Determinants of Airbnb demand in Vienna and their implications for the traditional accommodation industry. *Tour. Econ.*

Guttentag, D., 2015. Airbnb: disruptive innovation and the rise of an informal tourism accommodation sector. *Tour. Manag.* 66, 26–37.

Guttentag, D., Smith, S., Potwarka, L., Havitz, M., 2018. Why tourists choose airbnb: a motivation-based segmentation study. *Tour. Econ.*

Heo, C.Y., Hyun, S.S., 2015. Do luxury room amenities affect guests' willingness to pay? *Int. J. Hosp. Manag.* 46, 161–168.

Heo, C.Y., Blal, I., Choi, M., 2019. What is happening in Paris? Airbnb, hotels, and the Parisian market: a case study. *Tour. Manag.* 70, 78–88.

Horn, K., Merante, M., 2017. Is home sharing driving up rents? Evidence from Airbnb in Boston. *J. Hous. Econ.* 38, 14–24.

Kirkos, E., 2021. Airbnb listings' performance: determinants and predictive models. *European Journal of Tourism Research*

Lee, S.K., Jang, S.C., 2012. Premium or discount in hotel room rates? The dual effects of a central downtown location. *Cornell Hosp. Q.* 53 (2), 165–173.

Liang, S., Schuckert, M., Law, R., Chen, C.C., 2017. "Be a "Superhost": the importance of badge systems for peer-to-peer rental accommodations.". *Tour. Manag.* 60, 454–465.

Liang, L.J., Choi, H.C., Joppe, M., 2018. Exploring the relationship between satisfaction, trust and switching intention, repurchase intention in the context of Airbnb. *Int. J. Hosp. Manag.* 69, 41–48.

Masiero, L., Nicolau, J.L., Law, R., 2015. A demand-driven analysis of tourist accommodation price: a quantile regression of room bookings. *Int. J. Hosp. Manag.* 50 (1), 8.

Mauri, A.G., Minazzi, R., Nieto-García, M., Viglia, G., 2018. Humanize your business. The role of personal reputation in the sharing economy. *Int. J. Hosp. Manag.* 73, 36–43.

Mody, M., Suess, C., Dogru, T., 2017. Comparing apples and oranges? examining the impacts of Airbnb on hotel performance in Boston. *Boston Hospitality Rev.* 5 (2), 1–15.

Qiu, R.T., Fan, D.X., Liu, A., 2018. Exploring the Booking Determinants of the Airbnb Properties: An Example of the Listings of London. *Information and Communication Technologies in Tourism Springer, Cham*, pp. 44–51.

Saló, A., Garriga, A., Rigall-I-Torrent, R., Vila, M., Fluvà, M., 2014. Do implicit prices for hotels and second homes show differences in tourists' valuation for public attributes for each type of accommodation facility? *Int. J. Hosp. Manag.* 36, 120–129.

Schamel, G., 2012. Weekend vs. Midweek stays: modelling hotel room rates in a small market. *Int. J. Hosp. Manag.* 31 (4), 1113–1118.

Sundararajan, Arun, 2016. *The Sharing Economy: the End of Employment and the Rise of Crowd-based Capitalism.* Mit Press. Mit Press.

Thrane, C., 2007. Examining the determinants of room rates for hotels in capital cities: the Oslo experience. *J. Revenue Pricing Manage.* 5 (4), 315–323.

Tussyadiah, I.P., 2016. Factors of satisfaction and intention to use peer-to-peer accommodation. *Int. J. Hosp. Manag.* 55, 70–80.

Varma, A., Jukic, N., Pestek, A., Shultz, C.J., Nestorov, S., 2016. Airbnb: Exciting innovation or passing fad? *Tourism Manage. Perspect.* 20, 228–237.

Wang, D., Nicolau, J.L., 2017. Price determinants of sharing economy based accommodation rental: a study of listings from 33 cities on Airbnb.cOm. *Int. J. Hosp. Manag.* 62, 120–131.

Yang, Y., Mueller, N.J., Croes, R.R., 2016. Market accessibility and hotel prices in the Caribbean: the moderating effect of quality-signaling factors. *Tour. Manag.* 56, 40–51.

Yannopoulou, N., Moufahim, M., & Bian, X., 2013. User-Generated Brands and Social Media: Couchsurfing and Airbnb. *Contemporary Management Research*, Vol. 9 No. 1, pp. 85-90.

Zervas, G., Proserpio, D., Byers, J.W., 2017. The rise of the sharing economy: estimating the impact of Airbnb on the hotel industry. *J. Mark. Res.* 54 (5), 687–705.

Zhang, Z., Chen, R.J.C., Han, L.D., Yang, L., 2017. Key factors affecting the price of Airbnb listings: a geographically weighted approach. *Sustainability (Switzerland)* 9 (9).