



**Politecnico
di Torino**

Politecnico di Torino

Corso di Laurea Magistrale in Ingegneria Gestionale

Anno Accademico 2021/2022

Sessione di laurea Luglio 2022

I fattori di successo degli host di Airbnb

Relatrice

Prof.ssa Laura Abrardi

Candidato

Simona Siclari

A mia nonna Lina

INDICE

INTRODUZIONE	10
 CAPITOLO 1	12
1.1 SHARING ECONOMY	12
1.1.1 Aspetti critici della sharing economy	12
1.1.2 Tipologie di sharing economy più diffuse	14
1.2 AIRBNB	15
1.2.1 La figura del Superhost	17
1.2.2 Confronto tra Airbnb e Hotel	20
1.3 SODDISFAZIONE DEL CLIENTE	22
 CAPITOLO 2	25
2.1 ANALISI DELLA LETTERATURA	25
2.2 EVIDENZE DALLA LETTERATURA	34
 CAPITOLO 3	38
3.1 DESCRIZIONE DATABASE	38
3.2 ANALISI DELLE VARIABILI	41
3.2.1 Host_Location	41
3.2.2 Host_Is_Superhost	42
3.2.3 Host_has_profile_pic	42

3.2.4 Host_Identity_Verified	43
3.2.5 Instant_bookable	43
3.2.6 Room_Type	44
3.2.7 Host_response_time	45
3.2.8 Host_response_rate e Host_acceptance_rate	46
3.2.9 Host_listings_count	47
3.2.10 Accommodates	48
3.2.11 Bathrooms	49
3.2.12 Bedrooms e Beds	50
3.2.13 Price	51
3.2.14 Number_of_reviews	52
3.2.15 Review_scores	53
3.3 VARIABILI AGGIUNTE	56
3.3.1 Host_about(length)	56
3.3.2 Description(length).....	58
3.3.3 Distance	60
CAPITOLO 4	62
4.1 MODELLI DI REGRESSIONE LINEARE	62
4.2 MODELLI DI REGRESSIONE LINEARE SEMPLICE	65
4.2.1 Domanda e location dell'host	65
4.2.2 Domanda e lunghezza descrizione dell'host	66
4.2.3 Domanda e tempo di risposta dell'host	67

4.2.4 Domanda e tasso di risposta dell'host, domanda e tasso di accettazione dell'host	68
4.2.5 Domanda e titolo di superhost	69
4.2.6 Domanda e numero di annunci dell'host	70
4.2.7 Domanda e foto profilo	71
4.2.8 Domanda e identità verificata dell'host	72
4.2.9 Domanda e lunghezza descrizione dell'host	73
4.2.10 Domanda e distanza dell'alloggio	74
4.2.11 Domanda e tipologia dell'alloggio	75
4.2.12 Domanda e numero di persone ospitabili	76
4.2.2 Domanda e numero di bagni, domanda e numero di camere da letto, domanda e numero di letti	77
4.2.14 Domanda e prezzo.....	79
4.2.15 Domanda e rating complessivo	80
4.2.16 Domanda e prenotazione immediata	81
4.3 MODELLI DI REGRESSIONE LINEARE MULTIPLA PER TIPOLOGIA DI VARIABILI	83
4.3.1 Variabili relative all'host	83
4.3.2 Variabili relative all'alloggio	85
4.4 MODELLO FINALE	87
4.4.1 Modello con variabile dipendente "review_scores_rating"	89
CAPITOLO 5	91
CONCLUSIONI	91

BIBLIOGRAFIA	94
---------------------------	-----------

SITOGRAFIA.....	95
------------------------	-----------

RINGRAZIAMENTI	96
-----------------------------	-----------

Indice delle figure

Figura 1: Modello conferma/disconferma	23
Figura 2: Grafico location	41
Figura 3: Grafico superhost	42
Figura 4: Grafico foto profilo	42
Figura 5: Grafico profilo verificato	43
Figura 6: Grafico prenotazione	43
Figura 7: Grafico tipologia alloggio	44
Figura 8: Grafico tempi di risposta	45
Figura 9: Andamento tasso di risposta	46
Figura 10: Andamento tasso di accettazione	47
Figura 11: Andamento numero di nnunci	48
Figura 12: Istogramma numero di persone ospitabili	49
Figura 13: Istogramma numero di bagni	50
Figura 14: Istogramma numero di camere da letto	50
Figura 15: Istogramma numero di letti	51
Figura 16: Andamento del prezzo	52
Figura 17: Andamento numero di recensioni	53
Figura 18: Istogramma rating complessivo	54
Figura 19: Istogramma pulizia	54
Figura 20: Istogramma precisione dell'host	54
Figura 21: Istogramma check-in	55
Figura 22: Istogramma comunicazione	55
Figura 23: Istogramma qualità	55
Figura 24: Istogramma location	55
Figura 25: Grafico descrizione host	56
Figura 26: Andamento lunghezza descrizione host	57
Figura 27: Grafico descrione annuncio	58

Figura 28: Andamento lunghezza descrizione annuncio	59
Figura 29: Istagramma distanza	61
Figura 30: Output Stata "Domanda e Location"	65
Figura 31: Output Stata "Domanda e Descrizione host"	66
Figura 32: Output Stata "Domanda e Tempo di risposta"	67
Figura 33: Output Stata "Domanda e Tasso di risposta"	68
Figura 34: Output Stata "Domanda e Tasso d'accettazione"	68
Figura 35: Output Stata "Domanda e Superhost"	69
Figura 36: Output Stata "Domanda e Numero annunci"	70
Figura 37: Output Stata "Domanda e Foto profilo"	71
Figura 38: Output Stata "Domanda e Profilo verificato"	72
Figura 39: Output Stata "Domanda e Descrizione annuncio"	73
Figura 40: Output Stata "Domanda e Distanza alloggio"	74
Figura 41: Output Stata "Domanda e Tipologia alloggio"	75
Figura 42: Output Stata "Domanda e Capienza alloggio"	76
Figura 43: Output Stata "Domanda e Numero bagni"	77
Figura 44: Output Stata "Domanda e Numero stanze"	77
Figura 45: Output Stata "Domanda e Numero di letti"	78
Figura 46: Output Stata "Domanda e Prezzo"	79
Figura 47: Output Stata "Domanda e Rating"	80
Figura 48: Output Stata "Domanda e Diversi rating"	81
Figura 49: Output stata "Domanda e Prenotazione immediata" ..	81
Figura 50: Output Stata "Domanda e Variabili host"	83
Figura 51: Output Stata "Domanda e Variabili alloggio"	85
Figura 52: Output Stata Modello Finale	87
Figura 53: Output Stata Rating	89

Indice delle tabelle

Tabella 1: Schema riassuntivo	37
Tabella 2: Tempi di risposta	45
Tabella 3: Statistiche descrittive tasso di risposta e di accettazione	46
Tabella 4: Statistiche descrittive numero di annunci	47
Tabella 5: Statistiche descrittive numero di persone ospitabili....	48
Tabella 6: Statistiche descrittive numero di bagni	49
Tabella 7: Statistiche descrittive numero di camere da letto e letti	50
Tabella 8: Statistiche descrittive del prezzo	51
Tabella 9: Statistiche descrittive numero di recensioni	52
Tabella 10: Statistiche descrittive dei vari rating	53
Tabella 11: Statistiche descrittive lunghezza descrizione host	56
Tabella 12: Schema lunghezza descrizione host	57
Tabella 13: Statistiche descrittive lunghezza descrizione annuncio	58
Tabella 14: Schema lunghezza annuncio	59
Tabella 15: Statistiche descrittive distanza	60

INTRODUZIONE

La sharing economy, o economia della condivisione, è un modello di distribuzione di beni e servizi affermatosi negli ultimi decenni, e che può essere declinato e applicato in diversi settori economici e sociali. Si trova sempre associato al concetto di *Peer-to-Peer Economy*, in quanto modello dove individui interagiscono per comprare o vendere beni e servizi direttamente l'uno con l'altro, senza l'intermediazione di una terza parte. Vi è dunque un contatto diretto tra domanda e offerta, grazie all'utilizzo di siti web e mobile app.

La rivoluzione digitale, infatti, ha permesso una maggior diffusione di Internet e lo sviluppo di App e piattaforme digitali. Tale progresso tecnologico ha quindi reso possibile la sharing economy, permettendo una fruizione semplice ed efficace di beni e servizi condivisi digitalmente, ed è per questo che si dice che la sharing economy sia imprescindibile dalla tecnologia, in quanto vive grazie ad essa.

L'emblema di sharing economy del settore turistico è "Airbnb", un portale online, con anche App mobile associata, che mette in contatto persone in cerca di un alloggio con persone che li mettono in affitto per periodi limitati. Con Airbnb la rental economy (termine utilizzato per sottolineare il concetto di dare o prendere in affitto o a noleggio) ha avuto una rapida diffusione, cambiando il modo in cui si condividono gli spazi abitativi, introducendo indubbiamente disruption nel mondo dell'accoglienza.

Con questo lavoro di tesi si intende analizzare quali sono i fattori che determinano il successo di un host sulla piattaforma Airbnb.

Nel primo capitolo vi è una panoramica sul fenomeno della sharing economy, una sorta di introduzione sul funzionamento di Airbnb e su quelli che sono i fattori che influenzano il comportamento del consumatore.

Nel capitolo successivo sono presentati i risultati ottenuti da studi precedenti, evidenziando quali sono gli elementi che i consumatori prendono in considerazione e quali variabili influiscono nella scelta di un alloggio.

Successivamente, partendo dalle variabili presenti sui database consultabili sul portale "insideairbnb", si andranno a selezionare quelle ritenute più interessanti e se ne farà un'analisi dettagliata per cercare di capirne la distribuzione.

Infine, nell'ultimo capitolo saranno presentati i modelli di regressione, rispettivamente lineare e multipla, per capire come le varie variabili selezionate influenzano il tasso di domanda, scelta come variabile dipendente.

CAPITOLO 1

1.1 SHARING ECONOMY

La sharing economy può essere considerata come un nuovo modello economico e culturale in grado di promuovere delle forme di consumo che prediligono l'utilizzo e lo scambio di beni e servizi piuttosto che il loro acquisto; quindi, è un modello che si basa sull'accesso piuttosto che sul possesso.

Proprio perché la sharing economy ha diverse declinazioni e può assumere diversi significati in base all'ambito di applicazione (economico, turistico, sociale), è associata ad una serie di espressioni, come peer-economy o access-economy.

Per quanto riguarda, però, l'espressione generale "**sharing economy**" si può considerare una delle definizioni più diffuse data da Rachel Botsman:

"La sharing economy è un sistema economico basato sulla condivisione di beni/servizi sottoutilizzati a pagamento o in maniera gratuita, gestito direttamente dagli individui"

1.1.1 Aspetti critici della sharing economy

Rachel Botsman e Roo Rogers (2010), nel loro lavoro incentrato sull'economia condivisa, individuano quelli che sono considerati i quattro principi del meccanismo della sharing economy:

1. Essenziale è il conseguimento della **massa critica**, cioè quel livello minimo che deve essere raggiunto affinché si abbia l'innescò di un fenomeno. Per i modelli di rete come la sharing economy è un fattore importante perché dipende molto da quanto i servizi sono

diffusi e conosciuti. A tal proposito, un esempio emblematico è Airbnb che riuscì in pochissimo tempo a raggiungere questa massa critica (si parla di pochissimi users iniziali alla sua nascita, fino a mezzo milione di utilizzatori dopo soli due anni).

La massa critica ha dei risvolti anche a livello psicologico: il suo raggiungimento è un punto di svolta per il numero di utilizzatori perché si attiva un meccanismo comportamentale che porta le persone a voler imitare i propri simili e a volersi conformare ad essi.

2. Si parla poi di **capacità inutilizzata**, fulcro centrale su cui si sono sviluppate moltissime piattaforme di scambio di beni. Se, infatti, un tempo vi era la preferenza ad acquistare qualsiasi tipo di bene, la consapevolezza del loro effettivo utilizzo parziale ha portato alla nascita della sharing economy, comprendendo sia beni tangibili che intangibili (si può parlare di condivisione, infatti, anche degli spazi lavorativi, il cosiddetto coworking, durante il quale si possono condividere idee e conoscenze).
3. L'**utilizzo di beni comuni** trova tipicamente esempio nei fenomeni come carsharing, coworking e nell'affitto di alloggi. Anche quest'aspetto non è da sottovalutare, perché se vi sono degli effetti positivi in termini di risparmio, in generale, quando si parla di condivisione vi è anche il rischio di andare incontro al fenomeno del "free-riding", un comportamento opportunistico da parte di persone che tendono ad usufruire di un bene/servizio senza contribuire alla sua costituzione e al suo mantenimento.
4. Infine, un ulteriore aspetto importante da considerare è quello della **fiducia**, che è alla base di ogni relazione all'interno della sharing economy e che può essere difficile da ottenere a causa di barriere innalzate dalla distanza o dalla estraneità che vi è tra le persone.

1.1.2 Tipologie di sharing economy più diffuse

Attraverso le sue applicazioni si può comprendere maggiormente il significato di sharing economy. Tra le tipologie più diffuse ci sono:

- Car sharing
È una delle forme di sharing economy che ha maggiormente preso piede negli ultimi anni. Anche in questo caso, è lampante la dipendenza dalla tecnologia, infatti, tramite l'uso di un'applicazione specifica si può sfruttare un servizio di mobilità e quindi usufruire di viaggi condivisi, risparmiando, così, sul costo complessivo del viaggio.
- Coworking
È uno stile lavorativo innovativo che prevede la condivisione di spazi di lavoro con altre persone che mantengono però un'attività indipendente. È questa la differenza rispetto ai tipici ambienti d'ufficio: chi si trova in coworking generalmente non appartiene alla stessa organizzazione.
- Crowdfunding
Per crowdfunding si intende una sorta di micro-finanziamento collettivo che ha luogo grazie a delle piattaforme specifiche (conti online) per la raccolta di denaro destinato ad un determinato progetto.
- Home sharing
È un modello di business immobiliare riguardante il settore turistico che prevede la condivisione della propria casa con altre persone a prezzi convenienti. Il tutto funziona tramite un apposito sito nel quale vengono caricati gli annunci relativi agli alloggi affittabili per un determinato periodo di tempo.

1.2 AIRBNB

Airbnb nasce verso la fine del 2007, a San Francisco, dall'idea di due studenti neolaureati che, trovandosi in difficoltà economiche, decisero di pubblicare un annuncio mettendo in affitto dei posti letto della propria abitazione durante un importante meeting che si svolgeva in città, che, portando una grossa affluenza di viaggiatori, aveva determinato la saturazione dei posti negli hotel.

Airbnb è una piattaforma che permette di prenotare e/o mettere a disposizione alloggi in ogni parte del mondo. Chiunque può diventare un host, offrendo uno spazio non utilizzato e affittandolo per un breve periodo.

I profili personali e un sistema di valutazione e revisione forniscono informazioni sull'host e su ciò che egli offre. Allo stesso tempo, gli host possono decidere in maniera autonoma a chi affittare il proprio spazio.

Per poter utilizzare Airbnb basta iscriversi gratuitamente, creando un account e associandone un metodo di pagamento valido. Una volta creato il profilo è possibile procedere con diverse azioni in base al fatto che l'utente sia un host (lato offerta) oppure un guest (lato domanda):

- Inserire annunci

L'host può facilmente creare un annuncio andando ad inserire informazioni come il tipo di alloggio che desidera affittare e il tipo di stanza, il numero massimo di persone che si possono ospitare e la città con l'indirizzo dell'abitazione, il numero di letti e bagni disponibili, i servizi offerti (per esempio TV, Wifi, aria condizionata), ed il prezzo.

Per avere un annuncio il più completo possibile e che attiri la clientela bisogna innanzitutto fornire una descrizione convincente dello spazio che si mette a disposizione, tenere costantemente aggiornato il calendario con le date in cui c'è disponibilità e pubblicare foto (fino a un massimo di 24 immagini).

La pubblicazione dell'annuncio è gratuita, e solo nel momento in cui viene effettuata una prenotazione Airbnb trattiene una fee per l'utilizzo del servizio.

- Effettuare una prenotazione

L'ospite può procedere con la ricerca dell'alloggio di interesse inserendo nella pagina principale del portale la località di destinazione, la data di arrivo e la data di partenza e il numero di persone per cui si intende prenotare. Per facilitare la ricerca, si possono applicare dei filtri di ricerca in relazione, per esempio, al tipo di stanza (casa intera, stanza privata), al numero di letti disponibili, ai servizi offerti...

Vi è, inoltre, anche la possibilità di visualizzare i risultati di ricerca su mappa, per avere un'immediata visione del posizionamento dell'abitazione.

Per completare la prenotazione, si può contattare tramite messaggio l'host, oppure, se quest'ultimo ne ha previsto la possibilità, si può procedere direttamente dal portale e concludere subito la pratica.

- Effettuare pagamenti

Per motivi di sicurezza, Airbnb prevede un pagamento immediato: al momento stesso della conferma della prenotazione i soldi vengono prelevati e "congelati" fino al giorno successivo al check-in nell'appartamento. È un meccanismo adottato per tutelare sia l'host che l'ospite.

Nel caso in cui la prenotazione non vada a buon fine (o perché rifiutata dall'host o per disdetta da parte del guest) i soldi vengono restituiti.

- Scrivere recensioni

Al termine del soggiorno, collegandosi alla pagina dell'annuncio che è stato utilizzato per la prenotazione, è possibile lasciare una recensione in merito all'esperienza nell'alloggio in cui si è stati ospitati. Questo è un passaggio molto importante e critico per gli

host, in quanto, ovviamente, recensioni positive attireranno maggior clientela.

Nel momento in cui un utente decide di diventare un host, procede con la compilazione dell'apposito modulo dell'annuncio in relazione allo spazio che vuole affittare. Tra le informazioni da inserire citate precedentemente, come la posizione, il tipo di alloggio e di stanza, i servizi offerti... il prezzo è sicuramente quello che svolge un ruolo fondamentale, in quanto uno dei principali fattori che regola la scelta del guest e che viene utilizzato per confrontare e valutare le varie offerte disponibili.

Il prezzo può essere modificato dall'host nei diversi periodi dell'anno: per esempio, durante le festività o in corrispondenza di rilevanti eventi caratterizzati da un aumento della domanda può alzare le fasce di prezzo; viceversa, in periodi con domanda minore può offrire tariffe giornaliere minori.

Airbnb non impone alcun vincolo a riguardo permettendo all'host di scegliere in maniera autonoma le fasce di prezzo per il proprio alloggio, limitandosi a fornire un supporto nella fase iniziale suggerendo una stima di prezzo in base alle informazioni inserite nell'annuncio (le dimensioni, la posizione nella città, la distanza da servizi principali come supermercati, negozi...).

1.2.1 La figura del Superhost

Quando si naviga sul sito di Airbnb tra i tanti filtri per focalizzare la ricerca si può utilizzare quello relativo al superhost.

Il superhost è semplicemente un host che si è contraddistinto rispetto agli altri per aver offerto ai propri ospiti un'esperienza di soggiorno con elevati standard qualitativi.

Il fatto che un host abbia questo titolo aggiuntivo è un'informazione molto importante per i guest: è un'indicazione abbastanza attendibile del fatto che si rapporteranno con un proprietario serio e affidabile, che offre un servizio comunemente apprezzato.

Lo status di superhost è rappresentato sul portale tramite un'icona a forma di medaglietta, posizionata in vista accanto all'immagine del profilo, ed è legato alla persona dell'host e non ai singoli alloggi che possiede. Allo stesso tempo, il legame con questi ultimi è estremamente importante, perché un superhost trae forza proprio dagli spazi che gestisce, e sono le valutazioni ottenute su tutti i suoi alloggi a determinare questo suo riconoscimento.

È importante aggiungere che la posizione di superhost non è permanente: infatti, Airbnb quattro volte all'anno procede con una mappatura di tutti gli host presenti sulla piattaforma per verificare chi è idoneo a diventare superhost e chi, invece, avendo avuto un calo di prestazione, non ne possiede più i requisiti e quindi ne viene esonerato.

Dal sito dell'Airbnb si trovano le seguenti informazioni relative ai diversi requisiti necessari per diventare un superhost:

- Bisogna che almeno il 50% degli ospiti abbia lasciato una recensione. Queste, infatti, non sono obbligatorie, ma, nonostante ciò, la percentuale da raggiungere non rappresenta uno degli obiettivi più ardui (basti pensare al fatto che anche una semplice chiacchierata con l'ospite alla fine del suo soggiorno può spingerlo a lasciar scritto un commento sulla sua esperienza).
- È necessario avere più di 10 soggiorni all'attivo oppure di avere all'attivo nell'ultimo anno almeno 100 notti da minimo 3 guest diversi. La seconda opzione è pensata per garantire le stesse possibilità agli host che offrono soggiorni a lungo termine e che altrimenti sarebbero penalizzati. Questo requisito è una sorta di garanzia per il guest di non andare incontro a disagi dovuti alla poca esperienza del proprietario.

- L'host non deve aver annullato alcuna prenotazione. Non aver effettuato nessuna cancellazione è, infatti, indice di affidabilità e buona organizzazione.
- Il tasso di risposta deve essere almeno del 90% entro 24 ore. La tempestività della risposta è un parametro molto importante da rispettare: l'host deve garantire risposta entro un giorno ad almeno il 90% delle richieste che riceve.
- La valutazione media delle recensioni deve essere di almeno 4,8 stelle su un totale di 5. Questo è uno dei requisiti più difficili da ottenere, in quanto è legato molto alla comprensione e al significato e al peso che ogni persona dà al range di stelle che si può assegnare.

Dunque, per ottenere il riconoscimento di superhost, è chiaro che bisogna soddisfare molteplici requisiti non sempre facili da ottenere.

Nonostante ciò, ha senso tentare di ottenerli in quanto il proprietario otterrebbe di conseguenza diversi vantaggi:

- Innanzitutto, come già menzionato precedentemente, sarebbe indice di affidabilità e professionalità, cosa che porta sicuramente ad essere preferiti rispetto ad altre opzioni, che possono essere ridotte anche grazie al filtro di ricerca superhost;
- Si entrerebbe a far parte dell'elenco di host per cui Airbnb invia spesso delle mail pubblicitarie ai guest con liste di alloggi di potenziale interesse;
- Si usufruirebbe di un bonus aggiuntivo del 20% per l'invito di nuovi host. Infatti, di base Airbnb offre un premio in denaro spendibile sul proprio sito per coloro che portano nuovi host sul portale, che quindi, per un superhost, sarebbe maggiorato;
- Infine, si godrebbe di premi esclusivi in denaro spendibili sempre sul sito per i viaggi, nel caso di mantenimento del titolo di superhost per oltre un anno.

1.2.2 Confronto tra Airbnb e Hotel

Nonostante Airbnb si definisca più come un servizio complementare che sostituto, è indubbiamente un vero e proprio concorrente per gli hotel.

Vi sono delle differenze sostanziali in relazione al meccanismo dei ricavi e costi da sostenere.

Per quanto riguarda gli host di Airbnb, i loro ritorni dipendono direttamente dai costi di affitto dell'alloggio e dal suo prezzo di mercato, mentre i costi sono relativi allo spazio che mettono a disposizione e alla sua preparazione per il soggiorno; inoltre, un costo che è anche il punto di forza della sharing economy è il tempo speso con gli ospiti in tutta la durata del viaggio. Il prezzo a cui affittano le stanze è influenzato dall'offerta delle camere d'hotel e dalla volatilità della domanda.

In relazione alla gestione di un hotel bisogna considerare innanzitutto il costo iniziale che deve essere sostenuto per la realizzazione di una struttura con capacità limitata: tale costo dipende da variabili legate all'offerta abitativa, quali la posizione geografica e le regolamentazioni del luogo. Altri costi da sostenere sono sostanzialmente quelli di mantenimento della struttura, come, per esempio, i servizi di pulizia.

È da considerare, inoltre, che se una struttura alberghiera nasce con la funzione di ospitare i viaggiatori, un host di Airbnb ha generalmente un altro lavoro primario, e quindi, la messa a disposizione dei suoi alloggi inutilizzati è legata al tempo che può dedicare a questa seconda attività, senza poi una formazione sull'aspetto dell'ospitalità.

Tornando a parlare della volatilità della domanda, quando questa risulta elevata rispetto alla capacità dell'hotel tendenzialmente si assisterà ad un aumento dei prezzi delle camere e quindi l'alternativa del servizio offerto da piattaforme come Airbnb risulterà più attrattiva.

Airbnb si differenzia dai classici hotel e strutture alberghiere non soltanto per l'offerta di un'opzione di viaggio più economica e conveniente, ma anche per provvedere l'accesso a spazi e alla cultura distintivi delle loro

destinazioni. Infatti, si propone di far sentire i viaggiatori a casa ovunque nel mondo, costruendo delle connessioni con gli host locali ospitanti. Questo aspetto non è da sottovalutare: diversi studi hanno dimostrato l'importanza della posizione dell'alloggio. Tra questi, nell'articolo scritto da Gutierrez (2007) è effettuata un'analisi relativa alla collocazione degli hotel e degli alloggi su Airbnb confrontandoli con i maggiori punti di interesse di Barcellona. I risultati mostrano una maggiore concentrazione di Airbnb nelle zone centrali e più rilevanti, mentre gli hotel sono collocati nelle zone più esterne e lontane.

Possiamo concludere che uno dei fattori che può spingere un visitatore ad orientarsi verso una struttura disponibile su Airbnb piuttosto che un hotel è appunto la vicinanza ai luoghi di interesse della città.

1.3 SODDISFAZIONE DEL CLIENTE

L'obiettivo finale dei fornitori di servizi è la soddisfazione del cliente, che è un fattore importante da considerare nella misurazione della qualità di un servizio.

La soddisfazione del cliente deriva dall'incontro tra le aspettative che si hanno e quelle che invece sono le performance effettive, e si può misurare su più livelli, considerando anche la comparazione che si può effettuare con servizi uguali o simili; se ne parla in relazione al soddisfacimento o meno di diversi elementi, quali, per esempio, i benefici, il rapporto qualità/prezzo, le promesse fatte...

Questa valutazione è basata su percezioni individuali, dipendenti da diversi aspetti che possono essere personali, sociali, culturali.

È importante dunque capire quali sono i requisiti da soddisfare per poter garantire la massima soddisfazione del cliente.

Generalmente, il grado di soddisfazione del cliente si misura tramite dei test realizzati partendo da questionari che permettono di individuare i propri punti di forza e gli aspetti a cui invece possono essere apportate delle migliorie. È logico arrivare alla conclusione che clienti insoddisfatti sono maggiormente meno propensi al riacquisto di quel determinato servizio/prodotto.

Nella letteratura sul tema della soddisfazione del cliente è spesso menzionato il paradigma della conferma/disconferma delle aspettative sviluppato da Richard L. Oliver.

Il modello può essere schematizzato come in figura, dove sono considerati quattro costrutti primari:

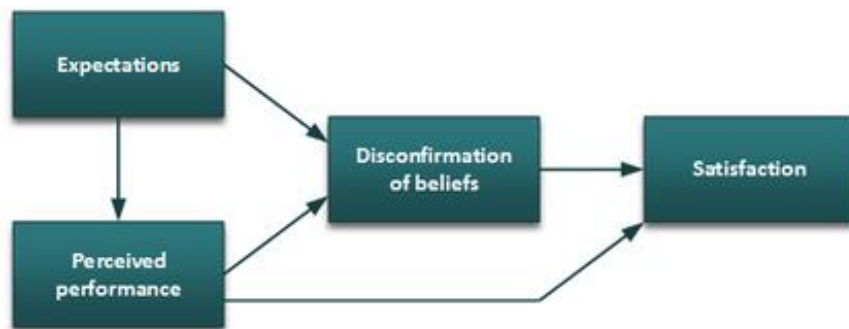


Figura 1: Modello conferma/disconferma

- Le *aspettative* fanno riferimento alle caratteristiche che un individuo immagina associate ad un'entità come un prodotto o un servizio. Secondo il modello queste hanno un'influenza diretta sia sulle performance percepite che sulla disconferma delle convinzioni, mentre indiretta sulla soddisfazione post-acquisto;
- La *performance percepita* si riferisce alla percezione che ha la persona sulle prestazioni effettive del prodotto/servizio acquistato; esercita influenza diretta sulla disconferma delle convinzioni e sulla soddisfazione finale.
- Con il termine *disconferma delle convinzioni* si indicano i giudizi e le valutazioni che un individuo fa rispetto ad un prodotto/servizio, che sono originati dalle aspettative che si hanno. Quando il prodotto/servizio supera le aspettative originali, si dice che la disconferma è positiva, in quanto si presume che aumenti anche la soddisfazione; viceversa, qualora la disconferma sia negativa, vi è un aumento dell'insoddisfazione.
- La *soddisfazione* fa riferimento alla misura in cui una persona è soddisfatta di un prodotto/servizio dopo averlo utilizzato. Secondo il modello, subisce influenza diretta dalla performance percepita e dalla disconferma delle convinzioni.

Un'estensione del paradigma di conferma/disconferma include ulteriori due determinanti della soddisfazione del cliente:

- L'*emozione*, in particolare quella associata al consumo, intesa come una serie di risposte emotive suscitate dall'uso del prodotto e dalla fruizione del servizio;
- Il *prezzo percepito*, che include non solo il prezzo effettivo che bisogna pagare, ma anche tutti gli altri costi associati che bisogna sostenere (come, per esempio, la disponibilità immediata).

Il modello mette in evidenza il ruolo delle aspettative e della percezione dei consumatori in relazione ai prodotti/servizi. Le aspettative possono essere considerate come il valore aggiunto che ci si aspetta di avere, e sono un elemento controllabile dalle aziende tramite, per esempio, azioni di marketing. Vi sono però altri elementi che influenzano le aspettative del consumatore e che non sono facilmente gestibili, come le recensioni dei consumatori, e, come già accennato, le esperienze pregresse che fungono da termine di paragone.

CAPITOLO 2

2.1 ANALISI DELLA LETTERATURA

In questo capitolo si andranno a presentare i risultati ottenuti da diversi studi precedenti riguardo ai fattori critici di successo e i limiti della sharing economy che portano gli utenti ad utilizzare la piattaforma di Airbnb.

Recentemente, la sharing economy è oggetto di un'attenzione significativa, in quanto mezzo di condivisione di risorse altrimenti inutilizzate. Questo fenomeno si differenzia dal classico modello di business per i suoi regimi di produzione-consumo, che prevedono per il fornitore di disporre in affitto per un breve termine i beni propri (come stanze o veicoli), per il consumatore di beneficiare della possibilità di utilizzo di tali beni ad un costo più accessibile.

Esempio lampante nel settore dell'alloggio peer-to-peer è appunto Airbnb, che per la sua popolarità e rapida crescita nel settore del turismo e dell'ospitalità è oggetto di studio da parte di molti ricercatori.

Nisar et al. (2019) cercano di comprendere quali siano le principali determinanti che influenzano le intenzioni di acquisto degli alloggi, basandosi su dati raccolti tramite dei questionari online.

Tra i fattori utilizzati come fulcro dell'analisi vi sono: il valore percepito dell'alloggio, il prezzo percepito dell'alloggio, le informazioni sull'alloggio, le recensioni online, la fiducia nei confronti dell'host.

I risultati del test rivelano che questi fattori appena elencati influenzano positivamente l'intenzione di acquisto: primeggia il valore percepito dell'alloggio, seguito dal prezzo; mentre l'elemento con l'effetto minore risulta essere quello delle recensioni online.

Nell'articolo scritto da Agapitou et al. (2020), attraverso un'indagine quantitativa tramite questionario, sono identificate le motivazioni che fungono da incentivo per i turisti nell'utilizzo della piattaforma, fornendo anche una comparazione con i servizi offerti dagli hotel. Tra i principali incentivi sono stati individuati i costi: lo studio dimostra che più del 70% delle persone utilizza Airbnb perché riescono a trovare facilmente una sistemazione adeguata che soddisfi le loro esigenze a prezzi accessibili, permettendo, di conseguenza, di utilizzare i soldi risparmiati per altre attività locali.

La ricerca ha evidenziato, invece, che le motivazioni esperienziali hanno rilevanza secondaria e che ciò di cui si vantano le attività di sharing economy in generale, cioè il fornire un'esperienza autentica con valore aggiunto legato all'aspetto sociale, non è tra i fattori che guida la soddisfazione del cliente.

Quindi, ruolo principale è svolto dal prezzo, in relazione al quale bisogna considerare il fatto che è associato ad offerte comprendenti interi appartamenti e case, quindi molto più spazio rispetto a quello offerto dai convenzionali hotel.

Secondo sempre questa ricerca, circa il 70% delle persone utilizza la piattaforma perché riesce a trovare una sistemazione in una posizione comoda e strategica. Questo è un vantaggio importante rispetto alle strutture alberghiere standard, che magari a causa di restrizioni nei requisiti locali di uso del suolo, non possono utilizzare la stessa posizione, e ciò permette ad Airbnb di sfruttare la domanda esistente ed ampliare l'offerta di alloggi.

La ricerca condotta da Cheng e Jin (2019) ha avuto come base un database scaricato dal sito "insideairbnb", relativo alla città di Sidney, e consiste in un'indagine su quali siano gli attributi che influenzano le esperienze degli utenti di Airbnb. Il risultato di tale analisi è una serie di

aspetti ritenuti importanti dal cliente, di cui ne sono stati evidenziati dall'esperienza degli ospiti qualità positive e negative.

Il primo attributo chiave è stato indentificato nella posizione, intesa come vicinanza geografica dell'alloggio rispetto a ristoranti, supermercati, negozi, trasporti. Se l'essere in una zona centrale comporta grossi vantaggi e comodità, tra cui la disponibilità dei servizi precedentemente citati, è stato evidenziato anche un aspetto negativo associato al rumore, che però è inevitabile data la vicinanza alle strade principali e al cuore della città.

Molta importanza è attribuita anche alla struttura e alle camere, sottolineando l'apprezzamento dell'alloggio, rispetto al soggiorno offerto dagli hotel, per la sensazione di familiarità e del sentirsi a casa.

Nella ricerca è stata posta molta attenzione alla figura dell'host in relazione a diversi aspetti:

- La comunicazione prima dell'arrivo del cliente è un aspetto importante per far fronte alla mancanza, tipica delle esperienze di transazioni online, dell'interloquire faccia a faccia prima dell'effettivo arrivo nell'alloggio. Dall'analisi risulta che gli ospiti, in generale, apprezzano ricevere più informazioni possibili prima del soggiorno e ciò è considerabile un elemento essenziale per la creazione di fiducia nei confronti dell'host.
- La disponibilità e la comunicazione durante il soggiorno sono ritenuti aspetti aggiuntivi importanti: in generale, infatti, gli host non sono disponibili 24 ore su 24 per rispondere a domande o risolvere problemi, e ciò è più critico nel caso in cui l'host non risieda nella stessa città dell'alloggio che mette in affitto. Quindi, poter contare sui consigli dell'host è una risorsa utile anche per migliorare l'esperienza del soggiorno in relazione, per esempio, a posti da visitare.
- La flessibilità dell'host fa riferimento a quanto l'host sia disponibile nell'accogliere le richieste degli ospiti, soprattutto per quanto riguarda gli orari di check-in e check-out. La ricerca ha risultati

contrastanti, in quanto una buona parte degli host ha avuto punteggi positivi, ma molti sono stati paragonati al servizio rigido ed inflessibile offerto negli hotel.

Un'altra evidenza ottenuta da questo studio è che, contrariamente a quanto ottenuto nell'articolo citato precedentemente, il prezzo non risulta essere un fattore rilevante che influenza la scelta degli utenti che utilizzano Airbnb.

Facendo riferimento ai criteri elencati nel capitolo precedente per ottenere e mantenere il titolo di superhost, Gunter (2017) ha cercato di capire in quale misura le variabili prese in considerazione (elencate di seguito) influenzassero tale status, eseguendo un'indagine sulla città di San Francisco:

- Considerando gli annunci postati dagli host di Airbnb, lo studio ha evidenziato come un numero maggiore ha un effetto negativo sul numero delle prenotazioni; questo trova spiegazione nel fatto che gli ospiti potrebbero non considerare gli annunci autentici;
- È in generale molto apprezzata l'interazione online tra i clienti e gli host di Airbnb, tramite messaggi o e-mail;
- Molto desiderato dagli ospiti risulta essere "l'instabook", opzione che permette la prenotazione dell'alloggio senza bisogno di inviare una richiesta di accettazione all'host e che può essere consentita in maniera arbitraria da quest'ultimo;
- Tra i fattori più importanti in assoluto vi è ancora una volta il prezzo: offrire annunci di alloggio a tariffe inferiori rispetto alle normali tariffe alberghiere è uno dei vantaggi principali della piattaforma Airbnb.

Xie e Mao (2017) hanno effettuato una ricerca sull'impatto degli attributi qualitativi e quantitativi degli host di Airbnb sulla performance delle inserzioni. I dati utilizzati sono relativi agli alloggi situati ad Austin, nel Texas.

Dalla ricerca è emerso che il badge di superhost influisce in modo significativo sulle prenotazioni, così come un alto tasso di risposta ha un impatto positivo; invece, impatto decisamente non rilevante, è risultato quello della verifica dell'identità dell'host.

Uno studio interessante è quello condotto da Han et al. (2019), che utilizzando le affermazioni di Aristotele, tentano di capire quali siano i fattori che influenzano le decisioni di acquisto dei clienti. Anche in questo caso l'indagine è sviluppata da dati provenienti da Airbnb, utilizzando come riferimento il numero di recensioni per andare a considerare gli acquisti effettivi.

Secondo Aristotele, le tre variabili fondamentali per analizzare un messaggio persuasivo e la comunicazione persuasiva sono: ethos, pathos e logos. Gli autori esprimono queste modalità utilizzando diverse variabili indipendenti ed i risultati suggeriscono che tra queste il badge di superhost, il prezzo, l'immagine del luogo e la valutazione a stelle hanno un impatto positivo sugli acquisti.

Uno tra i fattori di maggior influenza che guida gli ospiti nella scelta del luogo in cui soggiornare è il badge di superhost: gli host con tale titolo risultano avere più prenotazioni. In questa ricerca la percentuale di superhost è molto bassa, in quanto appellativo difficile da ottenere. Pertanto, i luoghi messi in affitto da superhost sono selezionati più frequentemente perché ritenuti più credibili.

Le conclusioni relative al prezzo sono interessanti, in quanto se generalmente si pensa che sia preferito un prezzo più basso, quest'analisi ha dimostrato l'opposto: maggiore è il prezzo, maggiore è la probabilità

di acquisto. Questo risultato può trovare spiegazione nel preconetto che ad un costo più alto è generalmente associata una qualità maggiore. Infatti, è difficile stabilire ed accertare la qualità reale di un'offerta, ed il primo metro di giudizio che gli utenti utilizzano è proprio il prezzo.

Questa constatazione sottolinea l'importanza anche dell'evidenza visiva degli annunci di Airbnb, e quindi si può supporre che prezzi più elevati e un numero maggiore di immagini relative all'alloggio abbiano effetti significativi sull'acquisto da parte dei clienti.

Considerando, adesso, le valutazioni, è facile comprendere il loro effetto positivo in quanto uno dei fulcri base delle piattaforme di sharing economy: avere un meccanismo di revisione è importante per guadagnare fiducia (annunci con rating alto sono considerati più affidabili e sicuri).

I risultati della ricerca mostrano che gli alloggi con una capacità di occupazione inferiore sono selezionati più frequentemente dagli utenti di Airbnb: più piccolo è lo spazio, più prenotazioni risulta avere. Gli autori hanno presupposto che le persone preferiscono sistemazioni peer-to-peer più piccole in quanto, rispetto alle prenotazioni per una o due persone, per un numero maggiore è più alta la possibilità di andare incontro a disagi durante il soggiorno.

Come conferma di studi precedenti, anche in questo caso la verifica del profilo dell'host non ha impatto rilevante sulla prenotazione di un alloggio.

Jan Van der Borg et al. (2017) hanno realizzato un'analisi approfondita in relazione ai fattori che influenzano il successo di Airbnb nella regione Veneto. Rispetto alle variabili già presentate in altri studi, sono stati ottenuti i seguenti risultati:

- La tipologia di alloggio ha un impatto significativo sulla scelta del turista; in particolare, anche in questo caso, è dimostrato che le persone sono più propense a prenotare un'intera casa o appartamento, piuttosto che una stanza privata o condivisa;

- Uno dei risultati più importanti riguarda il prezzo: l'elasticità del tasso di occupazione rispetto al prezzo per notte è negativa e all'aumentare del prezzo dell'1% l'occupazione diminuisce dello 0,37%. L'effetto negativo è minimo, quindi si può affermare che il prezzo non influenzi la scelta dei clienti, in quanto a seguito di un aumento del prezzo l'occupazione ne risente negativamente ma in maniera meno che proporzionale;
- Il numero delle camere da letto disponibili nell'alloggio non è rilevante, probabilmente perché agli ospiti interessa maggiormente il numero di letti a disposizione;
- Il numero di recensioni non è importante per il processo decisionale del turista, che dà più peso alla qualità di esse piuttosto che alla quantità;
- La valutazione complessiva è un altro dei fattori chiave che prende in considerazione il turista, in quanto è espressione della qualità percepita del servizio offerto dalla struttura;
- Coerentemente con altri studi, uno dei fattori più influenti è risultato essere lo status di superhost, che ha un'influenza positiva sul tasso di occupazione;
- L'opzione instant booking non ha portato dei risultati significativi;
- Il tempo di risposta influisce sul processo decisionale, ed è una variabile con coefficiente negativo, ciò vuol dire che un tempo di risposta più basso alla richiesta di prenotazione garantisce un tasso di occupazione più alto;
- La posizione dell'alloggio su Airbnb è un fattore chiave che influenza in modo significativo il tasso di occupazione.

Una delle analisi più recenti è stata svolta da Kirkos (2022), che ha esaminato le prestazioni degli annunci di Airbnb con focus nella città di Salonicco in Grecia. Anche in questo caso la base dello studio è stato il set di dati scaricato dal portale "insideairbnb", e ha previsto la

realizzazione di tre differenti modelli con tre variabili dipendenti (occupazione, prenotazioni ed entrate).

I risultati dei modelli sono in linea tra di loro, in quanto il fattore che risulta essere più influente in maniera positiva è il titolo di superhost. Altri elementi, che secondo questa ricerca porterebbero al successo dell'host, sono:

- il tempo di risposta;
- una presentazione dettagliata dell'host;
- dei rating complessivi alti.

Come in altre ricerche, anche in questo caso il prezzo non risulta avere un ruolo importante, anzi contribuisce in maniera marginale sulle prestazioni e non è tra le principali preoccupazioni degli ospiti.

Ulteriore variabile analizzata in questo studio è la distanza degli alloggi rispetto al centro di Salonicco, che però in tal caso risulta non influenzare fortemente le variabili dipendenti. L'autore trova spiegazione nelle caratteristiche specifiche della città, in quanto la stragrande maggioranza degli alloggi di Airbnb si trovano in un'ottima posizione centrale.

Non significativo risulta essere anche l'aspetto della privacy, studiato tramite le variabili relative al tipo di alloggio "private room", "shared room" e "entire room/apt". Come per la distanza, anche in questo caso il risultato può essere spiegato per la questione che a Salonicco la maggior parte degli annunci Airbnb sono relativi ad intere case/appartamenti, dunque non vi è una varietà in relazione alle diverse tipologie sopra citate. Come già detto, gli ospiti sono influenzati dal punteggio della valutazione in stelle, e, inoltre, prestano molta attenzione anche a questioni come il check-in, la comunicazione, l'accuratezza.

In generale, questo studio dimostra il ruolo centrale dell'host, in quanto molti degli attribuiti, direttamente dipendenti dalla sua figura, sono stati evidenziati come forti influencer delle performance degli annunci.

Uno studio condotto su una piattaforma simile ad Airbnb, è quello di Wu et al. (2017), tramite un set di dati scaricato da "Xiaozhu" (la più famosa piattaforma di affitto di alloggi a breve termine in Cina), esplorando gli affetti degli attributi degli host sulla prenotazione delle stanze e degli appartamenti. I risultati evidenziano effetti particolarmente positivi in relazione:

- al tasso di accettazione;
- al numero di annunci relativi alle proprietà dell'host;
- ad una descrizione dettagliata dell'host.

Come in ricerche precedenti, anche in questo caso il prezzo risulta influente, ed alloggi con prezzo minore risultano avere più prenotazioni. Vi è una netta differenza anche sulla tipologia dello spazio affittato, avendo una prevalenza di prenotazioni per interi appartamenti piuttosto che per stanze singole.

Gli autori affermano poi che più informazioni sull'host sono fornite, più i consumatori si fideranno e orienteranno la loro scelta verso un determinato annuncio: infatti, host con pagine personali più curate rientrano tra le preferenze degli utenti.

Non rilevante, invece, risulta essere il tasso di risposta dell'host.

2.2 EVIDENZE DALLA LETTERATURA

Di seguito verrà presentata una schematizzazione dei risultati illustrati precedentemente, per avere una visione diretta sulle evidenze generali, che per alcune variabili son risultate contrastanti nei diversi studi effettuati, e che andranno ad essere testate nuovamente in questo lavoro di tesi.

VARIABILE	È RILEVANTE?		COMMENTO
	Si	No	
Prezzo	X	X	Vi sono risultati contrastanti: alcuni studi pongono la variabile prezzo come primo influente nelle prenotazioni degli alloggi, altri, invece, sono arrivati alla conclusione che non abbia alcun impatto su di esse.
Superhost	X		In generale in tutti gli studi il badge di superhost risulta avere un'influenza altamente positiva e significativa sulla scelta degli alloggi, perché è un indicatore di un buon livello di affidabilità e di un buon comportamento dell'host nella community di Airbnb.
Descrizione dell'host	X		Un host che fornisce una descrizione dettagliata sul suo profilo, e che quindi da più informazioni su di sé, risulta essere preferito in quanto è percepito

			come meno “estraneo” e tende ad essere considerato più affidabile.
Location dell’host	X	X	In generale, la presenza dell’host nella stessa città dell’alloggio può essere vantaggiosa in termini di disponibilità.
Verifica identità dell’host		X	In tutti gli articoli in cui è stata analizzata, questa variabile non risulta influenzare le scelte degli utenti.
Comunicazione con l’host	X		È un fattore ritenuto molto importante, sia prima dell’arrivo nell’alloggio, sia durante il soggiorno.
Check-in e check-out	X		Non è tra i fattori più importanti, ma gli ospiti non si aspettano un servizio rigido come quello offerto nelle strutture alberghiere, ritenendo importante una maggior flessibilità da parte degli host per quanto riguarda gli orari di check-in e check-out in base alle esigenze.
Opzione instant booking	X	X	Non è un’opzione che tutti gli host consentono ed i risultati degli studi precedenti sono contrastanti: a volte è considerata non rilevante, altre volte è dimostrato che host che la permettono ricevono più prenotazioni.
Tasso di risposta		X	Non risulta essere una variabile rilevante nella prenotazione dell’alloggio.
Tasso di accettazione e	X		Sono due variabili che influenzano la scelta dell’utente, che apprezza

Tempo di risposta			ricevere conferme e risposte nel minor tempo possibile.
Immagini dell'alloggio	X		È dimostrato che annunci correlati di immagini del luogo in affitto rientrano nelle preferenze degli ospiti, in quanto rappresentazione effettiva di dove andranno a soggiornare.
Numero di camere da letto		X	I turisti non sono interessati al numero di camere disponibile perché tendenzialmente affittano intere case; danno più peso al numero di letti presenti.
Tipologia dell'alloggio	X		Gli studi che prendono in considerazione questa variabile hanno concordatamente dimostrato che è influente nella scelta, ed in particolare, che sono maggiormente preferiti interi appartamenti alle singole stanze.
Posizione dell'alloggio	X	X	Sono selezionati maggiormente gli appartamenti o le stanze che si trovano più vicini a determinati servizi (ristoranti, musei, supermercati, trasporti...) che tendenzialmente sono presenti nel centro delle città. È una variabile importante che però può risultare rilevante o meno in base alla distribuzione geografica degli alloggi. Nella ricerca sopra citata della città di Salonicco, in Grecia, la vicinanza al centro risultava appunto irrilevante.

Rating complessivo	X		Un elemento decisivo per la scelta di un alloggio è sicuramente la valutazione che questo ha ottenuto dai precedenti clienti. Naturalmente, un utente tenderà a visionare con immediato interesse gli annunci con rating maggiore.
Numero di recensioni		X	Non è un aspetto su cui gli ospiti basano le loro scelte, in quanto interessati più alla qualità che alla quantità.

Tabella 1: Schema riassuntivo

CAPITOLO 3

3.1 DESCRIZIONE DATABASE

In questo capitolo sarà sviluppata una panoramica della selezione delle variabili presenti sul sito "insideairbnb", scelte perché ritenute utili alla creazione del database finale, input del modello di regressione.

Dal sito precedentemente citato, è stato scaricato gratuitamente il set di dati relativo agli host denominato come "listing.csv.gz" della città di Roma. Si è scelta tale città per il semplice motivo che si vuole incentrare l'indagine sull'Italia, considerandone la capitale.

A seguito del download del file si apre un foglio convertito in Excel riportante molteplici informazioni in colonne.

Di seguito vengono elencate con una breve descrizione solo le variabili selezionate e le variabili ricavate da informazioni già esistenti, trasformate e riadattate per l'analisi:

- Listing_id: numero intero che funge da identificatore univoco per l'alloggio;
- Description(length): lunghezza della descrizione dettagliata dell'inserzione;
- Host_id: numero intero identificatore univoco per l'host;
- Host_location: variabile che indica se l'host risiede a Roma o no;
- Host_about(length): lunghezza della descrizione che l'host dà di sé;
- Host_response_Time: velocità di risposta dell'host;
- Host_response_rate: tasso di risposta dell'host;
- Host_acceptance_rate: tasso di accettazione dell'host;
- Host_is_superhost: booleana che indica se l'host è un superhost o no;
- Host_listings_count: numero di annunci di cui dispone l'host;

- Host_has_profile_pic: indica se l'host ha caricato sul profilo una sua immagine;
- Host_identity_verified: indica se l'identità dell'host è verificata;
- Distance: variabile ricavata dalle informazioni sulla latitudine e longitudine dell'alloggio e ne esprime la distanza in Km rispetto al centro di Roma;
- Room_type: tipologia dell'alloggio (private room, entire room/apt);
- Accommodates: numero massimo di persone ospitabili;
- Bathrooms: numero di bagni presenti nell'alloggio;
- Bedrooms: numero di camere da letto presenti nell'alloggio;
- Beds: numero di letti disponibili nell'alloggio;
- Price: prezzo per notte;
- Number_of_reviews: numero di reviews per il listing;
- Review_scores_rating: media ottenuta dai punteggi delle recensioni;
- Review_scores_accuracy: punteggio sulla precisione dell'host;
- Review_scores_cleanliness: punteggio sulla pulizia dell'alloggio;
- Review_scores_checkin: punteggio relativo al check-in;
- Review_scores_communication: punteggio relativo alla comunicazione con l'host;
- Review_scores_location: punteggio relativo alla posizione dell'alloggio;
- Review_scores_value: punteggio sulla qualità del soggiorno;
- Instant_bookable: booleana che indica se l'ospite può prenotare direttamente l'alloggio senza che l'host debba accettare la richiesta di prenotazione.

Considerando le variabili sopra descritte, si è operata una pulizia del database andando ad eliminare le righe che risultavano vuote o con caratteri non validi in relazione ad alcune informazioni. Il file iniziale presentava 24626 righe; dopo la pulizia si è arrivati al file definitivo composto da un totale di 11385 righe.

Nel paragrafo successivo verranno analizzate le variabili considerate e, grazie all'utilizzo di Excel e del software Stata, verranno studiate le statistiche descrittive di tali variabili, che saranno utili nel capitolo successivo per effettuare l'analisi di regressione.

3.2 ANALISI DELLE VARIABILI

In questo paragrafo si mostrano le statistiche descrittive relative alle variabili selezionate, per analizzarle più nel dettaglio e cercare di stimarne la distribuzione.

Come prime saranno presentate le cosiddette “*variabili dummy*”, che sono delle variabili che possono assumere solo due valori, tipicamente 0 e 1; vengono chiamate anche *dicotomiche* e possono essere utilizzate per rappresentare due casi opposti (per esempio: la variabile assume valore 1 se l’elemento sotto esame è femmina, mentre assume valore 0 se maschio). Successivamente saranno presentate le restanti variabili ed infine quelle che sono state introdotte, ricavate da informazioni presenti nel file Excel.

3.2.1 Host_location

È una variabile dummy che indica se l’host è presente nella città di Roma: assume il valore 1 se è presente, 0 altrimenti.

Come si può notare dal grafico, sul totale dei dati a disposizione il 79% degli host risulta essere presente a Roma e solo il 21% risiede al di fuori della città. Questa è un’ottima percentuale se consideriamo l’opportunità d’interazione diretta tra host e ospiti.

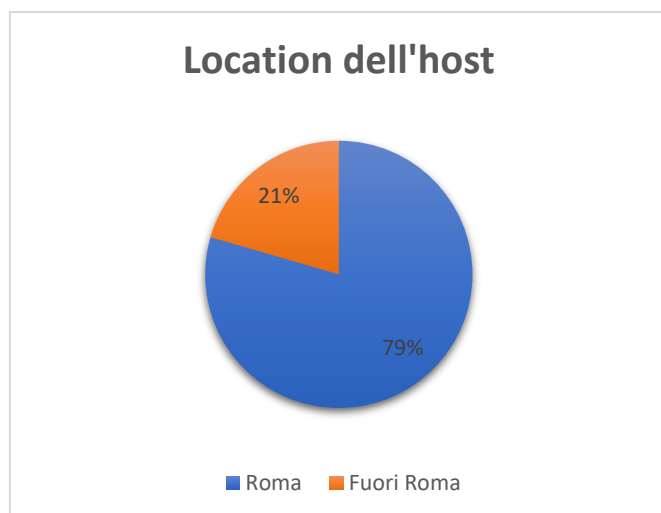


Figura 2: Grafico location

3.2.2 Host_is_superhost

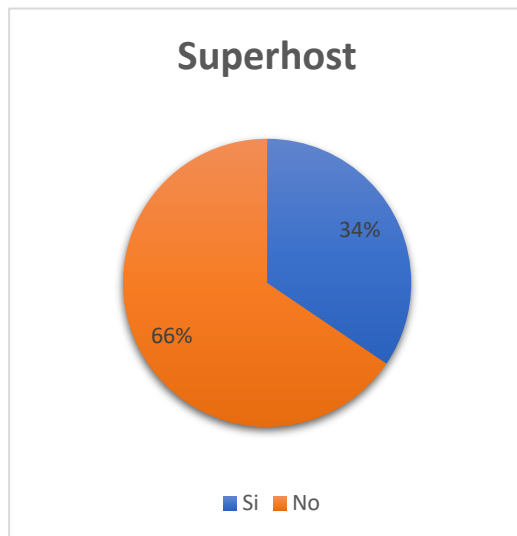


Figura 3: Grafico superhost

Si tratta di una variabile dummy che assume valore 1 nel caso in cui l'host abbia il titolo di superhost, 0 altrimenti. La figura mostra una percentuale del 34% di superhost, meno della metà del totale, ma ciò può trovare spiegazione dal fatto che, come illustrato nel capitolo precedente, vi sono molti vincoli da rispettare per ottenere tale titolo.

3.2.3 Host_has_profile_pic

Questa variabile assume valore 1 nel caso in cui l'host ha pubblicato una sua foto sul profilo, mentre 0 nel caso opposto. La percentuale delle persone sprovvista di foto profilo è quasi dello 0%, confermando l'importanza del farsi visivamente individuare e conoscere. In generale, infatti, l'ospite si sente più sicuro nel relazionarsi con un host di cui già ne conosce l'aspetto, abbattendo un piccolo strato di quella barriera di diffidenza che tutti possiedono nel momento in cui si ha un approccio con qualcosa/qualcuno di estraneo.

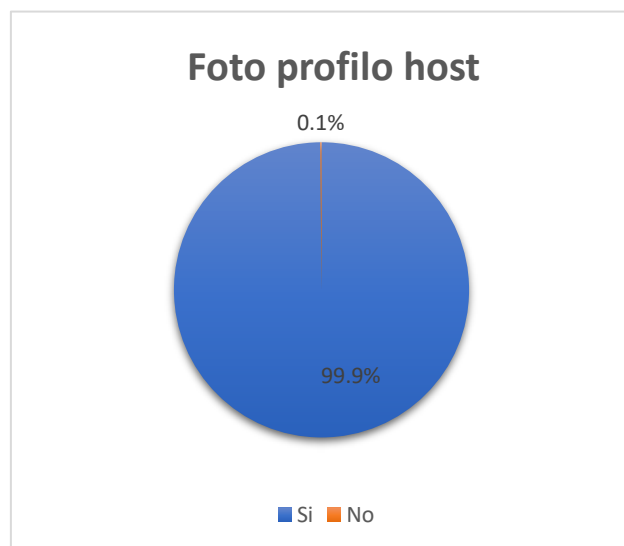


Figura 4: Grafico foto profilo

3.2.4 Host_identity_verified

Un'altra variabile importante è quella che indica se un host possiede un'identità verificata oppure no. La verifica da parte del sito di Airbnb è infatti un elemento che dà maggior sicurezza ai clienti riguardo alla credibilità e quindi anche all'affidabilità dell'host, ed è un fattore che sicuramente

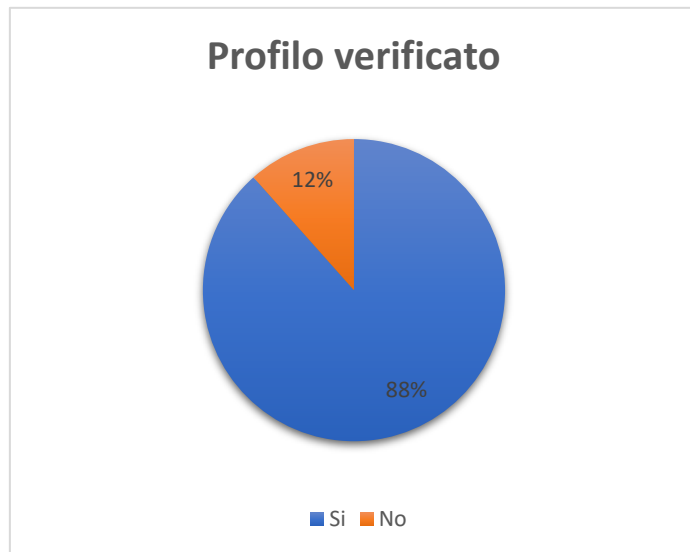


Figura 5: Grafico profilo verificato

viene preso in considerazione nel momento della scelta dell'alloggio. Anche in questo caso la variabile considerata assume due valori: 1 se il profilo risulta verificato, 0 altrimenti.

In figura è possibile notare che la maggior parte degli host ha un profilo verificato, mentre solo il 12% rientra nel caso apposto.

3.2.5 Instant_bookable

Come precedentemente accennato, tra le possibili azioni che un utente può compiere vi è quella della prenotazione immediata senza necessaria accettazione da parte dell'host, opzione che il singolo host può decidere autonomamente se rendere disponibile o meno.

"Instant_bookable" è una variabile dummy che assume valore 1 nel caso in cui è permessa la prenotazione immediata, 0 altrimenti.

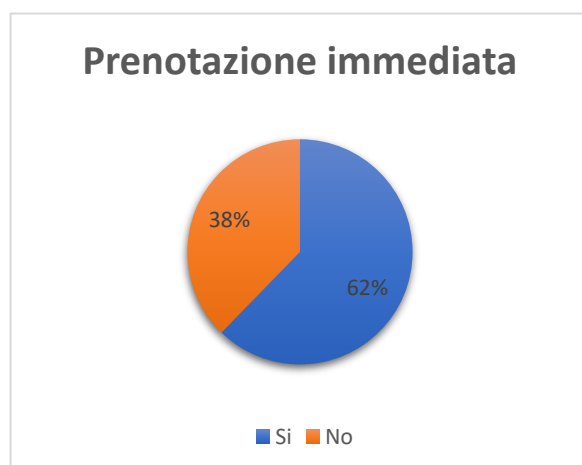


Figura 6: Grafico prenotazione

Il grafico mostra che il 62% degli host permette tale azione, mentre il restante 38% preferisce visionare ogni richiesta per poi eventualmente darne il consenso.

3.2.6 Room_Type

In relazione all'alloggio messo in affitto, sulla piattaforma Airbnb è possibile trovarne 4 tipologie: private room, entire room/apartment, shared room e hotel room.

In figura sono riportate le percentuali per ogni tipo di listing.

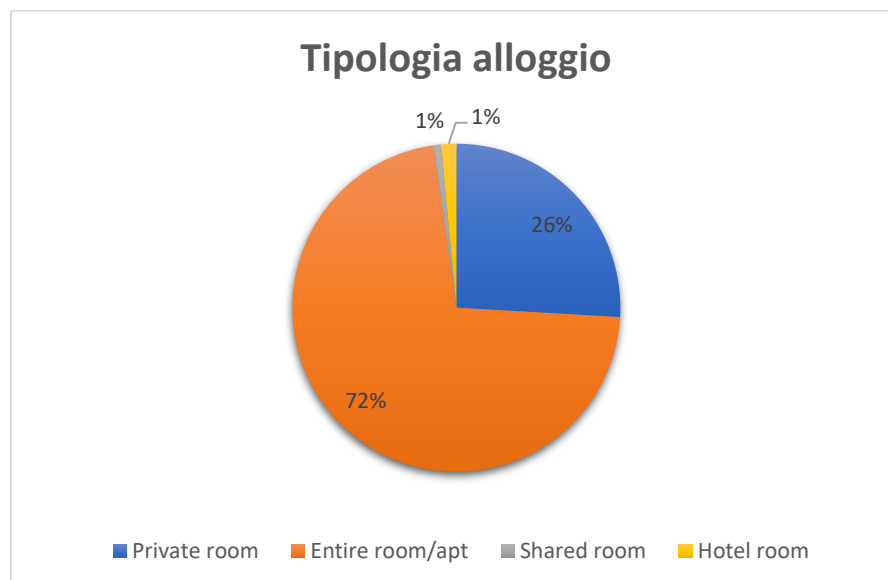


Figura 7: Grafico tipologia alloggio

Come si può notare, le percentuali di "shared room" e "hotel room" risultano essere pari all'1%, pertanto, non essendo una percentuale rilevante sul totale, nell'analisi verranno considerate solo le altre due categorie.

L'ospite, dunque, può decidere se usufruire di una camera singola oppure di affittare un intero appartamento. Il tipo di alloggio più richiesto risulta essere l'intero appartamento, con 72 punti percentuali, mentre la seconda scelta è la stanza singola non condivisa, con 26 punti percentuali.

La variabile "room_type" diventa dunque una dummy che assume valore 0 se l'alloggio è una stanza singola, 1 se è un intero appartamento.

3.2.7 Host_response_time

È una variabile che nel file scaricato si presenta in 4 possibili attributi, che sono stati tradotti in numero, in relazione al tempo che impiega l'host nel rispondere alle prenotazioni:

Within an hour	1
Within a few hours	2
Within a day	3
A few days or more	4

Il valore massimo (4) corrisponde alla situazione più negativa che si può presentare, mentre il valore minimo (1) corrisponde ad un tempo di risposta rapido.

Tabella 2: Tempi di risposta

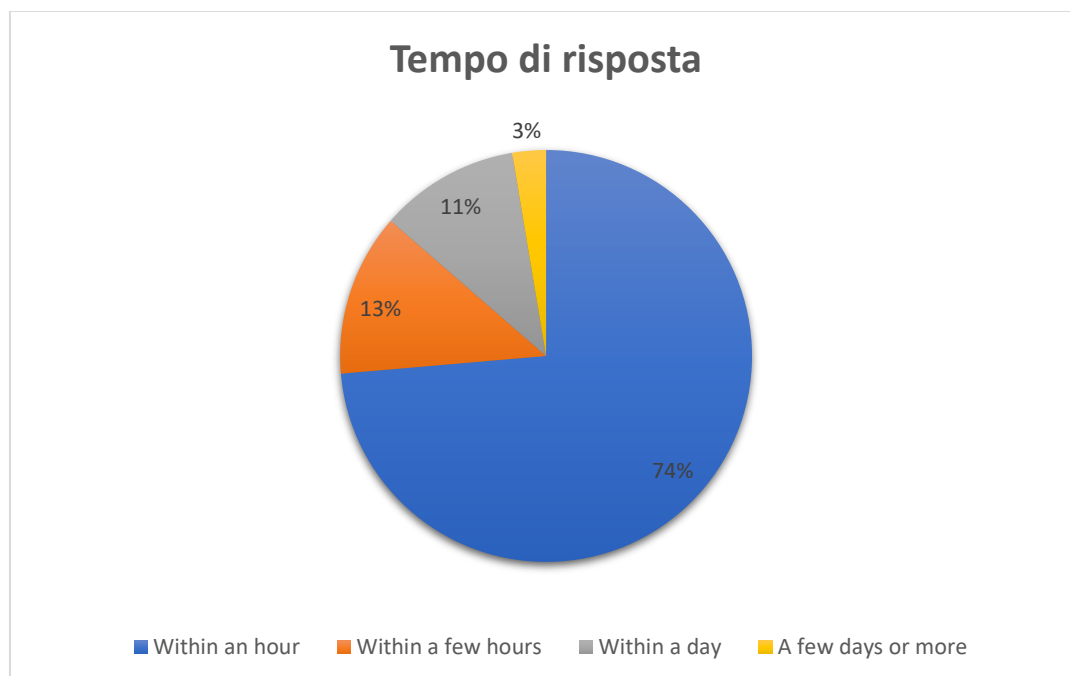


Figura 8: Grafico tempi di risposta

Dal grafico, è possibile avere una visione più chiara della distribuzione delle tempistiche di risposta: una buona percentuale degli host, il 74%,

risponde alle informazioni richieste dagli ospiti in un tempo inferiore ad un'ora; le restanti percentuali risultano rispondere entro comunque la giornata, mentre solo il 3% impiega un tempo che supera un giorno o anche più.

3.2.8 Host_response_rate e Host_acceptance_rate

Sono dei valori scaricati in percentuale, ma trasformati in numeri decimali per essere leggibili da Stata senza problemi.

Variabile	Valore minimo	Valore massimo	Media	Deviazione Standard
Host_response_rate	0	1	0.9410198	0.1623431
Host_acceptance_rate	0	1	0.8584014	0.2668422

Tabella 3: Statistiche descrittive tasso di risposta e di accettazione

Come si nota nel seguente grafico, circa 8500 host sul totale presentano un tasso di risposta massimo, e ciò non stupisce data l'importanza di tale elemento logistico che risulta fondamentale affinché possa avvenire la transazione online con il cliente. La media risulta essere circa del 94%, in quanto sono presenti anche pochi ma sparsi punti con percentuali più basse.

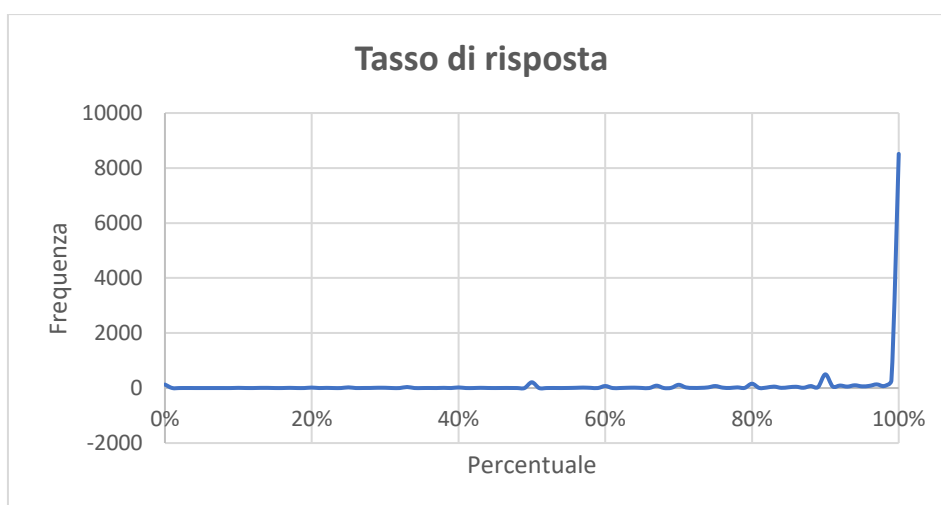


Figura 91: Andamento tasso di risposta

Per quanto concerne il tasso di accettazione, anche in questo caso vediamo la maggior parte dei valori distribuiti a partire dal 90% in poi. La media risulta più bassa, circa dell'86%, perché sporcata da una presenza maggiore di dati negli intervalli percentuali più bassi (circa 570 osservazioni ricadono tra l'0% e il 2%). Questi risultati possono dipendere da vari fattori: si può considerare il fatto che tendenzialmente gli host sono propensi ad accettare qualsiasi richiesta, o ancora l'opportunità dell'opzione di prenotazione immediata consentita dalla piattaforma, grazie alla quale l'utente non deve aspettare l'approvazione da parte dell'host, portando, quindi, il tasso di accettazione al valore massimo.

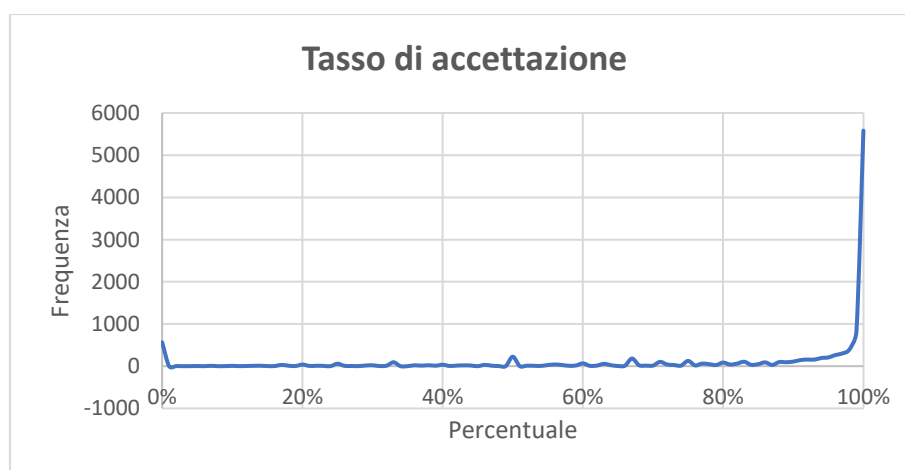


Figura 10: Andamento tasso di accettazione

3.2.9 Host_listings_count

Di seguito sono riportate le principali statistiche descrittive e un grafico relativo al numero di annunci per host.

Valore minimo	Valore massimo	Media	Deviazione Standard
1	277	12.67185	38.91948

Tabella 4: Statistiche descrittive numero di annunci

Come si nota chiaramente dal grafico, tendenzialmente la maggior parte degli host possiede un solo annuncio; si ha una lenta diminuzione del

numero di annunci, fino ai 13 per host, dove però si è ancora sopra le centinaia di affittuari. Successivamente, vi sono dei piccoli picchi in corrispondenza, per esempio, di 30-60 annunci, ed un numero consistente di host (155) risulta averne 248.

Una considerazione finale può essere fatta in merito al fatto che un numero maggiore di annunci può essere correlato ad un'esperienza superiore dell'host, ma anche che un minor numero di annunci da gestire comportano maggiore disponibilità per i singoli ospiti.



Figura 11: Andamento numero di nnunci

3.2.10 Accommodates

La variabile "accomodates" fa riferimento al numero di persone ospitabili nell'alloggio, e quindi dà un'informazione anche in relazione alla grandezza della struttura.

Valore minimo	Valore massimo	Media	Deviazione Standard
1	16	3.980061	2.178532

Tabella 5: Statistiche descrittive numero di persone ospitabili

La media complessiva sul totale degli annunci risulta essere di circa 4 persone ad annuncio, così come è mostrato nel grafico: più di 3000 alloggi sono adatti a 2-4 persone, circa 1500 a 3-6.

In base a questi dati, si può dedurre che una buona parte degli alloggi corrisponde ad appartamenti costituiti da almeno 2 stanze, riducendo al minimo la probabilità della presenza di semplici monolocali.

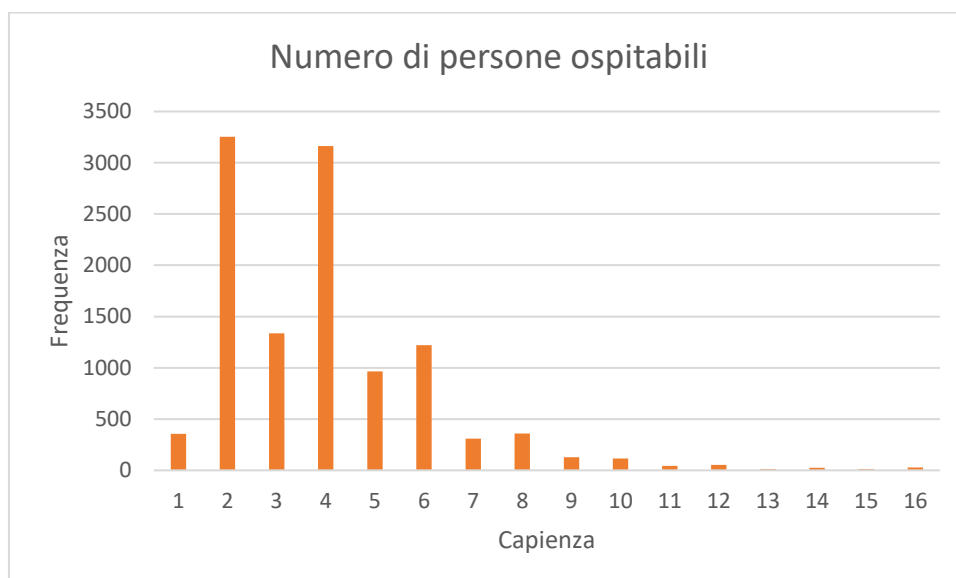


Figura 12: Istogramma numero di persone ospitabili

3.2.11 Bathrooms

Di seguito è stata analizzata la variabile relativa al numero di bagni presenti in ogni alloggio.

Valore minimo	Valore massimo	Media	Deviazione Standard
1	8	1.318665	0.679995

Tabella 6: Statistiche descrittive numero di bagni

Dal seguente grafico è evidente che quasi la totalità degli annunci risulta comprende un solo bagno per alloggio, circa 2000 almeno 2 bagni, mentre le restanti centinaia del totale sono sparse scarsamente tra le rimanenti cardinalità (infatti, solo 4 annunci corrispondono al numero massimo di letti disponibili).

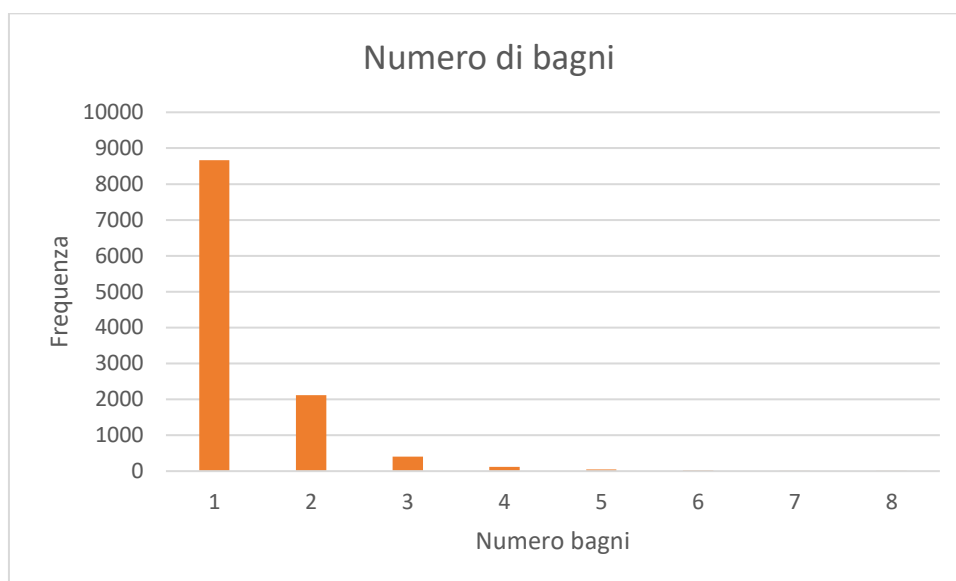


Figura 13: Istogramma numero di bagni

3.2.12 Bedrooms e Beds

Anche queste due variabili possono dare un'indicazione in merito alla grandezza dell'alloggio offerto.

Variabile	Valore minimo	Valore massimo	Media	Deviazione Standard
Bedrooms	1	10	1.524989	0.8503203
Beds	1	20	2.371981	1.664701

Tabella 7: Statistiche descrittive numero di camere da letto e letti

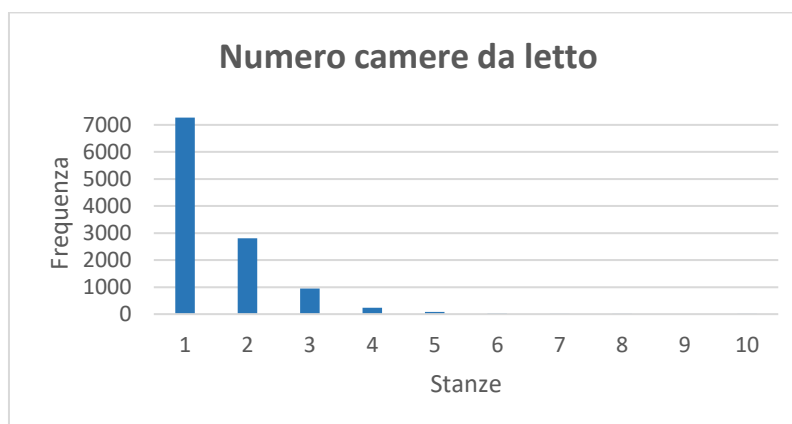


Figura 142: Istogramma numero di camere da letto

In relazione alle analisi svolte finora sulle altre variabili correlate alla tipologia dell'alloggio, sono stati ottenuti dei risultati che un po' stupiscono, in quanto, specie per il numero di letti, ci si poteva aspettare dei valori più alti rispetto alla media riportata in tabella. Tuttavia, bisogna considerare il fatto che non è specificato se si tratta di letti singoli o matrimoniali, quindi, anche i risultati ottenuti con la variabile "accomodates" non sono così contrastanti.

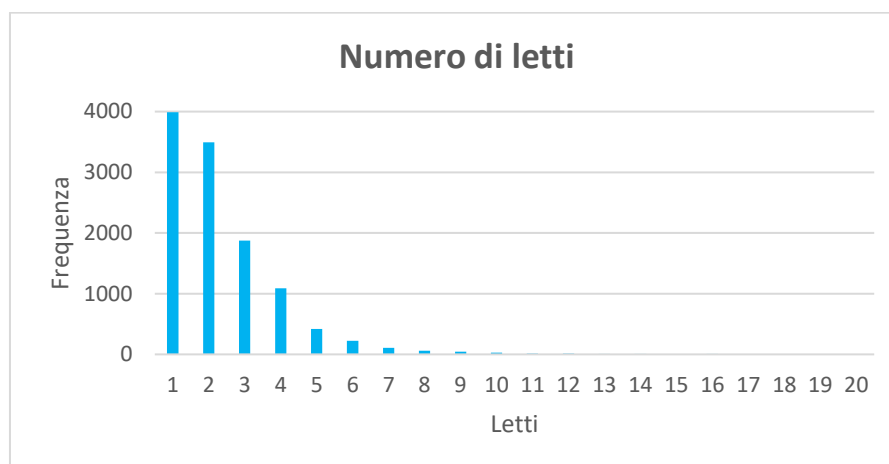


Figura 15: Istogramma numero di letti

Dal grafico si nota come quasi un terzo degli alloggi risulta avere solo un letto, e questa frazione potrebbe corrispondere alla percentuale di stanze singole individuate precedentemente.

3.2.13 Price

La variabile "price" è risultata una tra quelle di maggiore impatto nella letteratura.

I dati nel file scaricato erano disponibili in dollari, pertanto sono stati convertiti in euro ed il valore è riferito al prezzo per una notte.

Valore minimo	Valore massimo	Media	Deviazione Standard
14	885	91.5639	85.50109

Tabella 8: Statistiche descrittive del prezzo

Come mostrato nel grafico, l'andamento è abbastanza disperso, con una maggior concentrazione sotto i 100 euro. Infatti, la media risulta essere di circa 92 euro a notte, mentre si nota una piccola concentrazione di punti verso il valore massimo, che possiamo presupporre siano valori riferiti ad appartamenti di lusso.

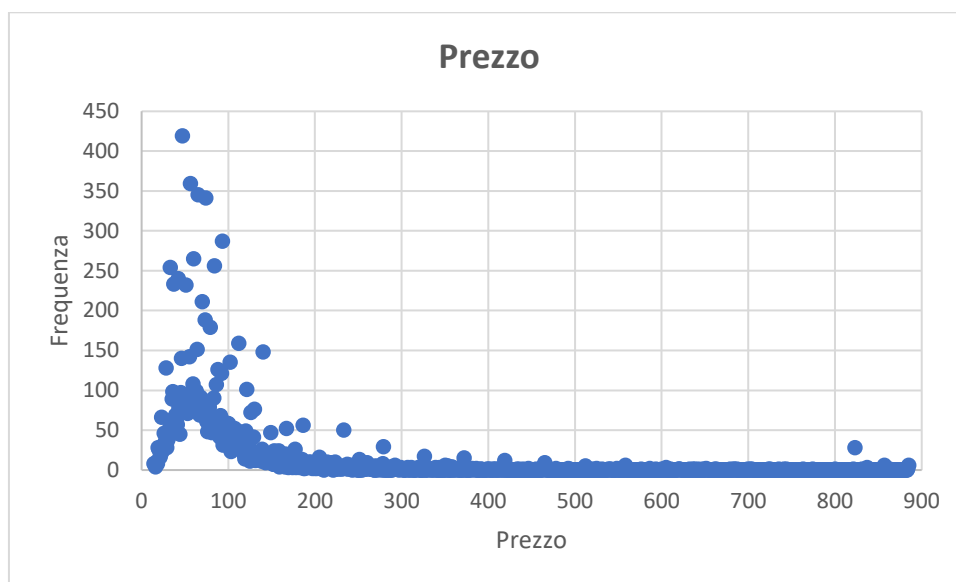


Figura 16: Andamento del prezzo

3.2.14 Number_of_reviews

Valore minimo	Valore massimo	Media	Deviazione Standard
1	829	63.78103	84.76314

Tabella 9: Statistiche descrittive numero di recensioni

La variabile "number_of_reviews" tiene conto del numero di recensioni che gli ospiti hanno lasciato dopo il soggiorno nell'alloggio scelto.

Graficamente si ha una sorta di curva iperbolica, con una concentrazione maggiore di dati nell'intervallo 1-100. La media è di circa 64 recensioni ad annuncio.



Figura 17: Andamento numero di recensioni

3.2.15 Review_scores

Dal file scaricato da "insideairbnb" sono presenti variabili relative a diversi aspetti i cui valori, non essendone data un'indicazione sul dizionario delle variabili presenti sul sito, si presuppone siano una media dei punteggi assegnati dagli ospiti alla fine del soggiorno. Come già detto nei capitoli precedenti, è una valutazione in stelle, con valore minimo 0 fino ad un valore massimo 5. I dati, tuttavia, si presentavano con valori compresi tra 0 e 499. Per un'analisi più semplice e per rendere tali valori in linea con il sistema di valutazione di Airbnb, si è deciso di dividere i dati per 100 e di procedere con un arrotondamento, così da avere una scala di numeri da 0 a 5.

Variabile	Valore Minimo	Valore Massimo	Media	Deviazione Standard
review_scores_rating	0	5	3.4426	2.061228
review_scores_accuracy	0	5	3.445235	2.073875
review_scores_cleanliness	0	5	3.432587	2.054393
review_scores_checkin	0	5	3.327361	2.08184
review_scores_communication	0	5	3.312956	2.077257
review_scores_location	0	5	3.3639	2.054454
review_scores_value	0	5	3.498199	2.075825

Tabella 10: Statistiche descrittive dei vari rating

In tabella sono elencate tutte le variabili con le principali statistiche descrittive e di seguito sono presentati i grafici relativi ad ognuna di esse.

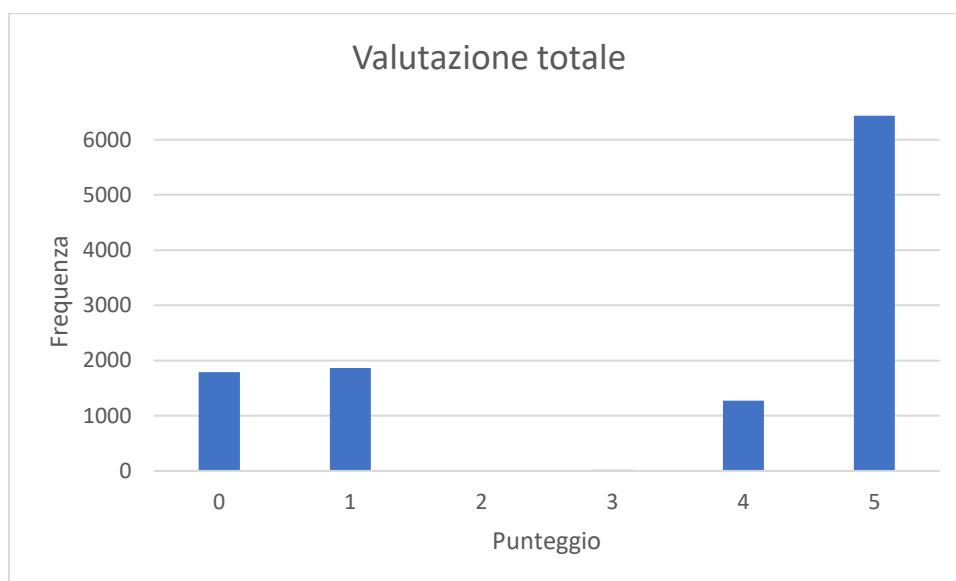


Figura 38: Istogramma rating complessivo

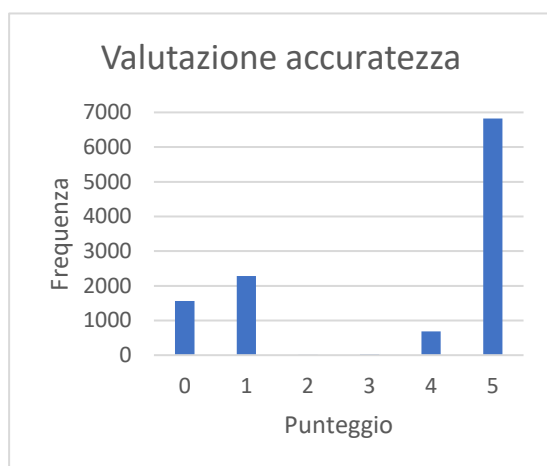


Figura 195: Istogramma precisione dell'host

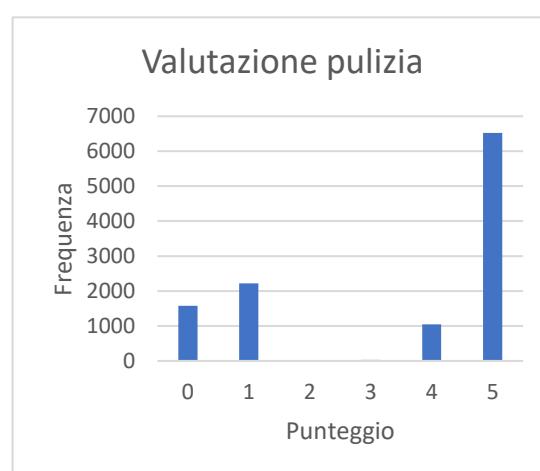


Figura 4: Istogramma pulizia

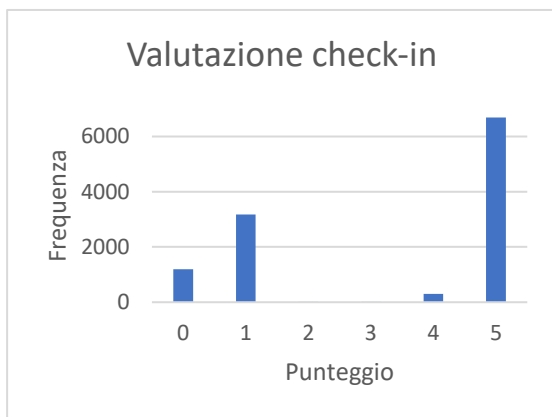


Figura 61: Istogramma check-in

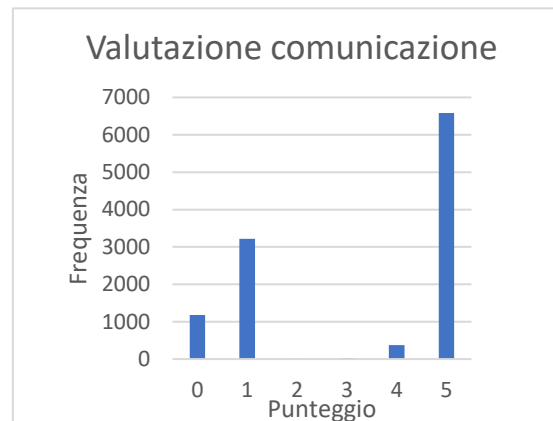


Figura 22: Istogramma comunicazione

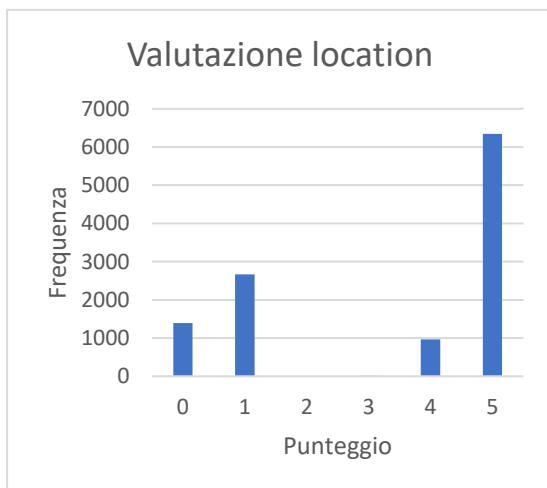


Figura 83: Istogramma location



Figura 7: Istogramma qualità

Tutti i grafici relativi alle diverse variabili presentano nella pratica lo stesso andamento: difatti, le medie e le deviazioni standard risultano pressoché simili.

La variabile "review_score_rating" rappresenta un punteggio generale, ricavato dalla media di tutte le altre valutazioni. Si nota uno sbilanciamento delle valutazioni verso il punteggio massimo, e una piccola concentrazione anche nei valori più bassi. Si presuppone, perciò, che vi sia una qualche distorsione nei giudizi degli ospiti, in quanto i valori intermedi non sono praticamente considerati.

3.3 VARIABILI AGGIUNTE

3.3.1 Host_about(length)

Questa variabile è stata ricavata dall'informazione relativa alla descrizione che l'host dà di sé. Innanzitutto, sono stati analizzati i dati a disposizione per capire quale percentuale del campione avesse effettivamente una descrizione dell'host: è stato posto come vincolo un minimo di 10 parole così da escludere eventuali semplici frasi come "ciao a tutti" presenti nel database e non ritenute idonee al significato della variabile.

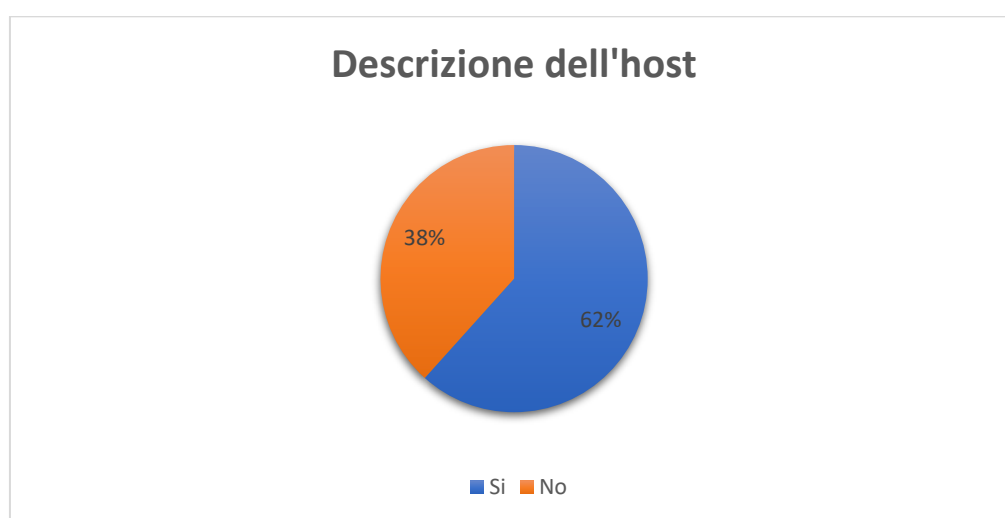


Figura 25: Grafico descrizione host

In base a questa analisi il 62% degli host risulta aver fornito una descrizione di sé, il restante 38% ne è sprovvisto.

È stata dunque ricavata la variabile "Host_about(length)", tramite la funzione di lunghezza su Excel che permette di calcolare il numero di parole della descrizione data dall'host.

Tabella 11: Statistiche descrittive lunghezza descrizione host

Variabile	Valore minimo	Valore massimo	Media	Deviazione Standard
Host_about(length)	0	3033	44.30435	81.52311

Dal grafico si nota subito il valore più alto corrispondente alla percentuale di persone che non fornisce una descrizione di sé. La rimanente percentuale di host è distribuita lungo l'asse, ed in particolare, come ci suggerisce anche la media, tendenzialmente la maggior parte delle persone sta sotto le 50 parole, e, come è maggiormente messo in evidenza dalla tabella in cui sono stati considerati degli intervalli da 50 parole, centinaia di host rientrano anche negli intervalli fino a 250 parole. Questa variabile è un'evidenza della volontà dell'host di raccontare di sé stesso, e dunque potrebbe anche essere un indizio positivo in merito alla sua predisposizione nella creazione di un rapporto con l'ospite.

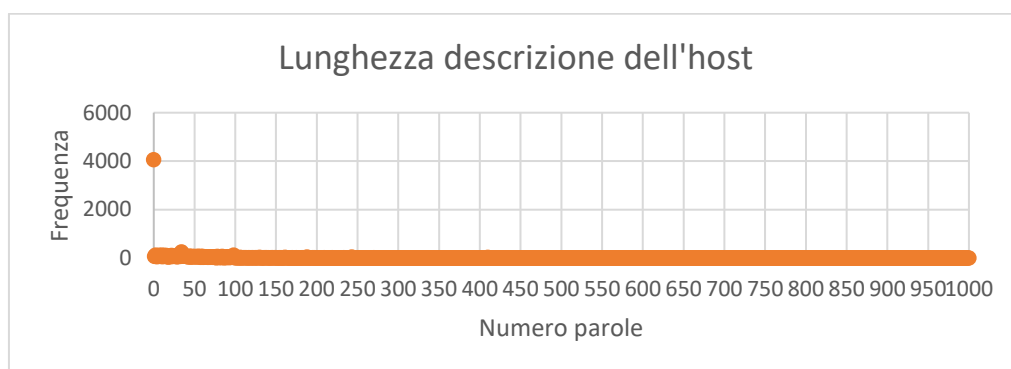


Figura 26: Andamento lunghezza descrizione host

Intervallo numero parole	Conteggio	Frequenza
0-10	4368	38.4%
11-50	3698	32.5%
51-100	1940	17.0%
101-150	659	5.8%
151-200	328	2.9%
201-250	192	1.7%
251-300	67	0.6%
301-350	44	0.4%
351-400	17	0.1%
401-450	37	0.3%
451-500	11	0.1%
501-550	2	0.0%
>550	22	0.2%

Tabella 12: Schema lunghezza descrizione host

3.3.2 Description(length)

La variabile "description" presente nel file fa riferimento alla descrizione che l'host fa dell'alloggio che mette in affitto. effettuando un'analisi preliminare simile alla precedente, in questo caso si ha che quasi la totalità degli host fornisce una descrizione dell'annuncio, e solo una minima porzione del 2% ne è sprovvista.

Ovviamente, più un ospite conosce dello spazio in cui soggiornerà, più è prepenso a visionare e scegliere quel determinato alloggio.

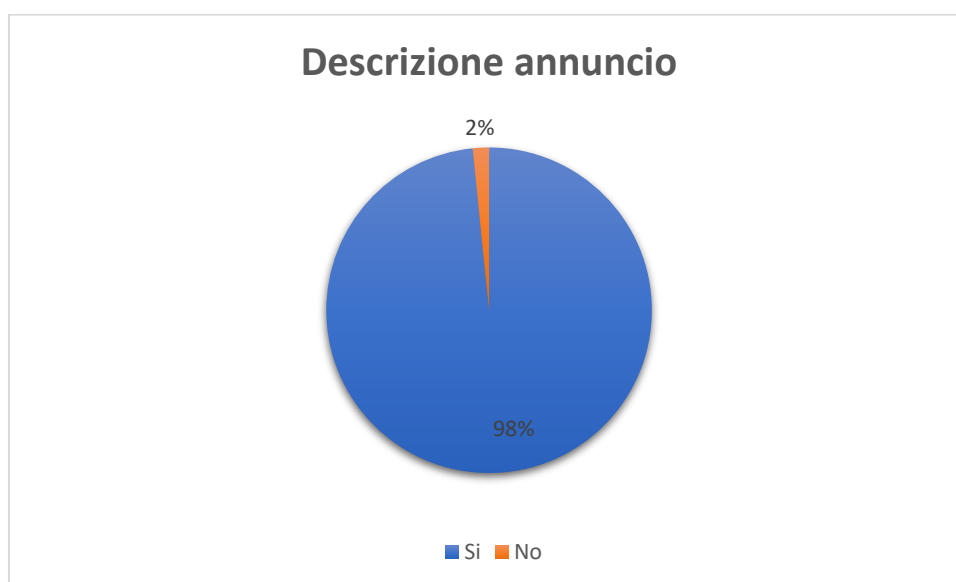


Figura 27: Grafico descrizione annuncio

È stata quindi calcolata la lunghezza della descrizione dell'annuncio fornita dall'host.

Variabile	Valore minimo	Valore massimo	Media	Deviazione Standard
Description(length)	0	194	125.9693	47.16041

Tabella 13: Statistiche descrittive lunghezza descrizione annuncio

Dal grafico si nota che non vi sono particolari concentrazioni di dati, che risultano abbastanza distribuiti tra il minimo e il massimo valore, come conferma la deviazione standard che è quasi del 50%. Il picco si ha

nell'intervallo tra le 140 e le 170 parole, e mediamente ogni host scrive circa 126 parole di descrizione.



Figura 28: Andamento lunghezza descrizione annuncio

Intervallo numero parole	Conteggio	Frequenza
0-10	234	2%
11-20	99	1%
21-30	160	1%
31-40	391	3%
41-50	328	3%
51-60	284	2%
61-70	364	3%
71-80	561	5%
81-90	438	4%
91-100	354	3%
101-110	344	3%
112-120	377	3%
121-130	374	3%
131-140	547	5%
141-150	1122	10%
151-160	2232	20%
161-170	2334	21%
171-180	758	7%
181-194	84	1%

Tabella 14: Schema lunghezza annuncio

Sarà interessante studiare se e come la lunghezza sia della descrizione dell'host, sia dell'annuncio, hanno influenza sulla scelta dei guest. Tenzionalmente si potrebbe pensare che una descrizione più lunga abbia degli effetti positivi, in quanto più completa e dettagliata.

3.3.3 Distance

La variabile "distance" è stata ricavata dalle informazioni della latitudine e longitudine relative ad ogni alloggio. In particolare, è stata utilizzata la seguente *formula di Haversine*, che permette, appunto, di calcolare la distanza in chilometri tra le coordinate di punti specificati:

$$ARCCOS(SEN(latitudine1)*SEN(latitudine2)+COS(latitudine1)*COS(latitudine2)*COS(longitudine2-longitudine1))*6371$$

Sono state utilizzate le seguenti coordinate come riferimento latitudine e longitudine del centro di Roma:

- Latitudine: 41°54'39"24 N
- Longitudine: 12°28'54"48 E

Valore minimo	Valore massimo	Media	Deviazione Standard
0.19	28.72	3.670565	3.93354

Tabella 15: Statistiche descrittive distanza

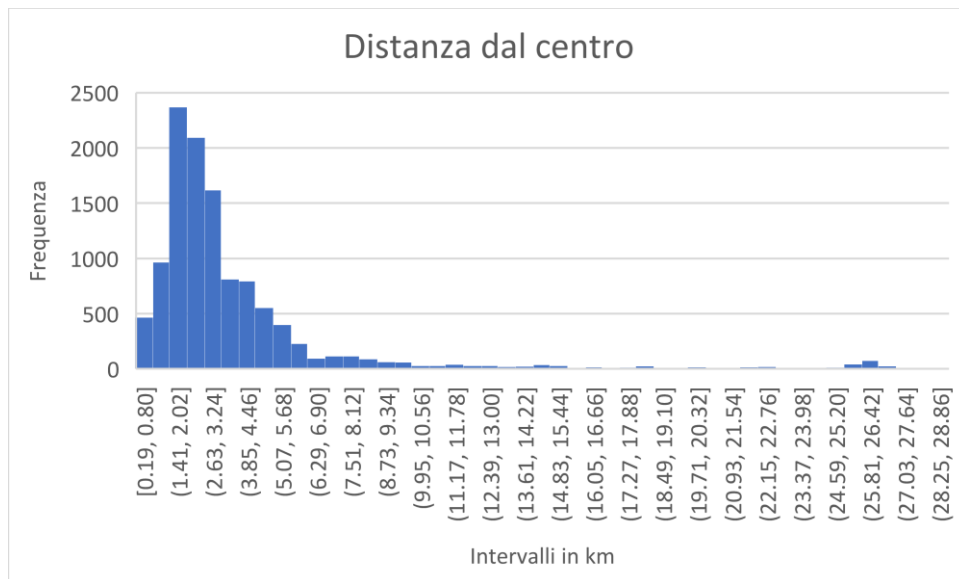


Figura 29: Istogramma distanza

La maggior parte degli alloggi risulta essere localizzata non oltre i 4 Km rispetto al centro, e solo poche centinaia si trovano oltre i 10 Km.

Possiamo presupporre che tendenzialmente gli host mettano a disposizione locali non troppo distanti dal centro, in quanto la vicinanza ad esso è ritenuta importante ed un fattore decisionale influente nelle scelte degli ospiti.

CAPITOLO 4

4.1 MODELLI DI REGRESSIONE LINEARE

La regressione lineare è una tecnica statistica che si utilizza per studiare la relazione tra due o più variabili. Si distinguono in generale due tipi di regressioni lineari in base al numero di variabili indipendenti del modello, dette **regressori**, ed entrambe le tipologie verranno utilizzate in questa analisi:

- **Regressione lineare semplice:** si studia la relazione tra due sole variabili, in quanto c'è una variabile dipendente (Y) e un regressore (X);
- **Regressione lineare multipla:** ci sono più regressori e dunque si studia l'effetto contemporaneo di più variabili X sulla variabile dipendente Y.

I modelli di regressione servono per stimare il valore della variabile Y con la maggior precisione possibile partendo dai valori delle variabili esplicative. La regressione lineare si utilizza perché si ipotizza che ci sia una correlazione lineare tra le variabili, e nella pratica vuol dire stimare la retta che meglio leghi le variabili: permette di analizzare la relazione tra di esse, studiarne la direzione e la significatività.

Nel caso di un modello di regressione lineare semplice, la retta assume la seguente forma:

$$Y_i = \beta_0 + \beta_1 X_i + u_i \quad i=1, \dots, n$$

Dove

- n è il numero di osservazioni
- X è la variabile indipendente o regressore

- Y è la variabile dipendente
- β_0 rappresenta l'intercetta
- β_1 è la pendenza della retta
- u_i è l'errore o residuo della regressione

L'output della regressione derivato dal software Stata si presenterà sotto questa forma:

- Nella prima colonna della tabella si trovano la variabile dipendente e al di sotto di essa quella indipendente;
- La seconda colonna presenta la stima dei parametri β_0 e β_1 ;
- Nelle colonne successive viene effettuato un test di ipotesi per verificare che i parametri siano diversi da 0 e viene presentato un intervallo di confidenza sul valore del parametro. Il test è di particolare importanza soprattutto per β_1 , poiché definisce statisticamente quanta probabilità ci sia che la relazione tra le due variabili in esame sia nulla. La distribuzione utilizzata è la T-student, e si verifica se la statistica t calcolata rientra nell'intervallo di confidenza posto al 95% (-1.96;1.96). Nel caso in cui la t non rientri nell'intervallo non si può accettare l'ipotesi iniziale che β_1 sia 0, e ciò viene confermato dal valore del p-value prossimo allo 0. Infatti, si vuole un p-value pari a 0 per affermare che la relazione tra le variabili esista con un grande margine di probabilità, e che quindi la variabile indipendente sia statisticamente significativa. Un'ulteriore conferma del test di ipotesi si ha se lo 0 non risulta essere compreso nell'intervallo di confidenza.

I successivi sottocapitoli saranno strutturati nel seguente modo: inizialmente sarà mostrata la correlazione della variabile dipendente con i singoli regressori; successivamente si andranno a strutturare dei modelli

di regressione multipla divisi in base alla tipologia delle variabili. Infine, verranno presentati i modelli completi di tutte le variabili.

4.2 MODELLI DI REGRESSIONE LINEARE SEMPLICE

Essendo lo scopo di questo lavoro di tesi cercare di capire quali sono i principali fattori di successo di un host, che quindi portano ad una maggiore richiesta dei propri alloggi, si è deciso di utilizzare come variabile dipendente quella relativa al numero di recensioni totali ricevute, ovvero "number_of_reviews".

La variabile, infatti, si presta bene a rappresentare il tasso di domanda di un alloggio, ed è stata scelta perché utilizzata anche in altri studi presentati nel capitolo sull'analisi della letteratura; dunque, è possibile fare una comparazione coi risultati ottenuti da analisi precedenti in modo coerente e non ambiguo.

4.2.1 Domanda e location dell'host

```
. regress number_of_reviews host_location, robust
```

Linear regression

Number of obs = 11385
F(1, 11383) = 310.61
Prob > F = 0.0000
R-squared = 0.0172
Root MSE = 84.037

number_of_r~s	Robust		t	P> t	[95% Conf. Interval]	
	Coef.	Std. Err.				
host_location	27.53105	1.562119	17.62	0.000	24.46903	30.59308
_cons	41.87709	1.250379	33.49	0.000	39.42614	44.32805

Figura 309: Output Stata "Domanda e Location"

Il primo modello che si va a presentare riguarda la variabile "host_location", che indica se l'host è presente fisicamente, quindi se l'host è residente, nella città di Roma. Si osserva una correlazione positiva tra la variabile e la domanda, in quanto il coefficiente è positivo, il valore della t risulta al di fuori dell'intervallo, e soprattutto il p-value è pari a 0.

Quest'ultimo valore indica un'alta significatività della variabile `host_location`.

Quindi, alla luce di questi risultati, possiamo concludere che il fatto che l'host risiede nella città in cui affitta l'alloggio è un elemento importante che può aggiungere valore all'esperienza del soggiorno dell'ospite. Infatti, come già detto nei capitoli precedenti, la presenza fisica comporta sicuramente una maggiore disponibilità da parte dell'host, che può interagire direttamente con il guest, ed implica un valore aggiunto anche in termini della qualità dell'esperienza a livello sociale.

4.2.2 Domanda e lunghezza descrizione dell'host

Il seguente modello mostra la correlazione tra la domanda e la lunghezza della descrizione che l'host dà di sé, disponibile sul suo profilo.

```
. regress number_of_reviews host_about_length, robust
```

Linear regression		Number of obs =	11385
		F(1, 11383) =	29.60
		Prob > F	= 0.0000
		R-squared	= 0.0054
		Root MSE	= 84.536

number_of_reviews	Coef.	Robust Std. Err.	t	P> t	[95% Conf. Interval]	
host_about_length	.0766242	.0140849	5.44	0.000	.0490153	.1042332
_cons	60.38624	.955208	63.22	0.000	58.51387	62.25861

Figura 31: Output Stata "Domanda e Descrizione host"

La relazione tra le due variabili risulta essere positiva, ma in misura minore rispetto alla variabile presentata precedentemente: il coefficiente non raggiunge neanche l'unità e il valore della t risulta di poco al di fuori dell'intervallo di confidenza.

Dal valore del p-value pari a 0 si può dedurre che la lunghezza della descrizione personale dell'host è un elemento con un impatto significativo sulla domanda. Agli ospiti interessa sapere in anticipo qualche aspetto della figura dell'host con cui si andranno a relazionare; il basso coefficiente indica che lo sforzo di fornire una descrizione dettagliata di sé non trova, però, adeguata ricompensa.

4.2.3 Domanda e tempo di risposta dell'host

```
. regress number_of_reviews host_response_time, robust
```

Linear regression

Number of obs =	11385
F(1, 11383) =	492.02
Prob > F	= 0.0000
R-squared	= 0.0250
Root MSE	= 83.702

	Coef.	Robust Std. Err.	t	P> t	[95% Conf. Interval]	
number_of_reviews						
host_response_time	-16.97594	.7653181	-22.18	0.000	-18.4761	-15.47578
_cons	87.98274	1.551695	56.70	0.000	84.94115	91.02433

Figura 3210: Output Stata "Domanda e Tempo di risposta"

Il modello di regressione semplice relativo al tempo di risposta dell'host dà una conferma numerica di quanto ci si aspettava: il coefficiente di correlazione risulta negativo, e quindi più il tempo di risposta si allunga, più l'impatto sulla domanda è negativo. È normale pensare, infatti, che l'ospite non gradisce l'attesa prolungata e che invece apprezza avere

risposta nell'immediato. Il p-value è 0 a conferma dell'alta significatività della variabile.

4.2.4 Domanda e tasso di risposta dell'host, domanda e tasso di accettazione dell'host

Altri due fattori relativi al comportamento dell'host sono il tasso di risposta e il tasso di accettazione, riportati nelle figure sottostanti.

```
. regress number_of_reviews host_response_rate, robust
```

Linear regression

Number of obs = 11385
F(1, 11383) = 375.46
Prob > F = 0.0000
R-squared = 0.0161
Root MSE = 84.082

number_of_reviews	Coef.	Robust Std. Err.	t	P> t	[95% Conf. Interval]	
host_response_rate	66.23626	3.418349	19.38	0.000	59.53571	72.93681
_cons	1.4514	3.101692	0.47	0.640	-4.628451	7.531251

Figura 33: Output Stata "Domanda e Tasso di risposta"

```
. regress number_of_reviews host_acceptance_rate, robust
```

Linear regression

Number of obs = 11385
F(1, 11383) = 587.65
Prob > F = 0.0000
R-squared = 0.0294
Root MSE = 83.511

number_of_reviews	Coef.	Robust Std. Err.	t	P> t	[95% Conf. Interval]	
host_acceptance_rate	54.47244	2.24708	24.24	0.000	50.06778	58.87711
_cons	17.02181	1.814875	9.38	0.000	13.46434	20.57927

Figura 34: Output Stata "Domanda e Tasso d'accettazione"

In entrambi i casi è possibile notare una relazione positiva con il tasso di domanda, e rispetto ai modelli precedenti visti finora, i coefficienti di correlazione risultano più alti (circa 66 per il tasso di risposta, 54 per il tasso di accettazione).

Il valore della t non rientra nell'intervallo di confidenza ed il p-value pari a 0 conferma l'elevata importanza delle due variabili sotto esame.

4.2.5 Domanda e titolo di superhost

```
. regress number_of_reviews host_is_superhost, robust
```

Linear regression

Number of obs = 11385
F(1, 11383) = 628.10
Prob > F = 0.0000
R-squared = 0.0641
Root MSE = 82.007

number_of_reviews	Coef.	Robust Std. Err.	t	P> t	[95% Conf. Interval]	
host_is_superhost	45.12945	1.800723	25.06	0.000	41.59972	48.65918
_cons	48.21067	.8126673	59.32	0.000	46.61771	49.80364

Figura 35: Output Stata "Domanda e Superhost"

Un altro elemento che è risultato rilevante in analisi precedenti è il badge di superhost. L'importanza dello status di superhost è confermata anche in questo studio; infatti, la correlazione con la domanda è positiva e quindi si può affermare che un host che possiede tale titolo risulta essere maggiormente scelto dai turisti.

Essere superhost, agli occhi dei clienti, implica serietà e affidabilità, nonché indice di esperienza consolidata; è una sorta di garanzia per l'ospite che il soggiorno sarà confortevole senza andare incontro a troppi disagi.

Il p-value pari a 0 conferma l'importanza della variabile.

4.2.6 Domanda e numero di annunci dell'host

```
. regress number_of_reviews host_listings_count, robust
```

Linear regression	Number of obs = 11385					
	F(1, 11383) = 349.10					
	Prob > F = 0.0000					
	R-squared = 0.0076					
	Root MSE = 84.443					

number_of_reviews	Coef.	Robust Std. Err.	t	P> t	[95% Conf. Interval]	
host_listings_count	-.1901285	.0101758	-18.68	0.000	-.2100748	-.1701821
_cons	66.19031	.8417429	78.63	0.000	64.54035	67.84027

Figura 36: Output Stata "Domanda e Numero annunci"

Il modello riportato in figura rappresenta la correlazione tra la domanda e il numero degli annunci pubblicati dagli host, che è anche una misura di quanti alloggi hanno in proprietà. In questo caso, il coefficiente di correlazione risulta negativo, indicando che ad un aumentare del numero di annunci, la domanda diminuisce.

Se si era ipotizzato, infatti, che un numero maggiore di proprietà gestite poteva implicare anche una maggior esperienza, d'altro canto, dal modello sembra aver trovato veridicità l'assunzione che un numero maggiore comporta minor attenzione e disponibilità da parte dell'host. I clienti, dunque, preferiscono un host che fornisce un servizio di qualità in termini di tempo a disposizione, rapida risposta e immediata disponibilità; la gestione contemporanea di molteplici alloggi, potrebbe far venir meno questi aspetti.

Il valore del p-value pari a 0 indica alta significatività della variabile analizzata.

4.2.7 Domanda e foto profilo dell'host

Di seguito è mostrata la correlazione tra la variabile "host_has_profile_pic", che indica se il profilo dell'host è fornito anche di foto identificativa, e la domanda.

```
. regress number_of_reviews host_has_profile_pic, robust
```

Linear regression		Number of obs =	11385
		F(1, 11383) =	12.93
		Prob > F	= 0.0003
		R-squared	= 0.0002
		Root MSE	= 84.757

number_of_reviews	Coef.	Robust Std. Err.	t	P> t	[95% Conf. Interval]	
host_has_profile_pic	34.32927	9.546762	3.60	0.000	15.61597	53.04257
_cons	29.5	9.513574	3.10	0.002	10.85176	48.14824

Figura 37: Output Stata "Domanda e Foto profilo"

Il coefficiente di correlazione risulta positivo, ma il valore della t, rispetto agli altri modelli, risulta di poco al di fuori dell'intervallo di confidenza.

L'ospite, dunque, nel momento della scelta dell'alloggio tiene in considerazione aspetti come la descrizione che l'host dà di sé precedentemente analizzata, ma si può anche assumere che opera una certa discriminazione in base all'aspetto esteriore.

Anche in questo caso, il valore del p-value indica alta significatività della variabile in esame.

4.2.8 Domanda e identità verificata dell'host

```
. regress number_of_reviews host_identity_verified, robust
```

Linear regression		Number of obs =	11385
		F(1, 11383) =	979.49
		Prob > F	= 0.0000
		R-squared	= 0.0271
		Root MSE	= 83.608

number_of_reviews	Coef.	Robust Std. Err.	t	P> t	[95% Conf. Interval]	
host_identity_verified	43.63633	1.394277	31.30	0.000	40.9033	46.36935
_cons	25.20015	1.085725	23.21	0.000	23.07194	27.32836

Figura 11: Output Stata "Domanda e Profilo verificato"

Nel modello sopra riportato, è mostrata la correlazione tra la domanda e "host_identity_verified". I risultati attestano un'elevata importanza di quest'ultima variabile: il coefficiente di correlazione è positivo, la t non è compresa nell'intervallo e il p-value risulta pari a 0.

Ciò vuol dire che il profilo verificato è un fattore che l'ospite tiene in considerazione nel momento in cui opera delle scelte. La conferma dell'identità dell'host è indice di affidabilità e simbolo di maggior sicurezza e garanzia di rapportarsi con un host serio.

4.2.9 Domanda e lunghezza della descrizione dell'annuncio

```
. regress number_of_reviews description_length, robust
```

Linear regression

Number of obs = 11385
F(1, 11383) = 195.40
Prob > F = 0.0000
R-squared = 0.0163
Root MSE = 84.073

number_of_reviews	Robust		t	P> t	[95% Conf. Interval]	
	Coef.	Std. Err.				
description_length	.2294313	.0164129	13.98	0.000	.1972592	.2616035
_cons	34.87971	2.117328	16.47	0.000	30.72938	39.03004

Figura 3912: Output Stata "Domanda e Descrizione annuncio"

Il modello in figura riporta la correlazione tra la variabile dipendente e la lunghezza della descrizione dell'annuncio. Il coefficiente, anche se basso, risulta positivo, ed il p-value suggerisce alta significatività del regressore.

Rispetto alla descrizione personale dell'host, possiamo affermare che un ospite dà più importanza alle informazioni relative all'alloggio. Infatti, è logico pensare che la scelta sia guidata anche dalle caratteristiche stesse della stanza/appartamento, per capire se sia adeguato alle esigenze del viaggio. Quindi, più un host dedica del tempo a scrivere una presentazione dettagliata dell'alloggio, più riceverà attenzioni da parte dei turisti.

4.2.10 Domanda e distanza dell'alloggio

```
. regress number_of_reviews distance, robust
```

Linear regression		Number of obs =	11385
		F(1, 11383) =	400.41
		Prob > F	= 0.0000
		R-squared	= 0.0155
		Root MSE	= 84.106

number_of_~s	Coef.	Robust Std. Err.	t	P> t	[95% Conf. Interval]	
distance	-2.6861	.1342366	-20.01	0.000	-2.949227	-2.422973
_cons	73.64053	1.050747	70.08	0.000	71.58089	75.70018

Figura 40: Output Stata "Domanda e Distanza alloggio"

Come ci si poteva aspettare, la correlazione tra la domanda e la distanza dell'alloggio risulta essere negativa: tramite il modello si ha una conferma numerica del fatto che più un alloggio risulta essere distante dal centro della città e più sarà considerato una scelta meno popolare rispetto alle altre. Questo risultato non stupisce, in quanto, come affermato precedentemente, un posizionamento abbastanza centrale dell'appartamento in affitto comporta notevoli vantaggi in termini di vicinanza alle principali attrazioni turistiche, o semplicemente a servizi necessari come supermercati, trasporti...

Tendenzialmente, quindi, gli ospiti scelgono un alloggio che risulta posizionato strategicamente, in modo da avere un soggiorno il più confortevole possibile, senza per esempio, dover affrontare chilometri di strada, ore di viaggio, per poter completare le tappe del proprio itinerario.

Il p-value pari a 0 conferma l'alta significatività della variabile distanza.

4.2.11 Domanda e tipologia dell'alloggio

```
. regress number_of_reviews room_type, robust
```

Linear regression

Number of obs =	11385
F(1, 11383) =	170.91
Prob > F	= 0.0000
R-squared	= 0.0132
Root MSE	= 84.208

number_of_~s	Coef.	Robust Std. Err.	t	P> t	[95% Conf. Interval]	
room_type	22.05213	1.686808	13.07	0.000	18.7457	25.35857
_cons	25.49728	2.946719	8.65	0.000	19.72121	31.27336

Figura 41: Output Stata "Domanda e Tipologia alloggio"

Dal modello in figura risulta evidente come la tipologia dell'alloggio influisce positivamente sulla domanda: tendenzialmente sono preferiti interi appartamenti alle stanze singole.

Il coefficiente di correlazione è, infatti, positivo, il valore della t è al di fuori dell'intervallo di confidenza, e il p-value indica alta significatività della variabile "room_type".

Non sorprende che la soluzione dell'intero appartamento sia preferito alla singola stanza, in quanto, gli ospiti preferiscono avere a disposizione tutti i servizi legati ad una casa e l'aspetto privacy è anche importante, in quanto difficilmente piace condividere con altri estranei gli spazi.

4.2.12 Domanda e numero di persone ospitabili

```
. regress number_of_reviews accommodates, robust
```

Linear regression

Number of obs =	11385
F(1, 11383) =	4.34
Prob > F	= 0.0374
R-squared	= 0.0003
Root MSE	= 84.756

number_of_~s	Coef.	Robust Std. Err.	t	P> t	[95% Conf. Interval]	
accommodates	-.6357889	.3053522	-2.08	0.037	-1.234332	-.037246
_cons	66.31151	1.547801	42.84	0.000	63.27755	69.34546

Figura 42: Output Stata "Domanda e Capienza alloggio"

La correlazione tra la domanda e il numero di persone ospitabili risulta essere negativa. Si osserva poi un valore della t non eccessivamente distante dagli estremi dell'intervallo di confidenza ed un p-value pari a 3.7% indica che la variabile non è estremamente significativa.

Ciò potrebbe essere giustificato dal fatto che chi usufruisce del servizio offerto da Airbnb non cerca appartamenti per un gruppo molto numeroso di persone; piuttosto, è preferibile soggiornare in uno spazio adatto come dimensioni e non eccessivamente grande, anche perché generalmente si è alla ricerca di un alloggio in cui non manchi nulla ma che abbia l'essenziale per la durata del soggiorno.

4.2.13 Domanda e numero di bagni, domanda e numero di camere da letto, domanda e numero di letti

Nel seguente paragrafo sono riportate le correlazioni della variabile dipendente con le singole variabili relative al numero di bagni, di camere da letto e di letti. Si è scelto di riportarle tutte insieme nello stesso paragrafo in quanto i risultati ottenuti e le considerazioni in merito che si possono fare sono simili.

```
. regress number_of_reviews bathrooms, robust
```

Linear regression

Number of obs = 11385
F(1, 11383) = 58.10
Prob > F = 0.0000
R-squared = 0.0028
Root MSE = 84.647

number_of_~s	Coef.	Robust Std. Err.	t	P> t	[95% Conf. Interval]	
bathrooms	-6.618035	.8682698	-7.62	0.000	-8.319994	-4.916077
_cons	72.508	1.53035	47.38	0.000	69.50825	75.50775

Figura 43: Output Stata "Domanda e Numero bagni"

```
. regress number_of_reviews bedrooms, robust
```

Linear regression

Number of obs = 11385
F(1, 11383) = 32.46
Prob > F = 0.0000
R-squared = 0.0019
Root MSE = 84.686

number_of_~s	Coef.	Robust Std. Err.	t	P> t	[95% Conf. Interval]	
bedrooms	-4.343773	.7623611	-5.70	0.000	-5.838132	-2.849413
_cons	70.40523	1.533472	45.91	0.000	67.39936	73.4111

Figura 44: Output Stata "Domanda e Numero stanze"

<code>. regress number_of_reviews beds, robust</code>						
Linear regression			Number of obs = 11385			
			F(1, 11383) = 22.75			
			Prob > F = 0.0000			
			R-squared = 0.0013			
			Root MSE = 84.711			
number_of_~s	Coef.	Robust Std. Err.	t	P> t	[95% Conf. Interval]	
beds	-1.851537	.3882166	-4.77	0.000	-2.612508	-1.090565
_cons	68.17284	1.308241	52.11	0.000	65.60846	70.73721

Figura 45: Output Stata "Domanda e Numero di letti"

La prima cosa che si nota è il coefficiente di correlazione negativo: tutte e tre i regressori considerati hanno una relazione negativa con la domanda. Ciò vuol dire che più aumenta il numero di stanze, letti e bagni, più la domanda risulta minore.

Anche questi risultati portano alla considerazione precedente che i dati sono relativi a guest che cercano alloggio per un numero limitato di persone, quindi per gruppi poco numerosi. In tal caso, è preferito maggiormente un appartamento bilanciato: un eccessivo numero a disposizione di letti/bagni potrebbe far somigliare l'alloggio ad un dormitorio.

Il valore del p-value pari a 0 indica significatività delle tre variabili in oggetto, e la variabile con maggiore coefficiente in valore assoluto risulta essere quella relativa al numero di bagni.

4.2.14 Domanda e prezzo

```
. regress number_of_reviews price, robust
```

Linear regression		Number of obs =	11385
		F(1, 11383) =	90.89
		Prob > F	= 0.0000
		R-squared	= 0.0054
		Root MSE	= 84.536

number_of_~s	Coef.	Robust Std. Err.	t	P> t	[95% Conf. Interval]	
price	-.0730819	.0076657	-9.53	0.000	-.088108	-.0580559
_cons	70.4727	1.158624	60.82	0.000	68.20159	72.7438

Figura 46: Output Stata "Domanda e Prezzo"

Un altro risultato atteso è confermato numericamente dal modello di regressione lineare in figura: il prezzo ha una correlazione negativa con la domanda. Più, infatti, il prezzo sale, minore sarà la richiesta per quella determinata stanza o per quel determinato appartamento.

Il coefficiente di correlazione è negativo ed il p-value mostra che la variabile ha elevata importanza statistica; il prezzo è sicuramente un fattore che gli ospiti prendono in considerazione nella pianificazione e nella scelta del loro soggiorno.

4.2.15 Domanda e rating complessivo

```
. regress number_of_reviews review_scores_rating, robust
```

Linear regression	Number of obs = 11385					
	F(1, 11383) = 1362.55					
	Prob > F = 0.0000					
	R-squared = 0.0978					
	Root MSE = 80.514					

number_of_reviews	Coef.	Robust Std. Err.	t	P> t	[95% Conf. Interval]	
review_scores_rating	12.86151	.3484298	36.91	0.000	12.17853	13.54449
_cons	19.50399	1.217836	16.02	0.000	17.11682	21.89116

Figura 47: Output Stata "Domanda e Rating"

Il modello in figura mostra la correlazione tra la variabile dipendente in esame e il rating complessivo. La relazione risulta positiva, con un coefficiente di circa 12.9; il p-value pari a 0 indica alta significatività della variabile.

Come ipotizzato, il rating gioca un ruolo essenziale: è un fattore che sicuramente viene preso in considerazione ed influenza le scelte di alloggio dell'ospite, in quanto, essendo originato dalle valutazioni lasciate dai precedenti guest, rispecchia maggiormente la qualità del servizio e dà un'idea dell'esperienza di soggiorno offerta dall'host.

Avendo spiegato che il rating complessivo è una media di altre 6 tipologie di aspetti valutati, è interessante notare quali tra questi risulti essere più influente sulla domanda.

Dalla seguente figura, è confermato ovviamente quanto detto sul rating generale: tutte le variabili risultano correlate positivamente alla domanda e con alta significatività.

Linear regression

Number of obs = 11385
 F(6, 11378) = 871.72
 Prob > F = 0.0000
 R-squared = 0.2196
 Root MSE = 74.898

number_of_reviews	Coef.	Robust Std. Err.	t	P> t	[95% Conf. Interval]	
review_scores_accuracy	3.170176	.4548979	6.97	0.000	2.278498	4.061854
review_scores_cleanliness	2.814442	.4729065	5.95	0.000	1.887464	3.74142
review_scores_checkin	5.68375	.4505869	12.61	0.000	4.800522	6.566978
review_scores_communication	6.308821	.488047	12.93	0.000	5.352165	7.265478
review_scores_location	5.628304	.4238421	13.28	0.000	4.7975	6.459107
review_scores_value	1.052009	.4431056	2.37	0.018	.1834453	1.920572
_cons	-19.22771	.7306064	-26.32	0.000	-20.65982	-17.79559

Figura 48: Output Stata "Domanda e Diversi rating"

Si nota che il coefficiente di correlazione più alto è quello relativo alla comunicazione con l'host, seguito dal punteggio sul checkin e sulla location. Questa è un'ulteriore conferma di quanto sostenuto finora in merito alla disponibilità e alla figura dell'host. Per gli ospiti è essenziale avere un'ottima comunicazione con l'host, punto di riferimento per qualsiasi tipo di inconveniente.

4.2.16 Domanda e prenotazione immediata

```
. regress number_of_reviews instant_bookable, robust
```

Linear regression

Number of obs = 11385
 F(1, 11383) = 284.31
 Prob > F = 0.0000
 R-squared = 0.0214
 Root MSE = 83.856

number_of_reviews	Coef.	Robust Std. Err.	t	P> t	[95% Conf. Interval]	
instant_bookable	25.56728	1.516313	16.86	0.000	22.59505	28.53952
_cons	47.86127	1.056409	45.31	0.000	45.79052	49.93201

Figura 49: Output stata "Domanda e Prenotazione immediata"

Il modello di regressione lineare semplice con regressore "instant_bookable", suggerisce l'importanza di permettere tale opzione per i clienti. Se, infatti, l'host decide di consentire all'ospite la prenotazione immediata, acquista un vantaggio rispetto agli altri. Chi cerca alloggio, in generale, non gradisce aspettare e preferisce poter procedere con le questioni relative alla prenotazione in maniera veloce e semplice.

Il coefficiente di correlazione con la domanda è positivo ed il p-value indica elevata importanza statistica della variabile di questo paragrafo in esame.

4.3 MODELLI DI REGRESSIONE LINEARE MULTIPLA PER TIPOLOGIA DI VARIABILI

Essendo abbastanza chiara e netta la suddivisione per tipologia delle variabili utilizzate nei modelli, si è deciso di sviluppare due modelli intermedi che analizzano quali degli aspetti riferiti strettamente alla figura dell'host e quali relativi all'alloggio risultano essere più influenti sulla domanda.

4.3.1 Variabili relative all'host

Il primo modello intermedio che si va ad analizzare è quello relativo alle caratteristiche dell'host. Le variabili incluse sono: la location dell'host, la lunghezza della descrizione dell'host, il tempo di risposta, il tasso di risposta, il tasso di accettazione, il titolo di superhost, il numero di annunci, la foto profilo e il profilo verificato.

Linear regression

Number of obs = 11385
F(9, 11375) = 203.85
Prob > F = 0.0000
R-squared = 0.1078
Root MSE = 80.095

number_of_reviews	Coef.	Robust Std. Err.	t	P> t	[95% Conf. Interval]	
host_location	17.0943	1.59022	10.75	0.000	13.9772	20.21141
host_about_length	.0403216	.0110478	3.65	0.000	.0186659	.0619773
host_response_time	-5.414067	1.123757	-4.82	0.000	-7.616824	-3.211309
host_response_rate	-2.946717	4.807688	-0.61	0.540	-12.37061	6.477181
host_acceptance_rate	29.54561	2.399956	12.31	0.000	24.84128	34.24994
host_is_superhost	33.19573	1.859756	17.85	0.000	29.55028	36.84117
host_listings_count	-.1453594	.0120805	-12.03	0.000	-.1690392	-.1216796
host_has_profile_pic	18.59699	10.22	1.82	0.069	-1.435974	38.62995
host_identity_verified	28.98882	1.388271	20.88	0.000	26.26757	31.71007
_cons	-20.28855	11.73885	-1.73	0.084	-43.29873	2.721632

Figura 50: Output Stata "Domanda e Variabili host"

In base ai risultati ottenuti dal modello è possibile notare una conferma della significatività delle variabili analizzate singolarmente nel precedente sotto capitolo, ad eccezione del tasso di risposta dell'host, che presenta adesso un valore del p-value pari al 54%, e della foto profilo, il cui valore del p-value è passato da 0 a 6.9%.

È interessante notare che il coefficiente di correlazione più alto in valore assoluto risulta essere quello relativo alla variabile superhost, pari a circa 33. Vediamo, invece, che tra le variabili significative quella con minore impatto risulta essere la descrizione dell'host. Anche l'influenza del tempo di risposta non è particolarmente elevata e risulta minore rispetto a quella del tasso di accettazione.

Quindi, secondo questo modello relativo alle caratteristiche dell'host, ciò su cui bisogna puntare per avere un'affluenza maggiore di clienti sono il titolo di superhost, il tasso di accettazione, l'avere un profilo verificato e l'essere presenti fisicamente nella città.

4.3.2 Variabili relative all'alloggio

Il secondo modello comprende le variabili direttamente associabili all'alloggio: la descrizione dell'annuncio, la distanza dell'alloggio dal centro della città, la tipologia dell'alloggio, il numero di persone ospitabili, il numero di bagni, il numero di camere da letto, il numero di letti, il prezzo, il rating complessivo e l'opzione di prenotazione immediata.

Linear regression					Number of obs = 11385	
					F(10, 11374) = 227.13	
					Prob > F = 0.0000	
					R-squared = 0.1432	
					Root MSE = 78.495	
number_of_reviews	Coef.	Robust Std. Err.	t	P> t	[95% Conf. Interval]	
description_length	.1462131	.0157925	9.26	0.000	.1152572	.177169
distance	-2.210718	.1281616	-17.25	0.000	-2.461937	-1.9595
room_type	18.88781	1.908097	9.90	0.000	15.14761	22.62801
accommodates	-1.356989	.707891	-1.92	0.055	-2.744578	.0305993
bathrooms	-.729346	1.508676	-0.48	0.629	-3.686612	2.22792
bedrooms	-1.200595	1.495287	-0.80	0.422	-4.131615	1.730425
beds	.0052803	.7085257	0.01	0.994	-1.383552	1.394113
price	-.0886251	.0087366	-10.14	0.000	-.1057502	-.0714999
review_scores_rating	11.29421	.3488506	32.38	0.000	10.6104	11.97801
instant_bookable	17.75637	1.452038	12.23	0.000	14.91012	20.60261
_cons	-12.95481	3.171945	-4.08	0.000	-19.17236	-6.737246

Figura 51: Output Stata "Domanda e Variabili alloggio"

Dall'output del modello si nota come la presenza delle altre variabili influisca e cambi i valori dei risultati precedentemente ottenuti.

In questo caso, variabili come il numero di bagni, le camere da letto e il numero di letti non risultano essere più significative come nei modelli di regressione lineare semplice.

Il numero di letti presenta, infatti, un valore della t compreso nell'intervallo di confidenza, ma soprattutto un valore del p-value pari al

99.4%; il numero di bagni non risulta essere più influente con un p-value di circa 63 punti percentuali, così come il numero delle stanze da letto presenta ora un p-value pari a 42.2 %.

Le restanti variabili si confermano altamente significative con un p-value pari a 0. In particolare, maggior influenza risulta essere esercitata dalle variabili `room_type` e `instant_bookable`, con dei coefficienti di correlazione prossimi a 20.

4.4 MODELLO FINALE

Di seguito è presentato il modello finale, comprendente tutte le variabili esaminate finora nei modelli precedenti.

Linear regression		Number of obs = 11385				
		F(19, 11365) = 162.49				
		Prob > F = 0.0000				
		R-squared = 0.1965				
		Root MSE = 76.043				
number_of_reviews	Coef.	Robust Std. Err.	t	P> t	[95% Conf. Interval]	
host_location	13.09258	1.499135	8.73	0.000	10.15401	16.03114
host_about_length	.0332799	.0106836	3.12	0.002	.0123382	.0542215
host_response_time	-2.180229	1.087491	-2.00	0.045	-4.3119	-.0485585
host_response_rate	-3.922678	4.54816	-0.86	0.388	-12.83786	4.992501
host_acceptance_rate	12.32882	2.488168	4.95	0.000	7.451585	17.20606
host_is_superhost	27.4568	1.779189	15.43	0.000	23.96928	30.94431
host_listings_count	-.1469306	.0126463	-11.62	0.000	-.1717195	-.1221416
host_has_profile_pic	10.48514	10.00055	1.05	0.294	-9.117666	30.08795
host_identity_verified	21.33386	1.333966	15.99	0.000	18.71906	23.94867
description_length	.1127016	.0155828	7.23	0.000	.0821567	.1432465
distance	-1.911764	.1255398	-15.23	0.000	-2.157843	-1.665684
room_type	16.05744	1.867193	8.60	0.000	12.39742	19.71746
accommodates	-1.68477	.6893385	-2.44	0.015	-3.035992	-.3335474
bathrooms	-3.039101	1.495687	-2.03	0.042	-5.970906	-.1072972
bedrooms	-.6200893	1.461214	-0.42	0.671	-3.484322	2.244143
beds	.7088058	.6935325	1.02	0.307	-.6506377	2.068249
price	-.0678767	.0094476	-7.18	0.000	-.0863957	-.0493578
review_scores_rating	9.759822	.3414785	28.58	0.000	9.090466	10.42918
instant_bookable	10.98481	1.657375	6.63	0.000	7.736071	14.23355
_cons	-48.14012	11.55594	-4.17	0.000	-70.79175	-25.48849

Figura 132: Output Stata Modello Finale

È interessante notare come i risultati ottenuti nei modelli intermedi sono confermati anche in questo caso, e tutte le variabili presentano lo stesso segno di correlazione con la domanda.

Facendo un confronto dei coefficienti di correlazione in valore assoluto, il più rilevante risulta essere quello associato alla variabile superhost, con un valore pari a circa 27.45; segue la verifica del profilo dell'host con un coefficiente di 21.33 e la tipologia della stanza con 16.06.

Gli aspetti legati alle caratteristiche dell'alloggio, ovvero il numero di stanze, il numero di letti e il numero di bagni risultano ancora con un p-value lontano dallo 0. Rispetto al modello intermedio presentato precedentemente, l'influenza di tutte le variabili fa sì che l'aspetto meno influente sulla domanda non risulti più essere il numero di bagni (che è passato da un p-value pari a 99.4% a 30.7%), ma il numero di camere da letto, con un p-valore del 67.1%.

In relazione alla numerosità dei bagni vediamo, invece, un miglioramento consistente: si passa da quasi 63 punti percentuali a poco più di 4. Quindi, col modello generale, acquista nuovamente significatività il numero di bagni a dispetto del numero di camere. Ciò può trovare spiegazione nel fatto che tendenzialmente si hanno più camere a cui sono associate meno bagni, e dunque l'ospite cerca il giusto equilibrio in base al numero di persone che devono soggiornare.

Anche il p-value della numerosità massima di persone ospitabili risulta ancora diverso da 0, anche se con un valore più basso rispetto al caso precedente. Possiamo comunque confermare le considerazioni fatte finora.

Conferme dei modelli precedenti si hanno anche in relazione alle seguenti variabili riferite alla figura dell'host:

- Tempo di risposta: la relazione con la domanda continua ad essere negativa, in quanto ad un tempo maggiore corrisponde minor richiesta, e presenta adesso un p-value pari a 4.5%. Quindi si può affermare che il tempo di risposta è un fattore da tenere in considerazione;
- Tasso di risposta: il coefficiente di correlazione è negativo, a conferma che, com'è logico pensare, l'host che risponde meno ha di conseguenza minor domanda. Il p-value risulta di poco più basso rispetto al caso del modello intermedio, ma comunque possiamo affermare che il tasso di risposta non ha elevata rilevanza;

- Foto del profilo: è l'unica tra le variabili riferite alla figura dell'host che ha subito un incremento in termini di p-value, che risulta essere pari a circa il 30%. Quindi, tale variabile non influisce significativamente sulla domanda.

Tutte le rimanenti variabili presentano un p-value pari a 0, e quindi sono altamente significative, anche se in misura minore rispetto a quelle discusse finora.

4.4.1 Modello con variabile dipendente "review_scores_rating"

Essendo il rating complessivo ricavato come media di valutazioni di altri aspetti (come la location, la precisione e la comunicazione con l'host...), si è deciso di indagare anche su di esso l'influenza delle variabili considerate.

In figura è dunque riportato il modello di regressione lineare multipla relativo al rating:

Linear regression

Number of obs = 11385
F(18, 11366) = 44.48
Prob > F = 0.0000
R-squared = 0.0662
Root MSE = 1.9934

review_scores_rating	Coef.	Robust Std. Err.	t	P> t	[95% Conf. Interval]	
host_location	.2433517	.0499277	4.87	0.000	.1454847	.3412186
host_about_length	-.0001521	.0002659	-0.57	0.567	-.0006733	.0003691
host_response_time	-.1435976	.0389529	-3.69	0.000	-.219952	-.0672432
host_response_rate	.2900733	.1719724	1.69	0.092	-.0470223	.6271689
host_acceptance_rate	.3989771	.0947243	4.21	0.000	.2133012	.584653
host_is_superhost	.3521423	.0412187	8.54	0.000	.2713465	.4329381
host_listings_count	-.0019206	.0005595	-3.43	0.001	-.0030173	-.0008239
host_has_profile_pic	-.0051165	.4650653	-0.01	0.991	-.9167247	.9064917
host_identity_verified	.3102829	.0637736	4.87	0.000	.1852757	.4352901
description_length	.0041998	.0004182	10.04	0.000	.00338	.0050195
distance	-.0179215	.0049684	-3.61	0.000	-.0276605	-.0081825
room_type	.4018954	.0505396	7.95	0.000	.302829	.5009619
accommodates	.0033026	.0196195	0.17	0.866	-.035155	.0417603
bathrooms	-.025235	.0443413	-0.57	0.569	-.1121515	.0616816
bedrooms	.0082358	.0433259	0.19	0.849	-.0766904	.093162
beds	-.0232434	.0213097	-1.09	0.275	-.0650141	.0185274
price	-.001216	.0002866	-4.24	0.000	-.0017777	-.0006543
instant_bookable	.0647435	.0462405	1.40	0.161	-.0258958	.1553829
_cons	1.451373	.5154669	2.82	0.005	.440969	2.461777

Figura 53: Output Stata Rating

Rispetto al caso del tasso di domanda, vediamo dei risultati simili in termini di correlazione con le variabili. È confermata la non significatività degli aspetti legati alla capienza dell'alloggio, alla numerosità delle camere, dei letti e dei bagni, che riportano rispettivamente i seguenti p-value: 86.6%, 84.9%, 27.5% e 56.9%.

Anche per il rating, risulta totalmente irrilevante la variabile "host_has_profile_pic", con un valore del p-value pari a 99.1%; ciò non stupisce anche per il fatto che il campione analizzato di host risultava avere quasi nella totalità dei profili muniti di foto. Il coefficiente negativo indica, però, che ai profili senza foto corrisponde un rating più basso.

È interessante notare come la lunghezza della descrizione dell'host non risulti essere influente sul rating; anzi, il coefficiente è negativo e ciò porta a pensare che gli ospiti non sono interessati molto alla persona dell'host, ma più alla descrizione dell'appartamento in cui alloggeranno. Quindi, dedicare del tempo alla creazione di una presentazione personale accurata non porta alcun vantaggio all'host.

Un'altra variabile che perde significatività in relazione al rating è l'opzione di prenotazione immediata. In questo modello, anche se di poco, il valore della t risulta all'interno dell'intervallo di confidenza ed il p-value è di circa il 16%. Questo significa che l'"instant_bookable" non ha correlazione con il rating complessivo.

CAPITOLO 5

CONCLUSIONI

In questo capitolo conclusivo si andrà a presentare un riassunto dei risultati e delle evidenze ottenute dai modelli costruiti in questa tesi.

Lo scopo di questo lavoro era quello di capire quali variabili relative all'alloggio e alla figura dell'host influenzano e risultano essere rilevanti nelle scelte dei consumatori. Sono stati identificati diversi fattori ritenuti di successo da tenere in considerazione e quindi da inserire in una strategia tale da massimizzare la domanda di un alloggio.

Nonostante i risultati ottenuti, è bene precisare che vi è una limitazione in questo studio, ovvero che la variabile utilizzata come proxy del tasso di domanda è il numero di recensioni per alloggio. Questa è una variabile aggregata, somma totale delle recensioni ricevute dal momento dell'inserimento dell'annuncio, ed è quindi anche una misura anche della presenza temporale dell'alloggio sulla piattaforma: è ovvio che annunci presenti da più tempo hanno delle performance più consolidate e robuste rispetto a listing attivi, per esempio, da pochi mesi. Proprio per questa ragione, per analisi future sarebbe più indicato considerare una variabile più accurata che non presenti questo limite.

Di seguito, sono riportate le principali variabili a cui, nel modello di regressione finale, come confermato anche dai modelli intermedi, sono stati associati dei valori di correlazione più alti e significativi rispetto al tasso di domanda.

Il primo punto su cui un host deve puntare per avere un incremento della domanda è l'ottenere e mantenere il titolo di superhost. La correlazione

di questa variabile con la domanda infatti risulta essere positiva e il coefficiente di correlazione è in generale il più alto tra tutti quelli analizzati, pari a 27.45 circa. Come spiegato nel capitolo 1, per ottenere il badge di superhost vi sono diversi requisiti da rispettare, tra cui quello di limitare il più possibile il numero delle prenotazioni annullate e cercare di rispondere tempestivamente alle richieste. Non è facile avere tale riconoscimento, ma è dimostrato che è uno dei fattori che gli utenti considerano nelle scelte relative al loro soggiorno.

La seconda variabile che risulta avere molta importanza dal modello di regressione finale è quella dell'identità verificata dell'host. Questo è un risultato che stupisce, in quanto gli studi affrontati negli articoli presi come riferimento nel capitolo dell'analisi della letteratura, e che hanno analizzato anche questa variabile, sono giunti alla conclusione che il profilo verificato non è un fattore che influisce sulla domanda.

In questa analisi, invece, è stato ottenuto il risultato opposto: il coefficiente di correlazione con il tasso di domanda è positivo con un valore di 21.33.

Possiamo concludere quindi, che il fatto che si abbia una verifica e una conferma del profilo operata dallo stesso sistema della piattaforma, dà una maggiore sicurezza agli ospiti dell'autenticità del profilo e degli annunci che stanno considerando.

In relazione alle caratteristiche strutturali dell'alloggio stesso, l'unico aspetto risultato rilevante è la tipologia della stanza. È dimostrato, in accordo anche con i precedenti studi, che tendenzialmente sono preferiti interi appartamenti alle singole stanze. Questo non stupisce perché, sia per un aspetto di privacy, che di comodità, è normale prediligere una soluzione che comprende sia camere da letto che bagni, magari anche una cucina, piuttosto che affittare una singola stanza e dover condividere con altri il resto dei servizi.

Con un p-value pari a circa il 67% del tutto irrilevante risulta il numero di camere da letto, ed anche se per poco, numero di bagni e numero di letti non sono influenti nella scelta dell'alloggio.

Per quanto riguarda altre variabili legate alla figura dell'host, con dei coefficienti minori rispetto a quelli discussi finora nel capitolo, risultano comunque rilevanti la location e il tasso di accettazione, con valori rispettivamente di 13.09 e 12.32. In maniera concorde alle precedenti analisi, il tasso di risposta non risulta influente, mentre il tempo di risposta in questo caso, con un p-value pari a 4.5% e con un valore della t pari a -2, risulta per poco al di fuori dell'intervallo di confidenza e quindi ininfluenza sulla domanda.

Altre variabili su cui un host dovrebbe puntare, risultate influenti anche se in misura minore rispetto alle precedenti sono: la lunghezza della descrizione dell'annuncio, la distanza dell'alloggio dal centro della città, il prezzo per notte, la prenotazione immediata ed il rating complessivo.

Per concludere, è interessante notare come anche il modello realizzato alla fine del capitolo 4, con variabile dipendente il rating complessivo, ha prodotto risultati simili rispetto al modello finale. Questo dà un'ulteriore conferma dell'importanza di alcuni aspetti su cui quindi un host dovrebbe puntare, in quanto anche il rating può essere considerato come linea guida dei fattori di successo, perché una misura diretta dell'apprezzamento del servizio offerto dall'host.

BIBLIOGRAFIA

Heejeong Han, Seunghun Shin, Namho Chung, Chulmo Koo, (2019) "Which appeals (ethos, pathos, logos) are the most important for Airbnb users to booking?", *International Journal of Contemporary Hospitality Management*, <https://doi.org/10.1108/IJCHM-12-2017-0784>

Jiang Wu, Panhao Ma, Karen L. Xie, (2017) "In sharing economy we trust: the effects of host attributes on short-term rental purchases", *International Journal of Contemporary Hospitality Management*, <https://doi.org/10.1108/IJCHM-08-2016-0480>

Tahir M. Nisar, Nick Hajli, Guru Prabhakar, Yogesh Dwivedi, (2019) "Sharing economy and the lodging websites: Antecedents and mediators of accommodation purchase intentions", *Information Technology & People*, <https://doi.org/10.1108/ITP-06-2018-0297>

Jan Van der Borg, Nicola Camatti, Dario Bertocchi, and Andrea Albarea (2017) "The Rise of the Sharing Economy in Tourism: Exploring Airbnb Attributes for the Veneto Region"

Chrysa Agapitou, Anna Liana, Dimitrios Folinas, and Aggeliki Konstantoglou, (2020) "Airbnb Is Customers' Choice: Empirical Findings from a Survey", doi:10.3390/su12156136

Ulrich Gunter, (2017) "What makes an Airbnb host a superhost? Empirical evidence from San Francisco and the Bay Area", <https://doi.org/10.1016/j.tourman.2017.11.003>

Karen L. Xie, Zhenxing Mao, (2017) "The impacts of quality and quantity attributes of Airbnb hosts on listing performance", *International Journal of Contemporary Hospitality Management*, <https://doi.org/10.1108/IJCHM-07-2016-0345>

Mingming Chenga, Xin Jin (2017) "What do Airbnb users care about? An analysis of online review comments", <https://doi.org/10.1016/j.ijhm.2018.04.004>

Efstathios Kirkos, (2022) "Airbnb listings' performance: determinants and predictive models"

Georgina Görög, (2018) "The Definitions of Sharing Economy: A Systematic Literature Review", DOI:10.26493/1854-4231.13.175-189

Rachel Botsman e Roo Rogers, (2010) "What's mine is yours: The Rise of Collaborative Consumption"

Gutierrez, (2007) "The eruption of Airbnb in tourist cities: Comparing spatial patterns of hotels and peer to peer accommodation in Barcelona"

SITOGRAFIA

<http://insideairbnb.com/>

<https://bmtoolbox.net/stories/airbnb/>

<https://www.insidemarketing.it/glossario/definizione/customer-satisfaction/>

RINGRAZIAMENTI

Ormai al termine di questo mio percorso universitario al Politecnico di Torino, vorrei dedicare poche righe a chi mi è stato accanto e mi ha sostenuto per tutto questo tempo.

Vorrei innanzitutto ringraziare la professoressa Laura Abrardi per avermi seguita nella realizzazione di questa tesi ed essersi dimostrata sempre disponibile, con risposte tempestive alle mie domande.

Un ringraziamento particolare va ai miei genitori che mi hanno da sempre permesso di seguire la strada che desideravo, non facendomi mancare mai nulla, e pazientando fino a questo giorno comprendendo le difficoltà affrontate durante questo percorso.

Ringrazio le mie sorelle Alessia e Pamela, con cui tuttora condivido lo stesso tetto a Torino. Grazie per aver sopportato le mie ansie e il mio umore ballerino che peggiorava all'avvicinarsi delle sessioni d'esame.

Ringrazio il mio fantastico "Truput team", le mie amiche Michela, Sara, Elena e Valentina. Ci siamo conosciute tra i banchi universitari, abbiamo condiviso ansie e frustrazione, lavorato insieme a diversi progetti, e nessuno meglio di voi comprende l'impegno e la volontà necessari per affrontare questo percorso. Grazie per essere tuttora presenti nella mia vita.

Infine, ringrazio tutti i miei familiari e amici che, anche con una semplice chiamata o un semplice messaggio, da lontano mi hanno sempre incoraggiata e sostenuta.