



**Politecnico
di Torino**

POLITECNICO DI TORINO

Corso di Laurea Magistrale in Ingegneria Gestionale

A.A. 2024/2025

Sessione di Laurea Luglio 2025

Studio empirico sulle performance di Airbnb: analisi semantica applicata al caso italiano

Relatore: Elisabetta Raguseo

Candidata: Assunta Parziale

Correlatore: Francesco Luigi Milone

Matricola: 317495

*Ai miei genitori,
fonte di ispirazione e immenso supporto,
non avrei potuto desiderare di meglio,
grazie di tutto.*

ABSTRACT

Il presente studio si pone l'obiettivo di analizzare il legame tra le caratteristiche semantiche dei titoli degli annunci Airbnb in Italia, in un arco temporale compreso tra il 2019 ed il 2024, e le rispettive performance. Le variabili prestazionali adoperate sono "ADR" che rappresenta il rapporto tra i ricavi espressi in dollari ed il numero di giorni prenotati ed "revPAN" che rappresenta il rapporto tra i ricavi espressi in dollari ed il numero totale di giorni disponibili (intesi come la somma tra il numero di giorni prenotati ed il restante numero di giorni disponibili). Le caratteristiche semantiche dei titoli sono state studiate mediante l'individuazione dei temi principalmente discussi all'interno dei titoli stessi. L'individuazione dei topic maggiormente trattati all'interno dei titoli degli annunci Airbnb, è avvenuta tramite l'utilizzo di un algoritmo che combina modelli di K-Means ed SBERT tramite l'utilizzo del software "Python". I topic individuati sono stati posti, poi, in relazione con le variabili prestazionali, mediante una funzione di regressione basata sul modello OLS, implementata sul software "Stata". I risultati ottenuti dall'analisi di regressione mettono in risalto le tematiche che impattano in modo significativo le variabili prestazionali e gli effetti positivi del trattare la definizione del titolo in modo congiunto alla realizzazione dell'annuncio nel suo complesso, dunque, mediante l'applicazione di un approccio olistico.

TABLE OF CONTENTS

Capitolo 1 – Introduzione	1
1.1 Introduzione agli obiettivi.....	1
1.2 Sharing economy.....	1
1.3 Airbnb	9
Capitolo 2 – Analisi della letteratura	17
Capitolo 3 – Analisi dei dati	21
3.1 Descrizione dataset	21
3.2 Metodologia di analisi.....	35
3.5 Sentiment Analysis	40
3.6 Analisi di regressione.....	44
Capitolo 4 – Discussione dei risultati	98
Capitolo 5 – Impatti pratici, limitazioni e ricerca futura.....	104
Capitolo 6 - Conclusioni	108
Riferimenti	113

Capitolo 1 – Introduzione

1.1 Introduzione agli obiettivi

La presente tesi si pone l'obiettivo di collocarsi come elemento di congiunzione con gli studi esistenti. In particolare, si vuole approfondire la comprensione della relazione tra le caratteristiche semantiche dei titoli degli annunci Airbnb e le rispettive variabili prestazionali. Al fine di raggiungere gli scopi preposti, la presente tesi si articola come segue: viene, innanzitutto, realizzata una panoramica generale in merito al modello economico di sharing economy ed al suo player principale, ovvero Airbnb. Si prosegue con l'analisi della letteratura, grazie alla quale è stato individuato il gap, in termini di studi scientifici, da colmare. L'analisi della letteratura è seguita da un capitolo dedicato alle metodologie di studio utilizzate e alle analisi numeriche, che rappresentano il mezzo attraverso il quale è stato possibile studiare il legame tra le caratteristiche semantiche e le variabili prestazionali. Vengono dapprima analizzate le caratteristiche generali del dataset utilizzato in questo studio, dopodiché sono stati individuati i principali argomenti trattati all'interno del campione dei titoli degli annunci; rispetto a tali topic sono state realizzate differenti funzioni di regressione che li collegano alle variabili prestazionali. È stata realizzata, inoltre, una breve sezione in merito alla sentiment analysis. Il lavoro di tesi si conclude con la discussione dei risultati ottenuti, gli impatti pratici, le limitazioni, le possibili ricerche future e, infine, le conclusioni.

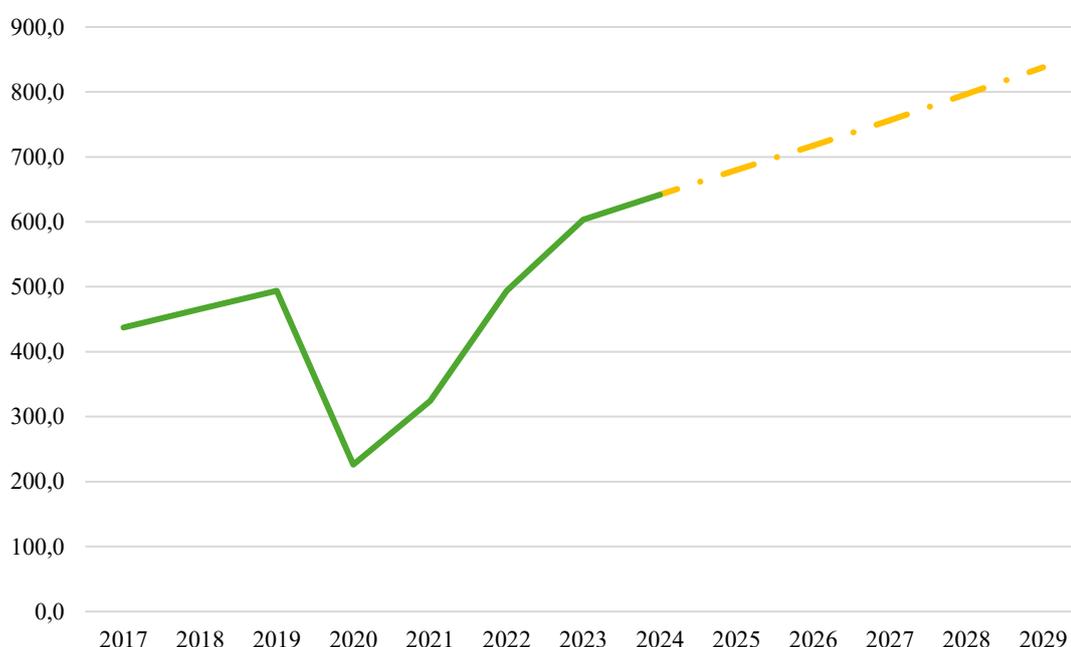
1.2 Sharing economy

La sharing economy è un modello economico estremamente interessante ed oggetto di molteplici studi. Non esiste una definizione univoca di sharing economy, dunque, per poterla caratterizzare è possibile far riferimento a un corpus di definizioni tra le quali emergono quella della Commissione Europea che la definisce come “*modelli imprenditoriali in cui le attività sono facilitate da piattaforme di collaborazione che creano un mercato aperto per l'uso temporaneo di beni o servizi spesso forniti da privati*” ([Commissione Europea, 2016, p. 3](#)) o ancora quella dell'Oxford Dictionary che la definisce come “*an economic system in which people can share possessions, services, etc., usually by means of the internet*” ([Oxford University Press. \(n.d.\)](#)). Si può notare

che, sebbene le definizioni siano differenti, sono presenti elementi comuni come il concetto di collaborazione e l'utilizzo di piattaforme che caratterizzano l'economia collaborativa stessa.

La sharing economy si è insidiata in diversi settori, come il settore del trasporto (di cui uno dei principali esponenti è Uber) ed il settore del turismo. Il settore del turismo è un settore in crescita, come mostrato dalla figura 1 ([Statista, 2024](#)) che rappresenta la capitalizzazione di mercato dei viaggi prenotati online.

Figura 1 Capitalizzazione di mercato dei viaggi prenotati online a livello mondiale (\$ Mld)



Fonte: Elaborazione dell'autore sui dati Statista, 2024

Dalla figura 1 ([Statista, 2024](#)) si evince un andamento crescente nel tempo, con l'unica eccezione nel 2020 dovuta probabilmente alla pandemia da Covid-19, ma seguita da una rapida ripresa; Statista ipotizza che tale ripresa continuerà nel tempo stimando una continua crescita della capitalizzazione di mercato fino all'anno 2029.

Figura 2 Dimensioni del settore turistico in termini di fatturato (\$ Bln)

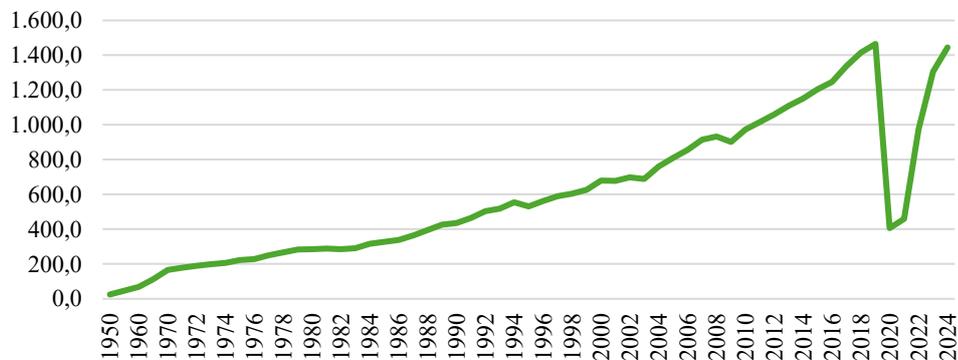


Fonte: Elaborazione dell'autore sui dati [IBISWorld, 2024](#)

La figura 2 ([IBISWorld, 2024](#)) conferma ulteriormente il trend di crescita del settore turistico, analizzandolo in termini di fatturato; anche in questo caso si verifica un drastico calo in corrispondenza del 2020 seguito però da una rapida crescita.

Il settore del turismo, oltre ad essere caratterizzato dal punto di vista economico, può essere analizzato in termini di andamento del numero di arrivi, per motivi turistici, nelle diverse città del mondo. Il trend presente nella figura 3 ([World Tourism Organization \[UN Tourism\], 2025](#)) è crescente, con una battuta di arresto nel 2020 nel periodo della pandemia da Covid-19, dal 2022 in poi però si è verificata una rapida ripresa.

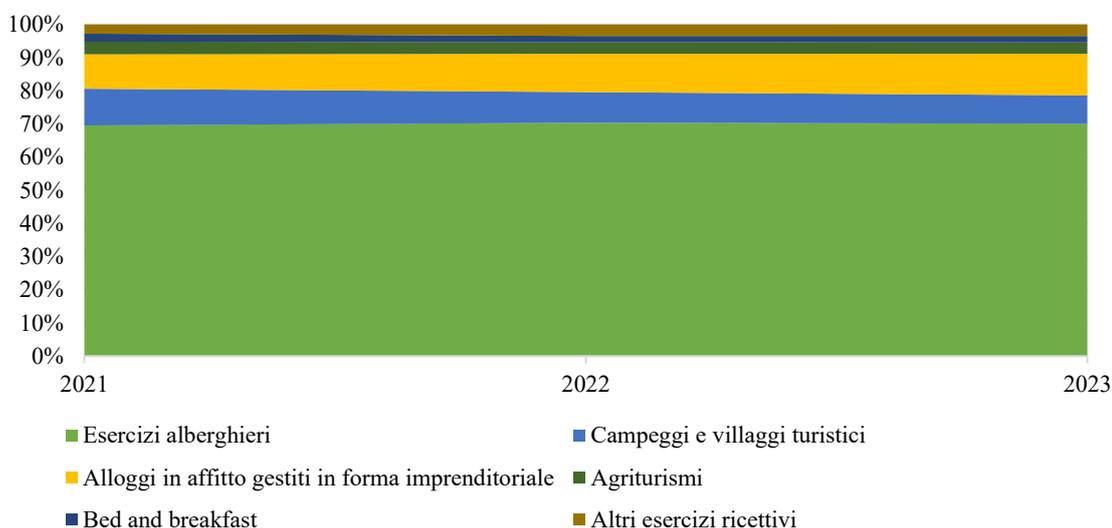
Figura 3 Numero di arrivi turistici nel mondo (M)



Fonte: Elaborazione dell'autore sui dati [World Tourism Organization \[UN Tourism\], 2025](#)

Facendo un focus sull'Italia, gli arrivi sono caratterizzati principalmente da soggiorni in strutture alberghiere, come mostrato dalla figura 4 ([Istituto Nazionale di Statistica \[ISTAT\], 2025](#)).

Figura 4 Composizione percentuale del numero di arrivi in Italia



Fonte: Elaborazione dell'autore sui dati [Istituto Nazionale di Statistica \[ISTAT\], 2025](#)

Analizzando, però, la variazione percentuale della composizione del numero di arrivi in Italia si può notare nella tabella 1 ([Istituto Nazionale di Statistica \[ISTAT\], 2025](#)) che l'aumento percentuale da un anno all'altro è superiore nel caso degli altri servizi ricettivi rispetto agli esercizi alberghieri.

Tabella 1 Variazione percentuale della composizione del numero di arrivi in Italia

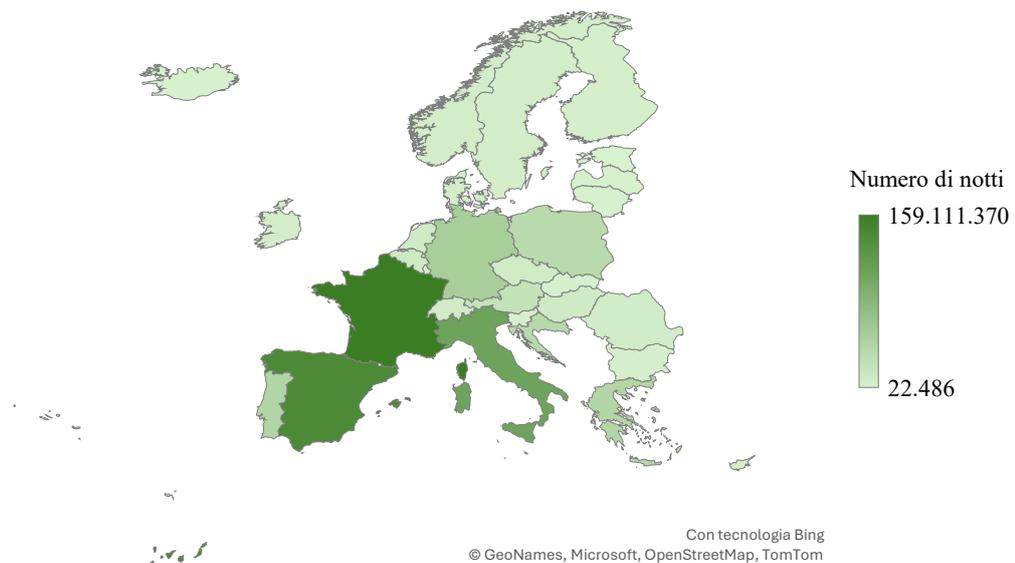
	$\Delta 2021-2022$	$\Delta 2022-2023$
Esercizi alberghieri	52,4%	12,4%
Campeggi e villaggi turistici	26,3%	3,5%
Alloggi in affitto gestiti in forma imprenditoriale	66,8%	23,2%
Agriturismi	35,0%	11,0%
Bed and breakfast	27,2%	9,5%
Altri esercizi ricettivi	86,2%	14,9%

Fonte: Elaborazione dell'autore sui dati [Istituto Nazionale di Statistica \[ISTAT\], 2025](#)

Sebbene, dunque, la maggioranza del numero di arrivi in Italia sia indirizzata agli esercizi alberghieri, l'aumento percentuale degli altri esercizi ricettivi è nettamente maggiore, facendo ipotizzare un trend crescente nel tempo.

A dimostrazione dell'importanza delle piattaforme di sharing economy nel settore dei viaggi, oltre agli arrivi, è possibile analizzare il numero di notti trascorse per soggiorni brevi mediante l'utilizzo di tali piattaforme.

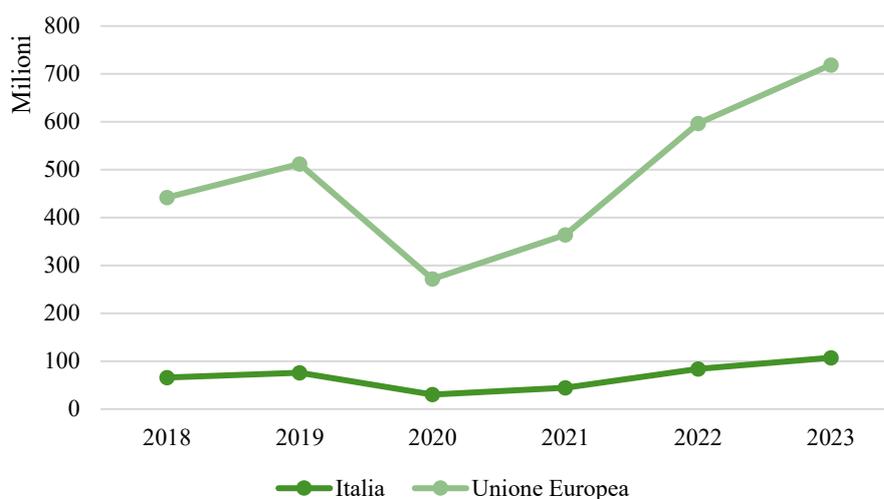
Figura 5 Numero di notti trascorse per soggiorni brevi offerti da piattaforme di sharing economy nel 2023



Fonte: Elaborazione dell'autore sui dati [Eurostat, 2024](#)

Le piattaforme di sharing economy permettono una forte diffusione del fenomeno turistico, come si evince in figura 5 ([Eurostat, 2024](#)), evidenziando in modo grafico l'eterogeneità della distribuzione del numero di notti trascorse, con un focus su Francia, Spagna ed Italia che presentano il numero maggiore di pernottamenti.

Figura 6 Numero di notti trascorse per soggiorni brevi offerti da piattaforme di sharing economy



Fonte: Elaborazione dell'autore sui dati [Eurostat, 2024](#)

Nella figura 6 ([Eurostat, 2024](#)), viene mostrato l'andamento nel tempo del numero di notti trascorse, confrontando il caso europeo con il caso italiano, come si può notare, l'andamento è crescente nel tempo ad eccezione di un calo nel 2020 in coincidenza con lo scoppio della pandemia da Covid-19, ma seguito subito da una rapida ripresa.

Figura 7 Numero di notti trascorse per soggiorni brevi offerti da piattaforme di sharing economy in Italia



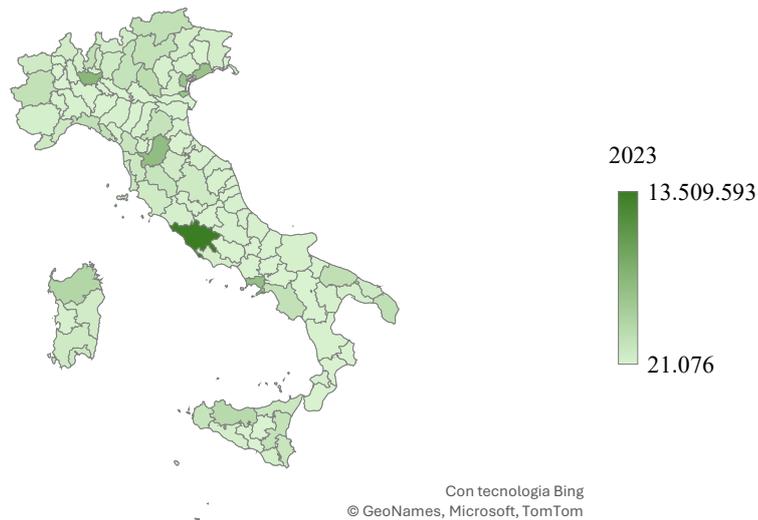
Fonte: Elaborazione dell'autore sui dati [Eurostat, 2024](#)

L'andamento del numero di notti nel corso degli anni nel caso italiano presenta un andamento simile a quello Europeo, come rappresentato nella figura 7 ([Eurostat, 2024](#)), anche se in figura 6 ([Eurostat, 2024](#)) l'andamento italiano risulta un po' compresso. Il

mercato italiano è una delle componenti principali dell'Europa, pesando infatti il 15% ([Eurostat, 2024](#)) nel 2023 in termini di notti trascorse.

Facendo un focus sull'Italia, si verifica la seguente distribuzione delle notti trascorse nel caso di soggiorni brevi offerti da piattaforme di sharing economy:

Figura 8 Numero di notti trascorse per soggiorni brevi offerti piattaforme di sharing economy in Italia



Fonte: Elaborazione dell'autore sui dati [Eurostat, 2024](#)

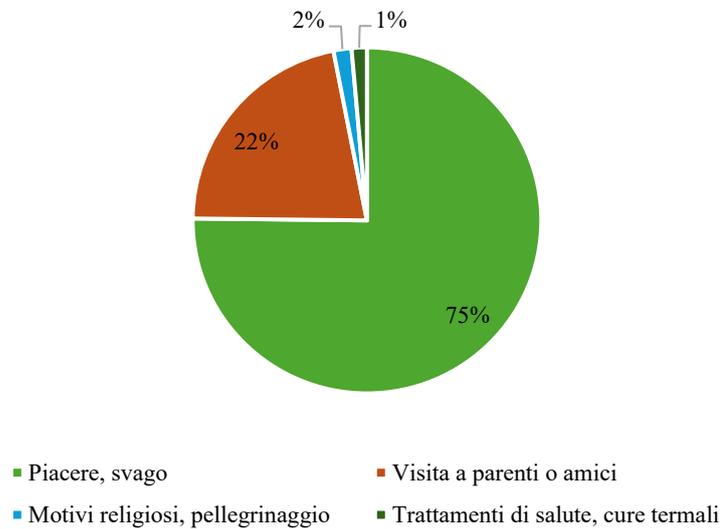
La figura 8 ([Eurostat, 2024](#)) permette di comprendere al meglio l'eterogeneità della distribuzione delle notti trascorse in Italia mediante l'utilizzo di piattaforme di economia collaborativa ed evidenziando Roma come principale città in termini di arrivi.

La motivazione alla base degli spostamenti, oltre che la rispettiva distribuzione in termini spaziali, è di fondamentale importanza per le analisi. Le figure 9 e 10 ([Istituto Nazionale di Statistica \[ISTAT\], 2025](#)) mostrano la composizione percentuale della motivazione alla base rispettivamente di:

- Vacanze brevi (ovvero vacanze caratterizzate da 1 a 3 notti di soggiorno, secondo l'ISTAT);
- Vacanze lunghe (ovvero vacanze caratterizzate da un numero di notti maggiore o uguale a 4, secondo l'ISTAT).

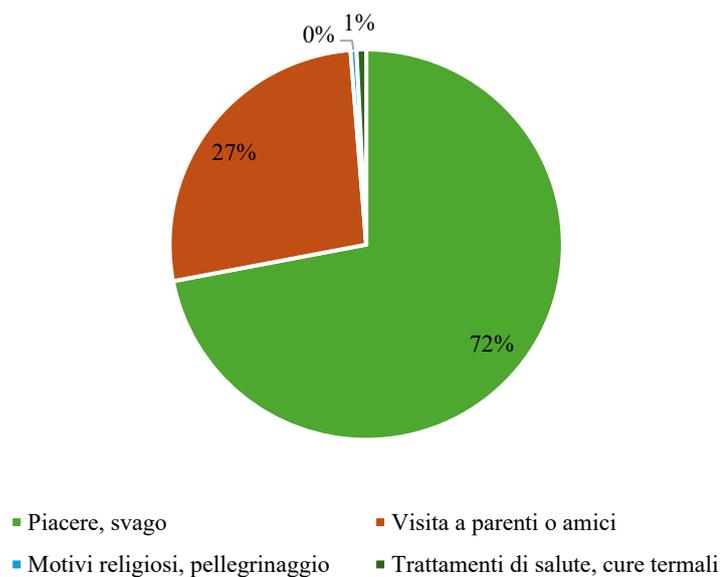
Si nota che in entrambi i casi la motivazione principale è legata al piacere e allo svago anche se per le vacanze lunghe la percentuale è leggermente inferiore rispetto alle vacanze brevi, a favore di un aumento della motivazione legata alla visita di parenti e/o amici.

Figura 9 Motivo delle vacanze brevi in Italia nel 2023



Fonte: Elaborazione dell'autore sui dati [Istituto Nazionale di Statistica \[ISTAT\], 2025](#)

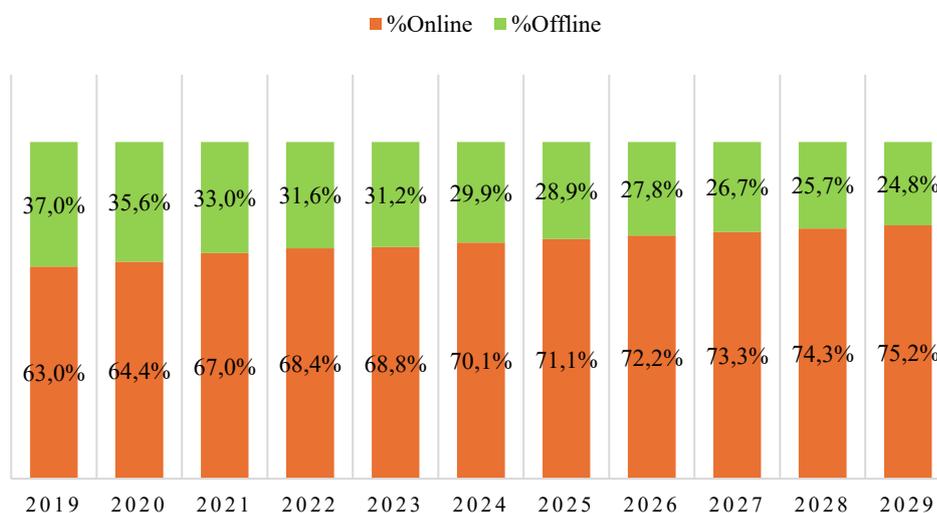
Figura 10 Motivo delle vacanze lunghe in Italia nel 2023



Fonte: Elaborazione dell'autore sui dati [Istituto Nazionale di Statistica \[ISTAT\], 2025](#)

In merito alla prenotazione di un viaggio, nella figura 11 ([Statista, 2024](#)) viene rappresentata la composizione percentuale dei canali di vendita nel mercato dei viaggi.

Figura 11 Composizione percentuale dei canali di vendita



Fonte: Elaborazione dell'autore sui dati [Statista, 2024](#)

Nel tempo, come mostrato nella figura 11 ([Statista, 2024](#)), i canali offline lasceranno sempre più posto ai canali online che già attualmente rappresentano la stragrande maggioranza dei canali di vendita del mercato dei viaggi. Le proiezioni fornite da Statista fino al 2029 prevedono un continuo aumento del peso del canale online a dimostrazione della crescente importanza delle piattaforme.

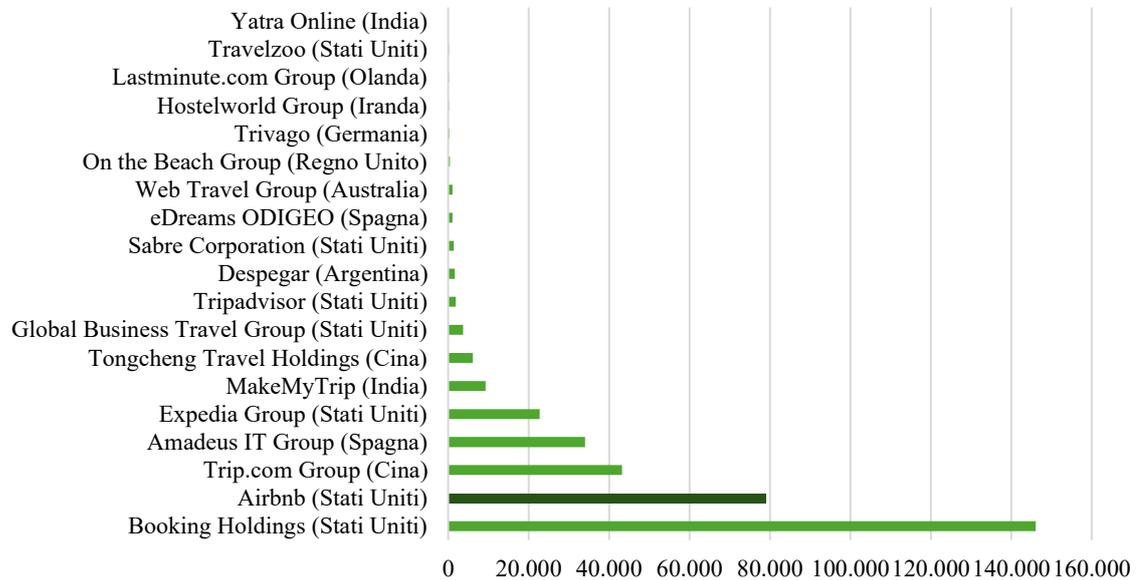
1.3 Airbnb

Dimostrata l'importanza della sharing economy e come questa si sia diffusa nel settore del turismo, si analizza uno dei player principali dell'economia collaborativa nel settore turistico: Airbnb. L'analisi della letteratura su Airbnb è ampia a tal punto da essere stata studiata l'evoluzione delle differenti tematiche alla base delle ricerche ([Andreu et al., 2020](#)).

Airbnb viene fondata nel 2008 ([Airbnb, n.d.](#)) e rappresenta una piattaforma multiversante che favorisce la comunicazione tra gli host che mettono a disposizione

proprietà/esperienze e gli ospiti che desiderano soggiornare in un luogo o vivere delle esperienze uniche.

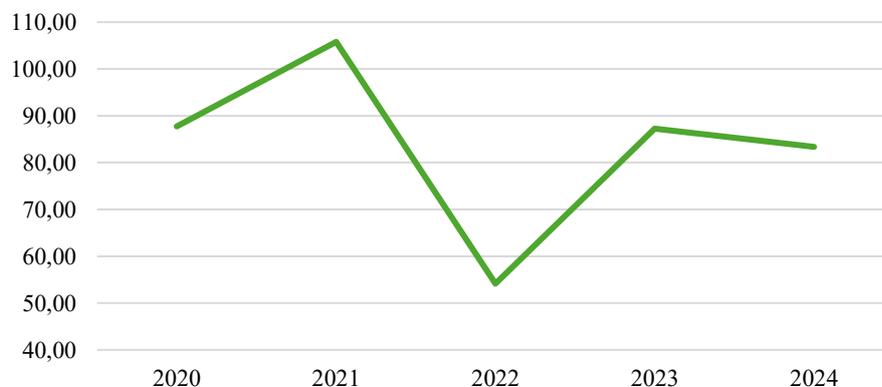
Figura 12 Capitalizzazione di mercato a livello mondiale a marzo 2025 (\$ Mld)



Fonte: Elaborazione dell'autore sui dati [Reuters, 2025](#)

Airbnb, come mostrato in figura 12 ([Reuters, 2025](#)), è il secondo tra le diverse società di viaggi in termini di capitalizzazione di mercato, rimarcando la sua importanza come player della sharing economy nel settore del turismo.

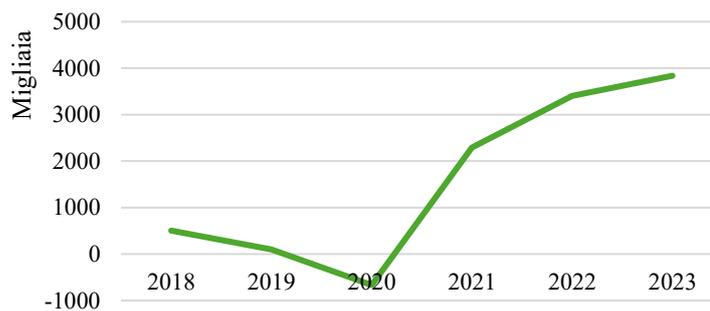
Figura 13 Capitalizzazione di mercato di Airbnb (\$ Mld)



Fonte: Elaborazione dell'autore sui dati [CompaniesMarketCap.com, 2025](#)

La figura 13 ([CompaniesMarketCap.com, 2025](https://www.companiesmarketcap.com)) mette in evidenza l'andamento altalenante della capitalizzazione di mercato di Airbnb che nel 2024 arriva a circa \$ 83,34 Mld ([CompaniesMarketCap.com, 2025](https://www.companiesmarketcap.com)). La capitalizzazione di mercato viene definita dal prodotto del numero di azioni per il valore di queste ultime ([Oxford University Press, n.d.](https://www.oxforduniversitypress.com)), dunque, tale valore può essere influenzato da differenti fattori andando a sottolineare l'importanza di analizzare Airbnb dal punto di vista economico.

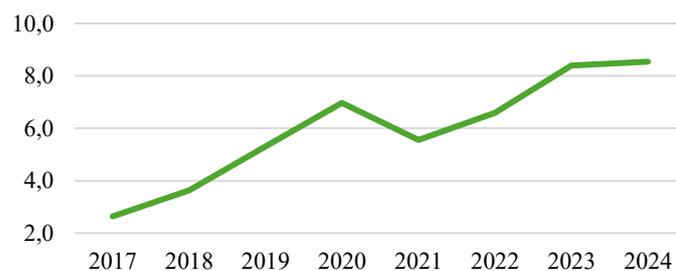
Figura 14 Free cash flow (k\$)



Fonte: Elaborazione dell'autore sui dati [Airbnb, Inc., 2020](https://www.airbnb.com); [Airbnb, Inc., 2021](https://www.airbnb.com); [Airbnb, Inc., 2022](https://www.airbnb.com); [Airbnb, Inc., 2023](https://www.airbnb.com)

In termini di free cash flow, come mostrato in figura 14 ([Airbnb, Inc., 2020](https://www.airbnb.com); [Airbnb, Inc., 2021](https://www.airbnb.com); [Airbnb, Inc., 2022](https://www.airbnb.com); [Airbnb, Inc., 2023](https://www.airbnb.com)), l'andamento è crescente nel tempo, con un unico picco verso il basso in corrispondenza del 2020 ma subito seguito da una rapida ripresa. Il free cash flow rappresenta un indicatore di salute finanziaria che però esclude le spese non monetarie del conto economico ([Fernando, 2024](https://www.fernando.com)), dunque per avere una visione completa in termini economici di Airbnb è necessario analizzare anche altri aspetti come l'utile netto ed i costi sostenuti.

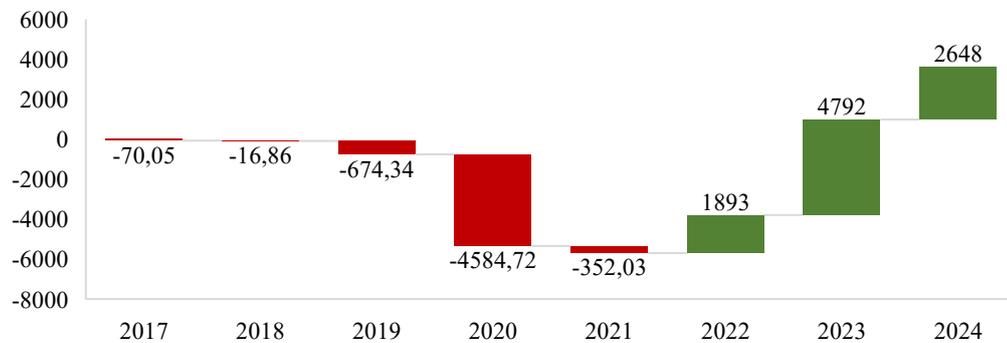
Figura 15 Costi e spese di Airbnb (\$ Mld)



Fonte: Elaborazione dell'autore sui dati [U.S. Securities and Exchange Commission, 2025](https://www.sec.gov)

I costi e le spese sostenute da Airbnb, come mostrato in figura 15 ([U.S. Securities and Exchange Commission, 2025](#)), sono crescenti nel tempo facendo ipotizzare dei crescenti investimenti all'interno della piattaforma stessa.

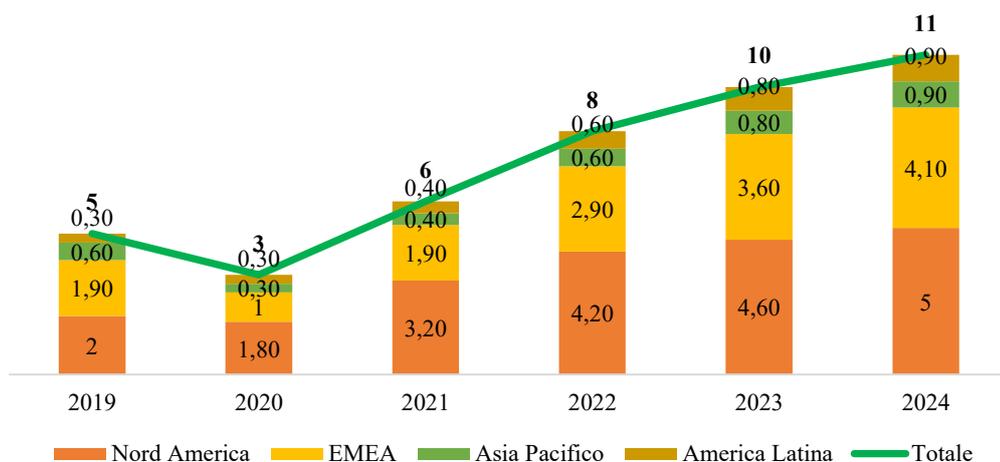
Figura 16 Utile Netto Airbnb (M\$)



Fonte: Elaborazione dell'autore sui dati U.S. Securities and Exchange Commission, 2025

In termini di utile netto, come mostrato dalla figura 16 ([U.S. Securities and Exchange Commission, 2025](#)), sebbene inizialmente si siano verificate delle perdite anche normali per un'azienda in crescita, con un picco nel 2020 (probabilmente dovuto al Covid-19), poi l'utile netto dal 2022 diventa positivo con una crescita graduale andando a dimostrare la rapidità di adattamento della piattaforma e la capacità di generare utili nel tempo.

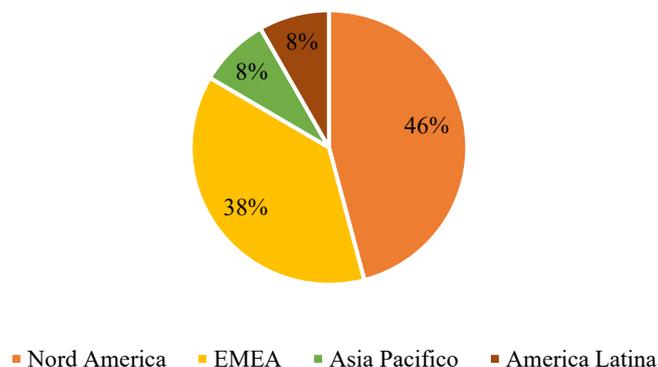
Figura 17 Fatturato di Airbnb (\$ Mld)



Fonte: Elaborazione dell'autore sui dati U.S. Securities and Exchange Commission, 2025

In termini di fatturato, come mostrato nella figura 17 ([U.S. Securities and Exchange Commission, 2025](#)), si può notare una rapida crescita dal 2020 fino a raggiungere nel 2024 \$ 11 Mld.

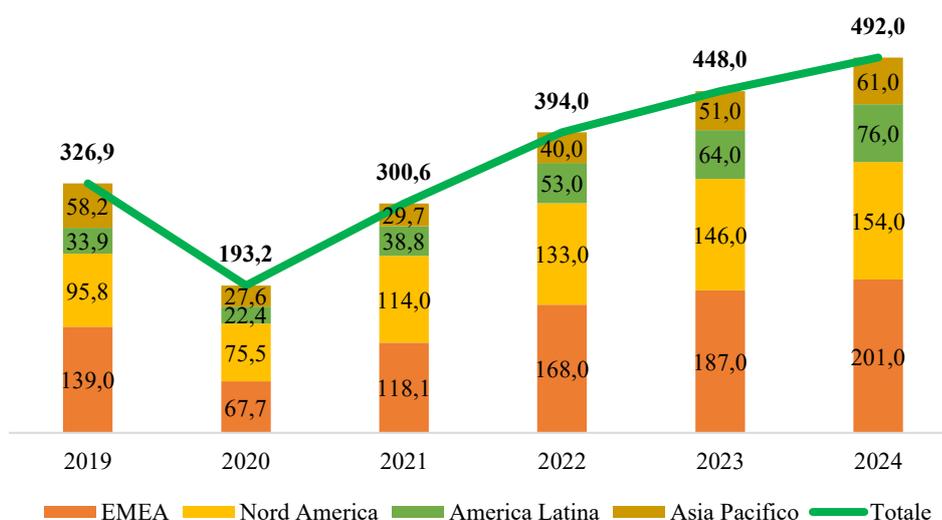
Figura 18 Composizione percentuale del fatturato di Airbnb nel 2024



Fonte: Elaborazione dell'autore sui dati [U.S. Securities and Exchange Commission, 2025](#)

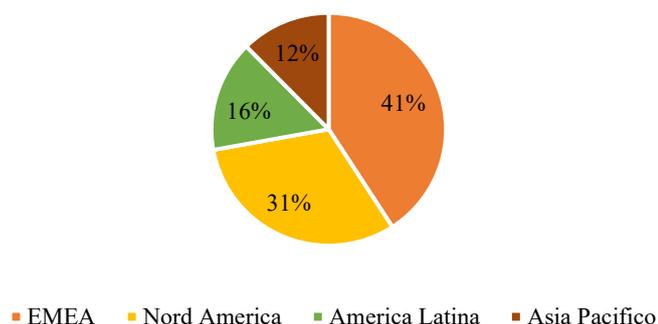
In termini di composizione percentuale, la figura 18 ([U.S. Securities and Exchange Commission, 2025](#)) mostra che il fatturato proviene in gran parte dal Nord America subito seguito dai paesi EMEA (acronimo di Europa, Medio Oriente e Africa ([IBM, 2024](#))).

Figura 19 Notti ed esperienze prenotate su Airbnb (M)



Fonte: Elaborazione dell'autore sui dati [U.S. Securities and Exchange Commission, 2025](#)

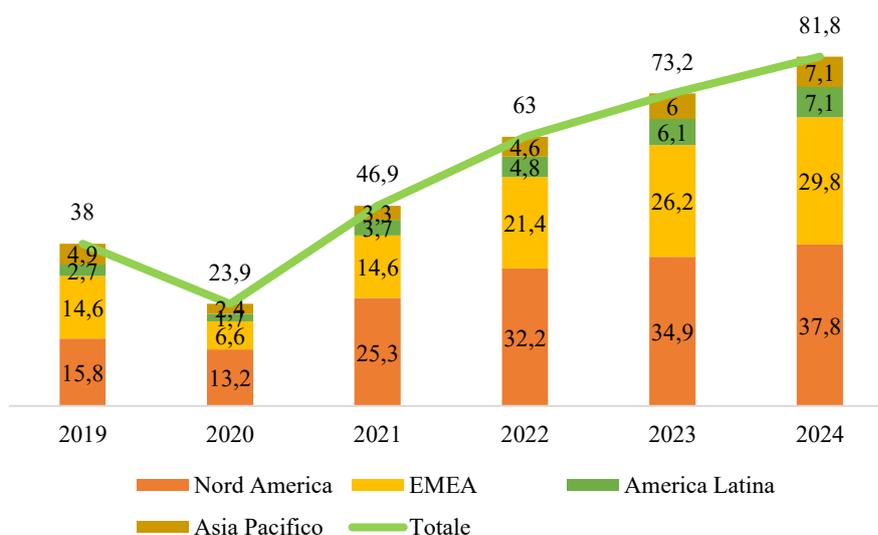
Figura 20 Composizione percentuale delle notti ed esperienze prenotate su Airbnb nel 2024



Fonte: Elaborazione dell'autore sui dati [U.S. Securities and Exchange Commission, 2025](#)

La figura 19 ([U.S. Securities and Exchange Commission, 2025](#)) presenta un andamento crescente del numero di notti ed esperienze prenotate su Airbnb. La figura 20 ([U.S. Securities and Exchange Commission, 2025](#)), in contraddizione con la figura 18 ([U.S. Securities and Exchange Commission, 2025](#)), mostra che la percentuale maggiore di notti prenotate è da attribuire ai paesi EMEA; dunque, anche se il Nord America rappresenta in termini percentuali il mercato che contribuisce maggiormente alla generazione di fatturato, i paesi EMEA risultano quelli caratterizzati dal maggior numero di prenotazioni. Quanto detto in precedenza può essere spiegato dal valore lordo della prenotazione.

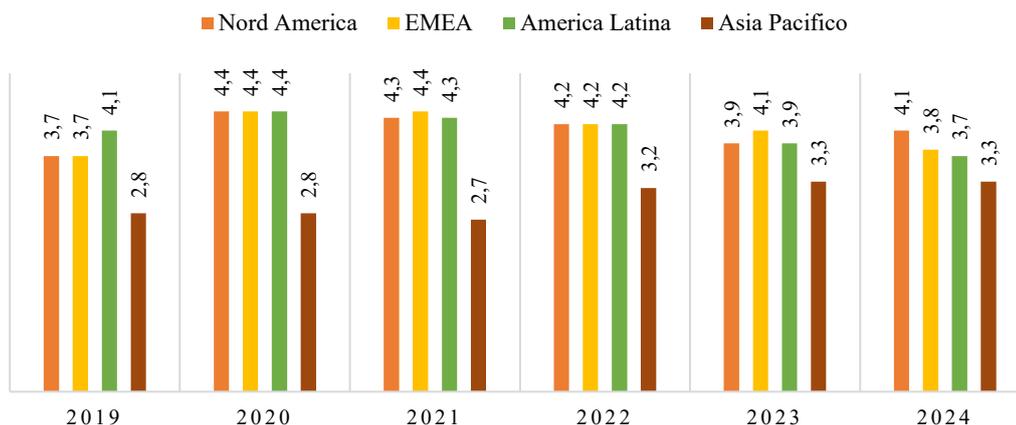
Figura 21 Valore lordo della prenotazione (\$ Mld)



Fonte: Elaborazione dell'autore sui dati [U.S. Securities and Exchange Commission, 2025](#)

La figura 21 ([U.S. Securities and Exchange Commission, 2025](#)) mostra il valore lordo crescente delle prenotazioni su Airbnb. Un dato particolarmente interessante è la composizione del valore lordo della prenotazione nel 2024: si può notare che il valore lordo delle prenotazioni in Nord America è maggiore rispetto ai paesi EMEA, questo potrebbe spiegare il perché, sebbene il numero di notti prenotate sia maggiore nei paesi EMEA, il fatturato derivi principalmente dal Nord America.

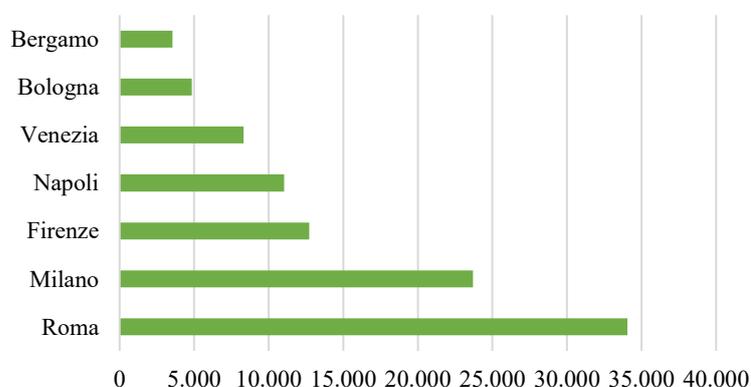
Figura 22 Numero medio di notti per prenotazione



Fonte: Elaborazione dell'autore sui dati [U.S. Securities and Exchange Commission, 2025](#)

In figura 22 ([U.S. Securities and Exchange Commission, 2025](#)) viene rappresentato il numero medio di notti trascorse per prenotazione dal 2019 al 2024, si nota come dal 2019 al 2024 il numero medio di notti sia diminuito per l'America Latina ed aumentato per i paesi EMEA, per il Nord America e per l'Asia Pacifico.

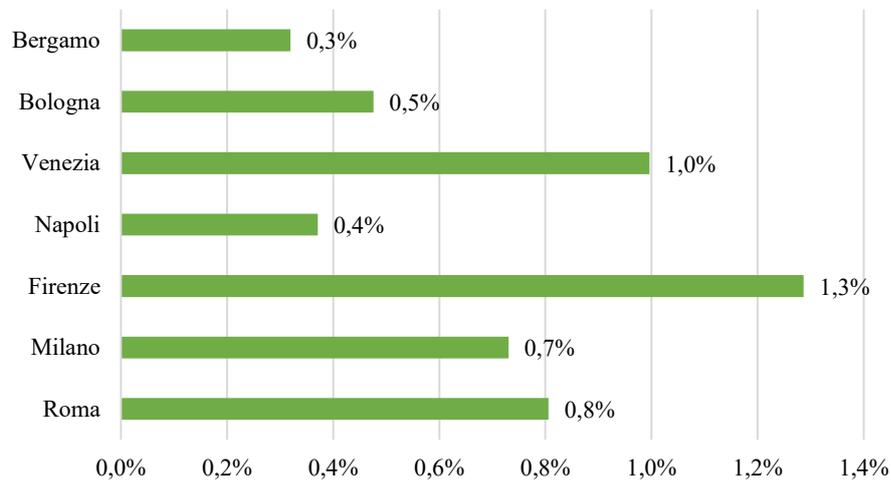
Figura 23 Numero di annunci Airbnb a Settembre 2024



Fonte: Elaborazione dell'autore sui dati [Inside Airbnb, 2024](#)

Facendo un focus sull'Italia, come si può notare nella figura 23 ([Inside Airbnb, 2024](#)), le città caratterizzate da un maggior numero di annunci sono rispettivamente Roma, Milano, Firenze e Napoli.

Figura 24 Proporzione (Numero di annunci/numero popolazione residente)



Fonte: Elaborazione dell'autore sui dati [Inside Airbnb, 2024](#); [Istituto Nazionale di Statistica \[ISTAT\], 2025](#)

Il solo numero di annunci però non è molto indicativo così, come viene mostrato in figura 24 ([Inside Airbnb, 2024](#); [Istituto Nazionale di Statistica \[ISTAT\], 2025](#)), il numero di annunci viene riproporzionato in funzione della popolazione residente al primo gennaio 2025; da tale rapporto si può notare che in ordine le città con proporzione maggiore sono Firenze, Venezia, Roma e Milano. L'ordine, dunque, mostrato in figura 24 ([Inside Airbnb, 2024](#); [Istituto Nazionale di Statistica \[ISTAT\], 2025](#)) è differente rispetto a quello della figura 23 ([Inside Airbnb, 2024](#)).

Capitolo 2 – Analisi della letteratura

Le ricerche in merito ad Airbnb fanno riferimento a differenti tematiche, una tra queste è il comportamento pre/durante/dopo l'acquisto ([Andreu et al., 2020](#)). Prima di vivere un'esperienza su Airbnb, i clienti non godono di informazione perfetta; dunque, potrebbero crearsi problemi di “azzardo morale”¹ e/o “selezione avversa”², necessitando, così, la presenza di un segnale credibile di qualità che non sia “cheap talk”. I problemi di asimmetria informativa sono stati ampiamente studiati in letteratura soprattutto in merito al mondo del lavoro come nel caso di [Spence \(2002\)](#) e [Spence \(1981\)](#). Airbnb secondo [Vinogradov, Leick e Kivedal \(2020\)](#) continuerà a crescere anche dopo che la domanda sarà uguale all'offerta, dunque, è importante studiare i fenomeni che caratterizzano il suo funzionamento, anche alla luce di quanto detto da [Ajzen \(1991\)](#) sulla teoria del comportamento pianificato, per comprendere al meglio il comportamento di acquisto degli utenti.

Dall'analisi della letteratura si possono individuare i temi caldi a partire dai quali incentrare le analisi. Come individuato da [Santos et al. \(2022\)](#), grazie all'analisi delle recensioni, i temi a cui gli utenti si riferiscono principalmente sono le caratteristiche dell'host, le caratteristiche dell'alloggio e la posizione di quest'ultimo. Tali tematiche sono state individuate anche da [Demirdelen et al. \(2020\)](#) mediante l'analisi delle recensioni positive. [Sthapit, Björk, Coudounaris e Stone \(2022\)](#) hanno confermato gli argomenti definiti in precedenza come tematiche che suscitano emozioni memorabili sia esse positive e/o negative, andandone a rimarcare l'importanza. Il servizio fornito dall'host, le caratteristiche dell'appartamento e la posizione di quest'ultimo, vengono definiti come argomenti core che influenzano la soddisfazione del cliente da [Cavique et al. \(2022\)](#). Anche [Li, Chen e Huang \(2020\)](#) hanno studiato le tematiche che suscitano emozioni nei clienti Airbnb andando a definire i servizi speciali, atmosfera e flessibilità come elementi che suscitano emozioni positive, mentre il disturbo arrecato dall'host come emozione negativa, ma pur sempre significativa. [Klegr \(2024\)](#) ribadisce le principali

¹ L'azzardo morale viene definito dall'Oxford University Press come “*lack of reason to try to avoid risk when protected from its consequences, for example by insurance*” ([Oxford University Press, n.d.](#)).

² La definizione di “selezione avversa” è disponibile presso [Edizioni Simone \(n.d.\)](#). *Dizionari Simone*. Presente all'indirizzo <https://dizionari.simone.it/6/adverse-selection>.

argomentazioni individuate in precedenza come fondamentali, in particolare definisce l'host ed il quartiere come elementi distintivi di Airbnb mentre le caratteristiche fisiche dell'appartamento come caratteristiche distintive degli hotel che però rappresentano un punto di debolezza per Airbnb. In termini di dimensioni della qualità, [Amat-Lefort, Barravecchia e Mastrogiacomo \(2023\)](#), mediante l'analisi delle recensioni, ne hanno individuate diverse sia in riferimento a caratteristiche tangibili sia in riferimento a caratteristiche intangibili. Il valore della qualità sugli atteggiamenti, così come il valore sociale sulle intenzioni di acquisto definito da [Tajeddini, Gamage, Hameed, Qumsieh-Mussalam, Chaijani, Rasoolimanesh e Kallmuenzer \(2022\)](#) come significativo, è stato ritenuto da [So, Kim e Min \(2022\)](#) determinante per i comportamenti di acquisto futuri dei clienti. Secondo [Qiu et al. \(2023\)](#) i fattori che influenzano positivamente l'intimità (componente fondamentale delle recensioni Airbnb) riguardano il luogo, il comportamento dell'host, l'emozione e la cultura. L'analisi delle recensioni è stata, dunque, al centro di diversi studi come: [Amat-Lefort, Barravecchia e Mastrogiacomo \(2023\)](#), [Cavique, Ribeiro, Batista e Correia \(2022\)](#), [Li, Chen e Huang \(2020\)](#), [Santos, A. I. G. P., Costa Perinotto, A. R., Rodrigues Soares, J. R. e Savi Mondo, T. \(2022\)](#), [Demirdelen, Dinçer e İstanbullu Dinçer \(2020\)](#), [Qiu, Yu, Tsai, Zhu, Chen e Kim \(2023\)](#), [Liang, Zhang e Li \(2020\)](#), [Klegr \(2024\)](#) e così via. Lo studio delle recensioni, oltre che per scopi accademici, è stato adoperato anche da Airbnb stesso per migliorare le sue performances, infatti, [Garay-Tamajón e Morales-Pérez \(2023\)](#) hanno riscontrato come l'analisi di queste ultime sia stata utilizzata per la realizzazione delle guide fornite dalla piattaforma al fine di toccare i principali punti di interesse espressi dagli utenti. Airbnb ha un impatto, oltre che sull'immagine della destinazione ([Garay-Tamajón & Morales-Pérez, 2023](#)) anche in termini di gentrificazione ([Gurran, Zhang, & Shrestha, 2020](#)), dunque, è importante comprendere come vengono strutturati gli annunci e quale impatto hanno questi sulla scelta dei clienti.

In termini di composizione dell'annuncio, [Canziani e Nemati \(2021\)](#) ne analizzano le componenti ed in particolare gli spunti tonali, individuando un forte utilizzo di aggettivi la cui rilevanza è stata già definita in relazione al place branding ([Giovanardi, 2012](#)). [Liang, Zhang e Li \(2020\)](#) confermano l'effetto positivo della verbosità dell'annuncio andando ad attestare la correlazione positiva tra il dettaglio delle descrizioni e l'informatività delle recensioni. [Zheng e Zhang \(2023\)](#) hanno studiato il valore dello storytelling in relazione al legame che si crea tra il brand ed i clienti, mentre [Godovykh e](#)

[Baker \(2022\)](#) hanno individuato come la presenza dei quadri negli annunci abbia un impatto sulle intenzioni di acquisto degli utenti. In termini di studio delle performance dell'annuncio legate alle sue caratteristiche, [Xie e Mao \(2017\)](#) si soffermano sul legame tra le caratteristiche di qualità dell'host e le prestazioni dell'annuncio, [Kirkos \(2022\)](#) analizza le principali determinanti dell'annuncio che ne influenzano le performance, confermando come elementi principali, la descrizione dell'host e le sue caratteristiche come il possesso del badge superhost ed il tasso di risposta, mentre [Chung e Sarnikar \(2022\)](#) si soffermano sullo studio dei legami tra le descrizioni degli annunci e performance.

Questa tesi, pertanto, si propone di analizzare la correlazione tra le caratteristiche semantiche del titolo dell'annuncio e le performance dell'annuncio stesso. Le variabili prestazionali esaminate in questo studio sono³:

- $ADR = \frac{\text{revenue usd}}{\text{reservation days}}$
- $revPAN = \frac{\text{revenue usd}}{\text{reservation days} + \text{available days}}$

Dove:

- Revenue USD: rappresenta l'ammontare dei ricavi nell'anno di riferimento espresso in dollari e comprende sia il prezzo definito al momento della prenotazione sia l'ammontare delle spese di pulizia;
- Reservation days: rappresenta l'ammontare dei giorni prenotati nell'arco dell'anno di riferimento;
- Available Days: rappresenta l'ammontare di giorni disponibili nell'anno di riferimento⁴.

Si ha l'intenzione, dunque, di proporsi come strumento di congiunzione con gli studi precedenti ed introdurre nuovi elementi alla letteratura esistente. Si presenta, di

³ La definizione delle variabili "ADR" e "revPAN" è stata fornita dalla relatrice Elisabetta Raguseo e dal correlatore Francesco Luigi Milone.

⁴ Le definizioni delle variabili fanno riferimento al file Excel "Airbnb_variabili_fino_2022" fornito dalla relatrice Elisabetta Raguseo (E. Raguseo, comunicazione personale, 2025) e al dataset fornito dalla relatrice.

conseguenza, la seguente tabella riassuntiva delle tematiche principali derivanti dagli studi effettuati in precedenza, presi in esame durante lo svolgimento del lavoro di tesi.

Tabella 2 Riassunto delle principali tematiche degli studi effettuati in precedenza presi in esame durante lo svolgimento della tesi

<p>Analisi delle caratteristiche dell'host, delle caratteristiche dell'immobile e della posizione di quest'ultimo come elementi influenti sulle performance di Airbnb.</p>	<ul style="list-style-type: none"> - Amat-Lefort, Barravecchia e Mastrogiacomo (2023) - Cavique et al. (2022) - Demirdelen et al. (2020) - Klegr (2024) - Li, Chen e Huang (2020) - Qiu et al. (2023) - Santos et al. (2022) - Sthapit, Björk, Coudounaris e Stone (2022)
<p>Analisi dell'impatto delle tariffe separate sulla domanda di alloggi.</p>	<ul style="list-style-type: none"> - Dogru, Majid, Laroche, Mody e Suess (2021)
<p>Analisi delle caratteristiche dell'annuncio come elementi legati alle sue performance.</p>	<ul style="list-style-type: none"> - Canziani e Nemati (2021) - Chung e Sarnikar (2022) - Godovykh e Baker (2022) - Kirkos (2022) - Liang, Zhang e Li (2020) - Xie e Mao (2017) - Zheng e Zhang (2023)

Fonte: Le fonti sono presenti all'interno della tabella stessa

Capitolo 3 – Analisi dei dati

3.1 Descrizione dataset

Il dataset⁵ ([Raguseo, 2025](#)) iniziale presentava 3.176.475 osservazioni e 33 variabili riferite agli annunci Airbnb in Italia; le variabili di tale dataset assumono il seguente significato:

- Property ID: ID univoco associato alla proprietà offerta;
- Year: anno di riferimento;
- Number of Reservations: numero di prenotazioni;
- Revenue (USD): ricavi nell'anno di riferimento espressi in dollari che comprendono sia il prezzo definito al momento della prenotazione sia l'ammontare delle spese di pulizia;
- Reservation Days: numero di giorni prenotati;
- Available Days: giorni disponibili;
- Listing title: titolo dell'annuncio;
- Property type: tipologia di alloggio offerto;
- Listing type: fa riferimento a quattro categorie che possono essere "Intero appartamento", "Camera privata", "Camera condivisa" o "Hotel";
- Created Date: data di creazione dell'annuncio;
- Country: paese in cui si trova l'alloggio;
- Latitude: latitudine dell'alloggio;
- Longitude: longitudine dell'alloggio;
- State: regione in cui si trova l'alloggio;
- City: città in cui si trova l'alloggio;
- Bedrooms: numero di camere da letto presenti nell'alloggio;
- Bathrooms: numero di bagni presenti nell'alloggio;
- Max guests: numero massimo di ospiti ammessi nell'alloggio;
- Response Rate: tasso di risposta di un host entro 24 ore;

⁵ Il dataset è stato fornito dalla relatrice Elisabetta Raguseo ([Raguseo, 2025](#)).

- Airbnb Superhost: variabile che indica se l'host possiede o meno il badge superhost;
- Cancellation Policy: politica di cancellazione della prenotazione riferita all'annuncio;
- Security Deposit (USD): deposito di sicurezza espresso in dollari;
- Cleaning Fee (USD): spese di pulizia per prenotazione espresse in dollari;
- Published Nightly Rate (USD): costo a notte dell'alloggio offerto nell'annuncio, espresso in dollari;
- Published Weekly Rate (USD): costo a settimana dell'alloggio offerto nell'annuncio, espresso in dollari;
- Published Monthly Rate (USD): costo mensile dell'alloggio offerto nell'annuncio, espresso in dollari;
- Minimum Stay: numero minimo di notti richiesto per effettuare la prenotazione dell'alloggio offerto nell'annuncio;
- Number of Reviews: numero di recensioni associate all'annuncio;
- Number of Photos: numero di foto associate all'annuncio;
- Instantbook Enabled: variabile che indica se l'alloggio ammette o meno la possibilità di effettuare una prenotazione istantanea;
- Overall Rating: valutazione associata all'annuncio che varia tra un minimo di una stella ad un massimo di cinque stelle;
- Airbnb Host ID: ID univoco associato all'host;
- Amenities: servizi/comfort.⁶

Preliminarmente alle analisi, il dataset è stato suddiviso in relazione all'anno di riferimento (ovvero la variabile "year" del dataset), ottenendo così sei sottoinsiemi: 2019, 2020, 2021, 2022, 2023 e 2024. Per ogni cluster sono stati eliminati i valori "NA", ovvero non definiti, dalle variabili di interesse per evitare di inficiare le successive analisi; non sono state effettuate ulteriori pulizie per evitare di introdurre bias nella ricerca.

⁶ Le definizioni delle variabili fanno riferimento al file Excel "Airbnb_variabili_fino_2022" fornito dalla relatrice Elisabetta Raguseo (E. Raguseo, comunicazione personale, 2025) e al dataset fornito dalla relatrice.

Tabella 3 Riassunto delle variabili di cui sono stati eliminati i valori non definiti

Variabile	Cancellate osservazioni con valori non definiti (NA)
Numberofreservations	x
Revenueusd	x
Numberofreviews	x
Publishednightlyrateusd	x
Numberofphotos	x
Listingtitle	x
Reservationdays	x
Listingtype	x
Overallrating	x
Propertytype	x
Avaiabledays	x
State	x
Responcerate	x
City	x
Airbnbsuperhost	x

Cancellationpolicy	x
Securitydepositusd	x
Cleaningfeusd	x
Minimumstay	x
Instantbookenabled	x
Bedrooms	x
Bathrooms	x
Maxguests	x
Publishedmonthlyrateusd	x
Publishedweeklyrateusd	x

Fonte: Elaborazione dei dati del dataset fornito dalla relatrice Elisabetta Raguseo (Raguseo, 2025).

La colonna denominata “overallrating” presentava delle discrepanze in termini di valori presenti: da definizione sarebbe dovuta essere caratterizzata da una scala di valori compresi tra uno e cinque, mentre nei dataset relativi al 2022, 2023 e 2024 erano presenti valori che variavano tra 20 e 100, dunque è stata riproporzionata la colonna mediante una divisione per 20, gli outlier caratterizzati da valori inferiori a 20 sono stati ricondotti a 20 mentre gli outlier superiori a 100 sono stati ricondotti a 100.

Si ottengono così i seguenti dataset:

- 2019: con 84.630 osservazioni;
- 2020: con 65.111 osservazioni;

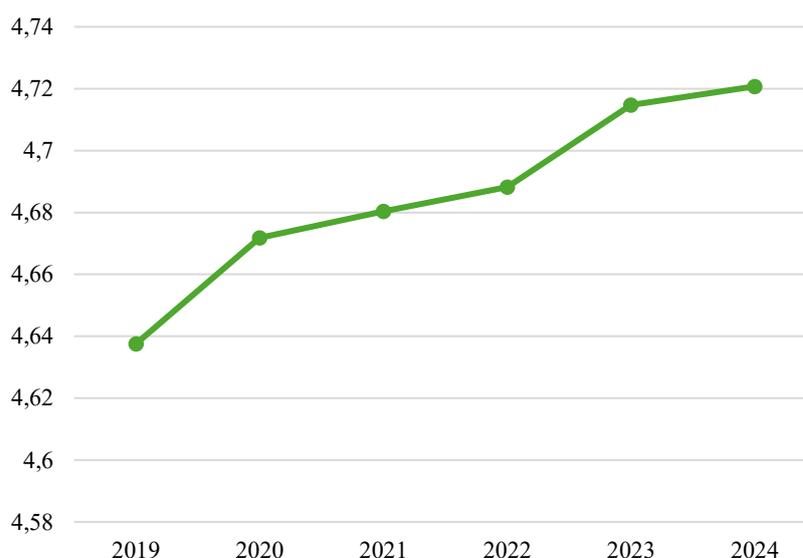
- 2021: con 57.091 osservazioni;
- 2022: con 48.824 osservazioni;
- 2023: con 104.091 osservazioni;
- 2024: con 67.211 osservazioni.

Dopodiché sono stati riuniti i file ottenendo un file finale caratterizzato da 426.959 osservazioni (i file erano stati in precedenza suddivisi per avere una miglior facilità computazionale per la pulizia del dataset).

Mediante l'utilizzo di tabelle pivot, sono stati realizzati differenti grafici che hanno permesso l'individuazione delle principali caratteristiche dell'insieme di dati.

In merito all'andamento della valutazione media associata agli annunci, che di per sé risulta essere già positiva, in accordo con quanto riscontrato da [Cavique et al. \(2022\)](#), si può notare come il trend di positività delle valutazioni degli annunci sia crescente negli anni. Il valor medio delle valutazioni degli annunci passa dall'essere pari a 4,64 nel 2019 a 4,72 nel 2024, indicando probabilmente una maggiore attenzione degli host al soddisfacimento delle esigenze degli utenti.

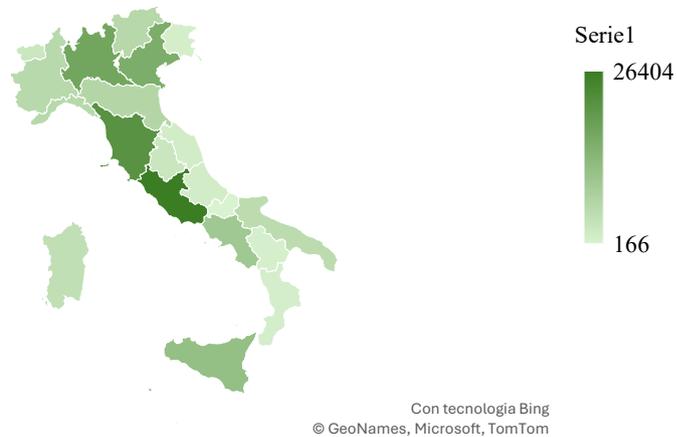
Figura 25 Andamento del valore medio della variabile "overallrating"



Fonte: Elaborazione dei dati del dataset pulito, dataset fornito dalla relatrice Elisabetta Raguseo (Raguseo, 2025).

In riferimento al numero di prenotazioni nel 2024, si può notare come la distribuzione di quest'ultimo sia eterogenea nel territorio italiano: si ha un minimo di 166 prenotazioni effettuate in Molise ed un massimo di 26.404 prenotazioni effettuate in Lazio.

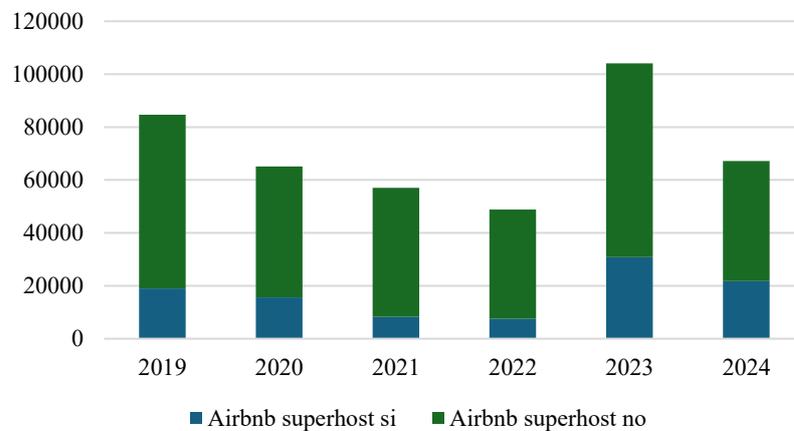
Figura 26 Distribuzione spaziale della somma del numero di prenotazioni nel 2024



Fonte: Elaborazione dei dati del dataset pulito, dataset fornito dalla relatrice Elisabetta Raguseo (Raguseo, 2025).

In merito alla composizione percentuale degli host aventi il badge “superhost”, si può notare che nel corso degli anni, sebbene la maggioranza dei proprietari non posseda tale distintivo, l’andamento risulti essere altalenante. Dal 2019 al 2022, si verifica un andamento decrescente del numero di host aventi la certificazione superhost, per poi crescere nel 2023 e successivamente diminuire nel 2024.

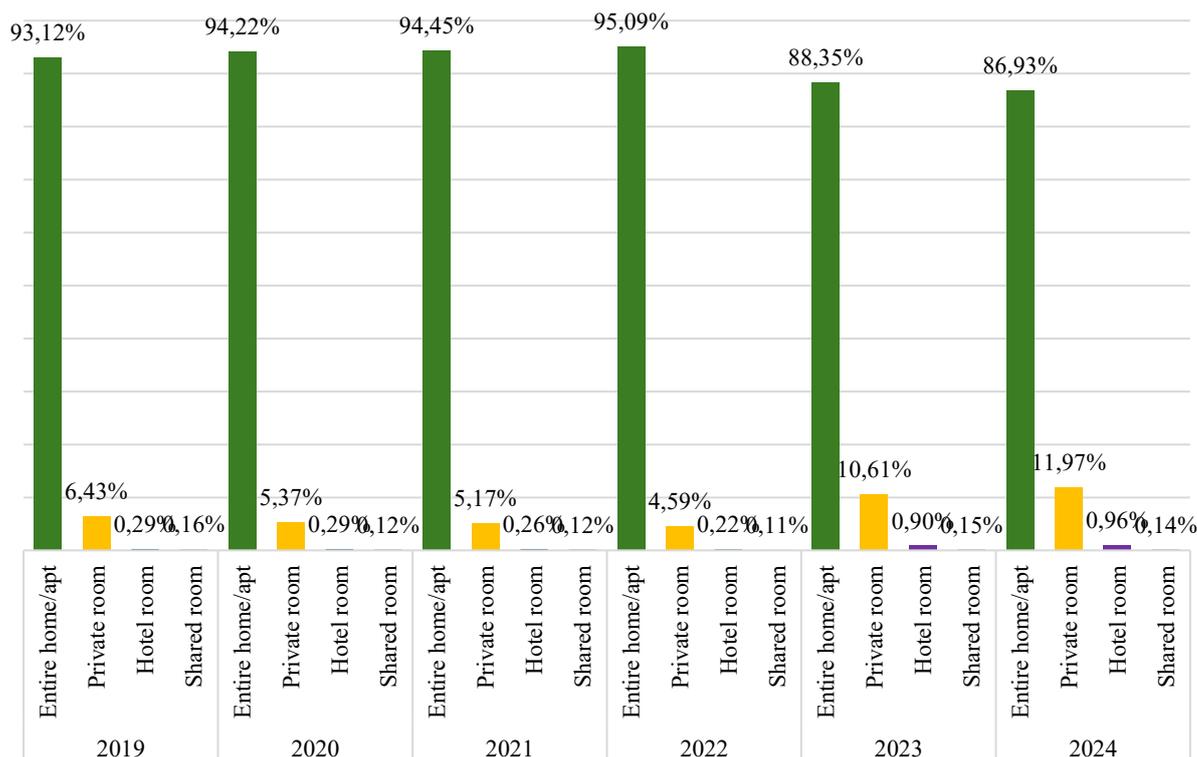
Figura 27 Composizione percentuale degli annunci in termini di possesso del badge superhost



Fonte: Elaborazione dei dati del dataset pulito, dataset fornito dalla relatrice Elisabetta Raguseo (Raguseo, 2025).

Il badge superhost potrebbe essere considerato come un indicatore di qualità dell’host poiché per poterlo ottenere bisognerebbe soddisfare diversi criteri in termini di performance, che comprendono: il superamento di un determinato numero di soggiorni offerti, il rispetto di soglie in merito al tasso di risposta, al tasso di cancellazione e alla valutazione complessiva ([Airbnb, n.d.](#)).

Figura 28 Composizione percentuale degli annunci in merito alla tipologia di alloggio offerto

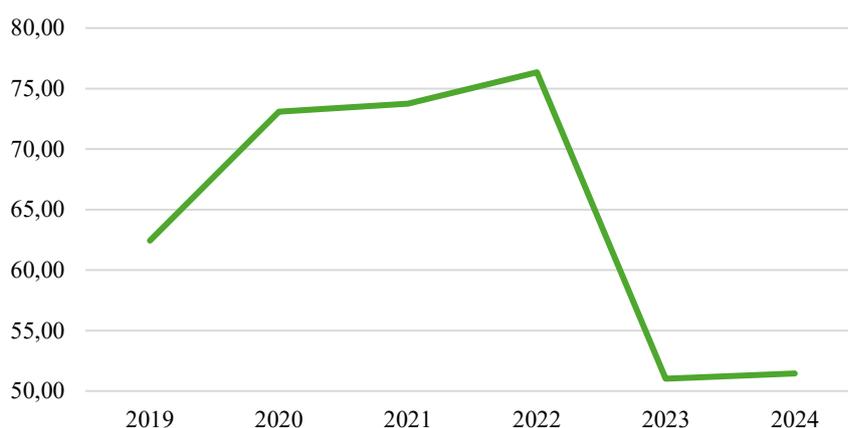


Fonte: Elaborazione dei dati del dataset pulito, dataset fornito dalla relatrice Elisabetta Raguseo (Raguseo, 2025).

In merito alla composizione percentuale della tipologia di alloggi offerti, si può notare che, in riferimento ad ogni anno, la maggioranza delle inserzioni fa riferimento ad interi alloggi, seguiti da camere private, Hotel ed infine camere condivise.

Dal 2019 al 2022 si può notare un aumento dell’offerta di annunci caratterizzati da un intero alloggio; tale incremento potrebbe essere legato al diffondersi della pandemia da Covid-19 nel 2020. Gli host, probabilmente, per speriare il diffondersi del corona virus potrebbero aver posto l’accento sull’offerta di interi appartamenti per ridurre le occasioni di contatto e conseguentemente di contagio.

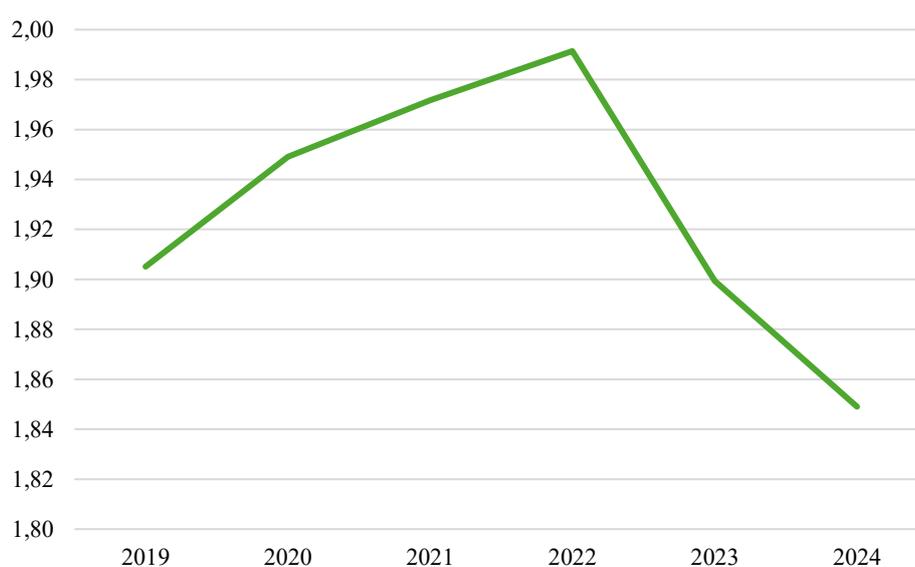
Figura 29 Andamento del valor medio delle spese di pulizia (\$)



Fonte: Elaborazione dei dati del dataset pulito, dataset fornito dalla relatrice Elisabetta Raguseo (Raguseo, 2025).

L'ammontare del valor medio delle spese di pulizia presenta un andamento crescente nel periodo tra il 2019 ed il 2022, per poi subire un drastico calo in corrispondenza del 2023. Il trend positivo che caratterizza gli anni dal 2019 al 2022 potrebbe essere spiegato dall'insinuarsi della pandemia da Covid-19 nel 2020; a seguito dell'insinuarsi della pandemia mondiale da Covid-19, gli host potrebbero aver reagito a tale shock esogeno incrementando le spese di pulizia per garantire determinati standard di pulizia e, dunque, ridurre il rischio di contagio.

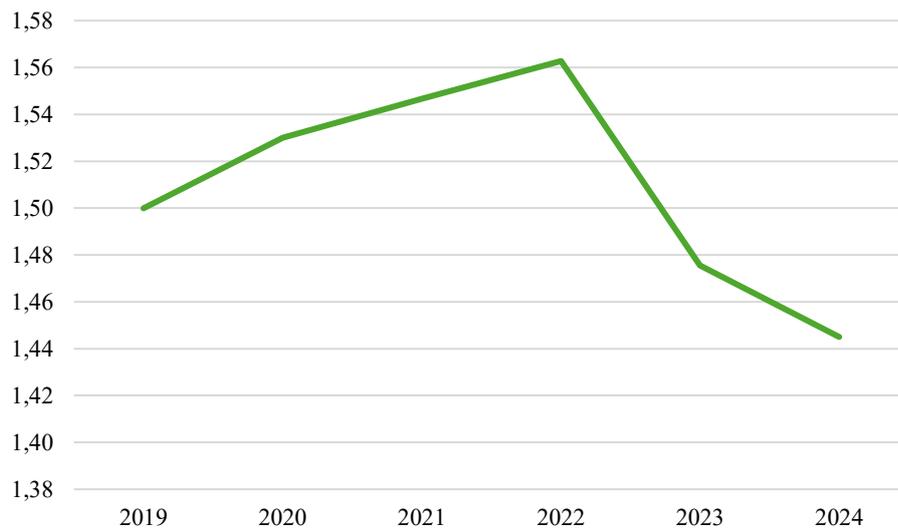
Figura 30 Andamento del valor medio del numero di camere da letto offerte nell'annuncio



Fonte: Elaborazione dei dati del dataset pulito, dataset fornito dalla relatrice Elisabetta Raguseo (Raguseo, 2025).

Il numero medio di camere da letto aumenta dal 2019 fino al 2022, dove raggiunge il suo picco massimo, dopodiché il trend è decrescente. Sebbene dal grafico si evinca una forte diminuzione del valore associato al numero medio di camere da letto offerte negli annunci, la variazione risulta essere di pochi decimali, infatti, si passa da un valore di picco del 2022 pari a 1,99 ad 1,85 nel 2024 che rappresenta il punto di minimo.

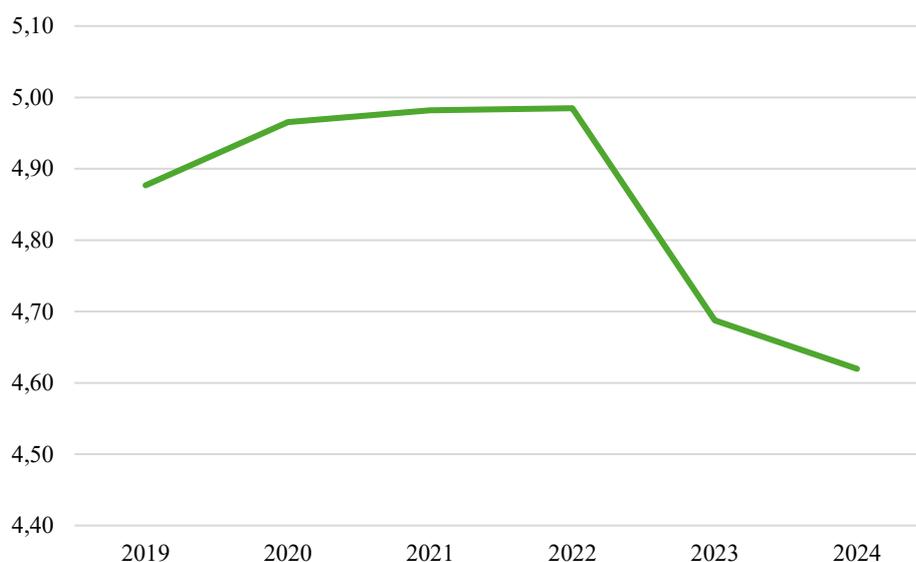
Figura 31 Andamento del valor medio del numero di bagni offerti nell'annuncio



Fonte: Elaborazione dei dati del dataset pulito, dataset fornito dalla relatrice Elisabetta Raguseo (Raguseo, 2025).

Il numero medio di bagni offerti all'interno di un annuncio presenta un andamento analogo al numero medio di camere da letto; si verifica un trend crescente dal 2019 al 2022, per poi diminuire drasticamente nel 2023 e nel 2024. In analogia con quanto riscontrato nella figura 30 ([Raguseo, 2025](#)), sebbene la figura 31 ([Raguseo, 2025](#)) mostri una forte diminuzione del valore associato al numero medio di bagni, la variazione risulta essere di pochi decimali, infatti, si passa da un valore di picco del 2022 pari a 1,56 ad 1,45 nel 2024 che rappresenta il punto di minimo.

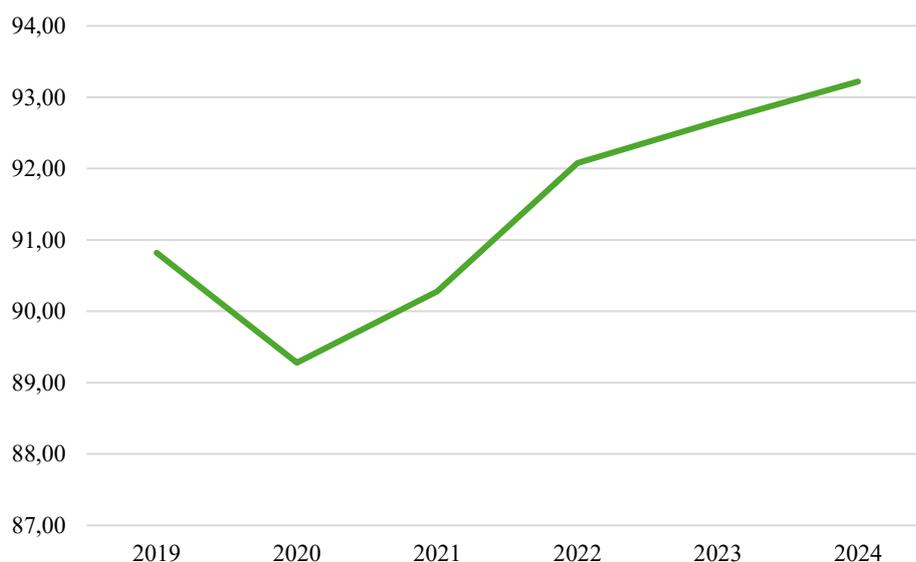
Figura 32 Andamento della media del numero massimo di ospiti



Fonte: Elaborazione dei dati del dataset pulito, dataset fornito dalla relatrice Elisabetta Raguseo (Raguseo, 2025).

L'andamento della media del numero massimo di ospiti è altalenante, ma le variazioni risultano essere minime; di fatti, il picco relativo al 2022 corrisponde al valore medio di 4,98, mentre il punto di minimo relativo al 2024 corrisponde al valore medio di 4,62.

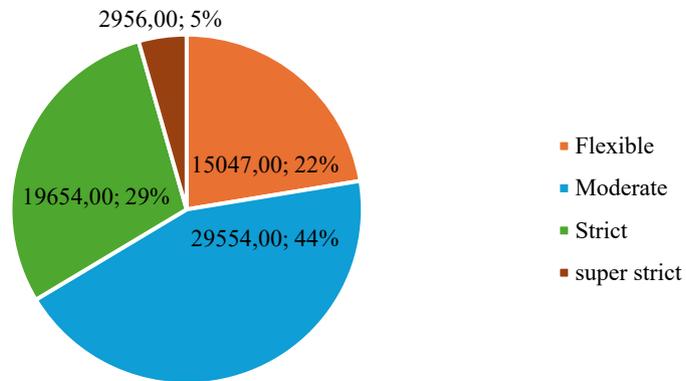
Figura 33 Andamento del valor medio del tasso di risposta (%)



Fonte: Elaborazione dei dati del dataset pulito, dataset fornito dalla relatrice Elisabetta Raguseo (Raguseo, 2025).

Il valor medio del tasso di risposta presenta un andamento crescente, simbolo, probabilmente, della crescente attenzione dedicata ai clienti.

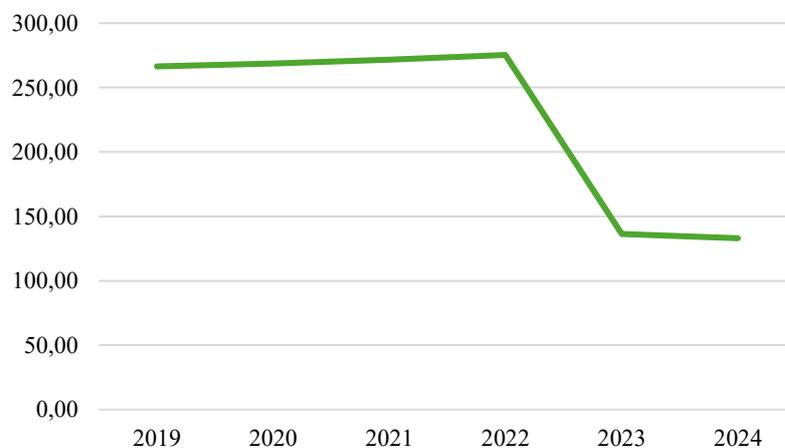
Figura 34 Composizione percentuale delle politiche di cancellazione applicate nel 2024



Fonte: Elaborazione dei dati del dataset pulito, dataset fornito dalla relatrice Elisabetta Raguseo (Raguseo, 2025).

In termini di politica di cancellazione, nel 2024, la maggioranza degli annunci fa riferimento ad una politica di cancellazione moderata, subito seguita da una politica di cancellazione ristretta, flessibile ed infine super ristretta. [Airbnb \(n.d.\)](#) prevede che passando dalla politica di cancellazione flessibile alla politica di cancellazione super ristretta, diminuisca l'arco temporale a disposizione per cancellare la prenotazione ottenendo il rimborso completo.

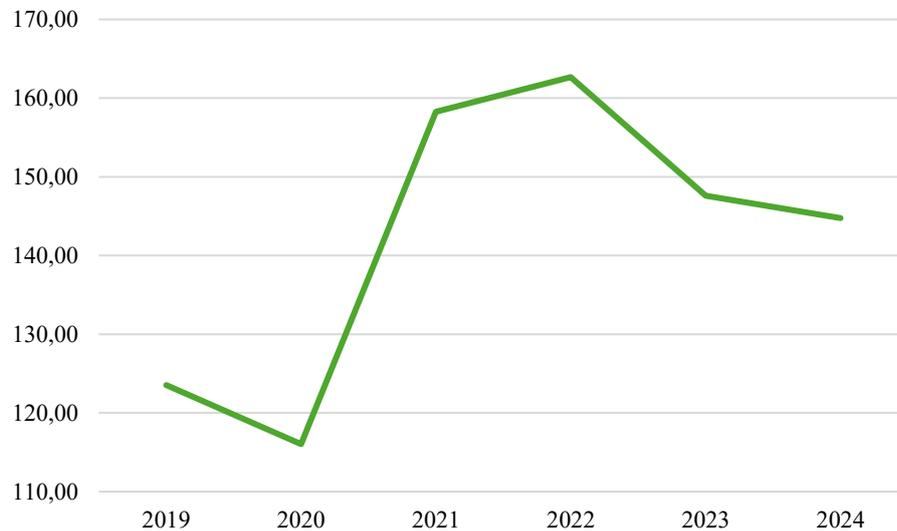
Figura 35 Andamento del valor medio del deposito di sicurezza richiesto (\$)



Fonte: Elaborazione dei dati del dataset pulito, dataset fornito dalla relatrice Elisabetta Raguseo (Raguseo, 2025).

Il valor medio del deposito di sicurezza espresso in dollari assume un valore più o meno costante tra il 2019 ed il 2022, per poi diminuire drasticamente nel 2023 e nel 2024.

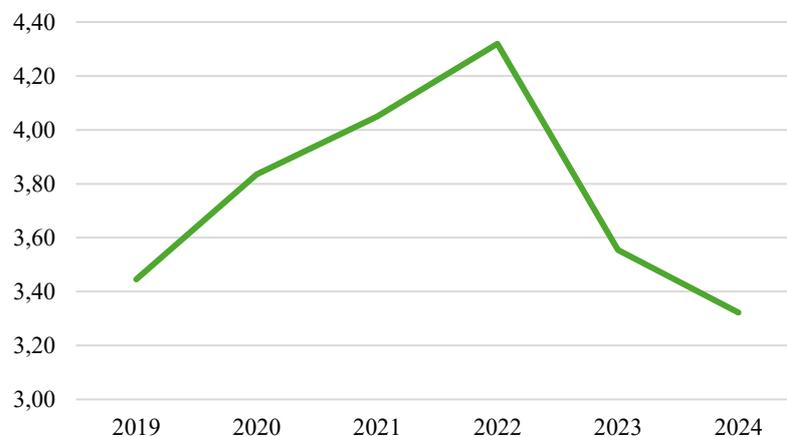
Figura 36 Andamento del valor medio del costo per notte richiesto (\$)



Fonte: Elaborazione dei dati del dataset pulito, dataset fornito dalla relatrice Elisabetta Raguseo (Raguseo, 2025).

Il valor medio del costo per notte risulta essere caratterizzato da un andamento altalenante: dal 2019 al 2020 diminuisce, probabilmente a causa dell'insidiarsi della pandemia da Covid-2019, per poi aumentare drasticamente nel 2021 e nel 2022 ed infine assumere un trend decrescente dal 2022 al 2024.

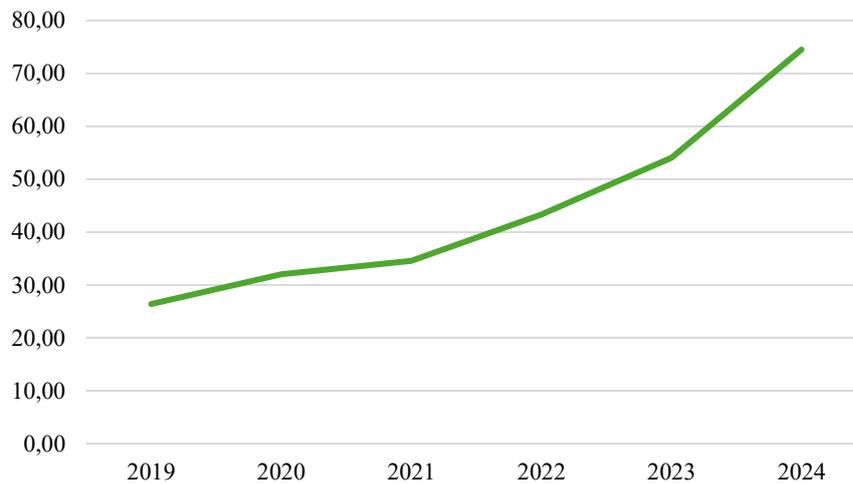
Figura 37 Andamento del valor medio del numero minimo di notti richiesto



Fonte: Elaborazione dei dati del dataset pulito, dataset fornito dalla relatrice Elisabetta Raguseo (Raguseo, 2025).

L'andamento della media del numero minimo di notti è crescente dal 2019 (anno in cui assume un valore pari a 3,45) al 2022 (anno in cui assume un valore pari a 4,32), per poi diminuire fino al 2024, anno in cui il parametro assume il valore minimo di 3,32.

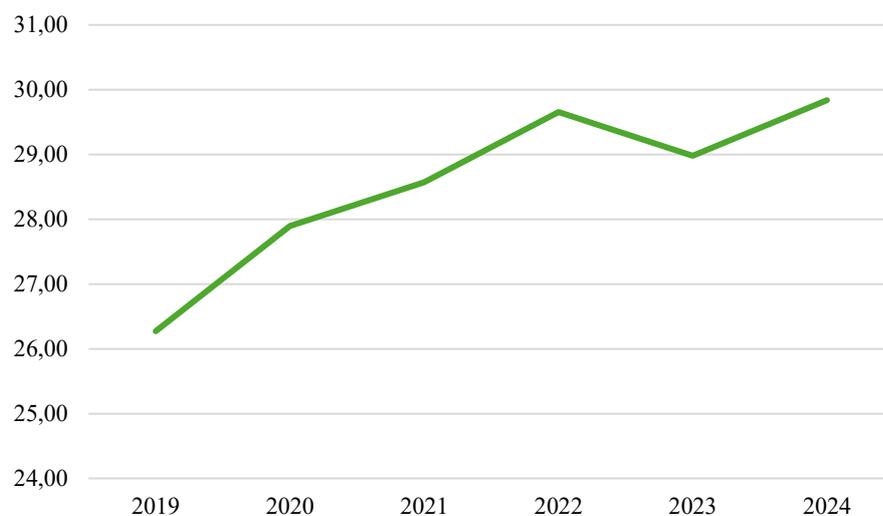
Figura 38 Andamento del valor medio del numero di recensioni possedute dagli annunci



Fonte: Elaborazione dei dati del dataset pulito, dataset fornito dalla relatrice Elisabetta Raguseo (Raguseo, 2025).

Il trend del numero medio di recensioni per annuncio cresce in modo quasi esponenziale dal 2019 al 2024, rendendo evidente la sempre maggiore partecipazione degli utenti sulla piattaforma di Airbnb.

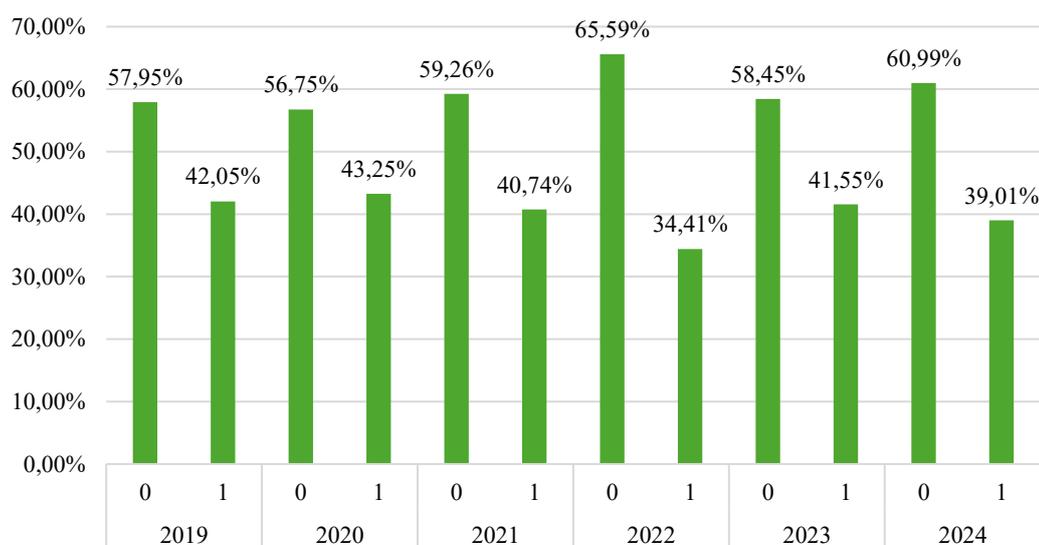
Figura 39 Andamento del numero medio di foto presenti negli annunci



Fonte: Elaborazione dei dati del dataset pulito, dataset fornito dalla relatrice Elisabetta Raguseo (Raguseo, 2025).

Il valor medio del numero di foto presenti all'interno di un annuncio aumenta dal 2019 (anno in cui assume il valore di 26,27) fino al 2024 (anno in cui assume il valore di 29,84). La presenza sempre più persistente di un numero elevato di foto potrebbe indicare una maggiore attenzione da parte degli host alla cura dell'annuncio al fine di ridurre l'asimmetria informativa presente tra host e cliente, facendo sì che il cliente possa farsi un'idea quanto più veritiera possibile del servizio offerto.

Figura 40 Composizione percentuale degli annunci in termini di possibilità di effettuare una prenotazione istantanea



Fonte: Elaborazione dei dati del dataset pulito, dataset fornito dalla relatrice Elisabetta Raguseo (Raguseo, 2025).

Il precedente grafico rappresenta la composizione percentuale in termini di annunci che prevedono o meno la possibilità di effettuare una prenotazione istantanea. Le colonne indicate con il valore “1” sono rappresentative della percentuale di annunci che prevedono la prenotazione istantanea, le colonne indicate con il valore “0”, invece, sono rappresentative della percentuale di annunci che non prevedono la prenotazione istantanea. Si evince dal grafico, dunque, che sebbene si verifichi un andamento altalenante in termini di composizione percentuale relativa alla possibilità di effettuare una prenotazione istantanea, la maggioranza degli annunci non prevede tale possibilità; suddetto risultato potrebbe essere indicativo della presenza di asimmetrie informative. L'host, dunque, non conoscendo i clienti che soggiureranno nell'alloggio offerto, preferisce riservarsi la possibilità di rifiutare la prenotazione.

3.2 Metodologia di analisi

Lo scopo della presente tesi è quello di individuare eventuali legami tra le performance dell'annuncio e le rispettive caratteristiche semantiche del titolo; a tal fine si è pensato di individuare gli argomenti maggiormente discussi all'interno dei titoli presenti in un campione del dataset pulito, per valutarne la correlazione con le variabili prestazionali. In una prima fase di analisi sono stati utilizzati algoritmi STM (structural topic modeling)⁷ implementati sul software "R-Studio" per la definizione di topic ([Barravecchia, 2024](#); [Barravecchia, 2024](#); [Barravecchia, 2024](#)); tali algoritmi, però, non sono risultati come miglior strumento di analisi a causa del numero limitato di parole presenti nei titoli degli annunci. Si è proceduto, dunque, all'utilizzo di algoritmi combinati di K-Means ed SBERT mediante l'utilizzo del software "Python". SBERT è un modello di embedding utilizzato per codificare frasi in vettori numerici ([Yilmaz, 2022](#)), mentre K-Means è un modello di clustering che associa gli embeddings a cluster in base alla distanza dal centroide mediante la minimizzazione della somma delle distanze tra ciascun punto ed il centroide del cluster assegnato ([Kavlakoglu & Winland, 2024](#)). L'algoritmo di clustering è stato realizzato da ChatGPT, un modello di intelligenza artificiale sviluppato da OpenAI ([OpenAI, 2025](#)).

⁷ L'applicazione di algoritmi STM è stata eseguita prendendo come esempio quanto appreso dal corso "Qualità nei servizi" del docente Federico Barravecchia nell'anno accademico 2023/2024 presso il Politecnico di Torino in particolare, facendo riferimento alle seguenti presentazioni PowerPoint utilizzate come ausilio didattico:

- Barravecchia, F. (2024). *Come analizzare la digital VoC: Analisi dei risultati* [Presentazione non pubblicata]. Corso "Qualità nei servizi", Politecnico di Torino.
- Barravecchia, F. (2024). *Come analizzare la digital VoC: Topic Modelling – Applicazione STM* [Presentazione non pubblicata]. Corso "Qualità nei servizi", Politecnico di Torino.
- Barravecchia, F. (2024). *Come analizzare la digital VoC: Topic Modelling – Selezione del numero di Topic da estrarre* [Presentazione non pubblicata]. Corso "Qualità nei servizi", Politecnico di Torino.
- Barravecchia, F. (2024). *Come analizzare la digital VoC: Topic Modelling – Labelling* [Presentazione non pubblicata]. Corso "Qualità nei servizi", Politecnico di Torino.

Figura 41 Algoritmo di clustering

```
|from sentence_transformers import SentenceTransformer
|from sklearn.cluster import KMeans
|import pandas as pd
|import numpy as np
|from sklearn.feature_extraction.text import TfidfVectorizer

# 📄 Carica i dati da un file Excel
df = pd.read_excel("C:\\Users\\assun\\OneDrive\\Documenti\\Datasetzenanacluster.xlsx")

# 📄 Estrai i titoli dalla colonna appropriata
titoli = df["listingtitle"].dropna().astype(str).tolist()

# 📄 Embedding con SBERT (multilingua e ottimizzato per testi brevi)
model = SentenceTransformer("distiluse-base-multilingual-cased-v2")

# 📄 Calcola gli embedding (batch-wise, usa GPU se disponibile)
embeddings = model.encode(titoli, show_progress_bar=True, device='cuda')

# 📄 KMeans clustering
NUM_TOPICS = 6 # Numero di cluster desiderati
kmeans = KMeans(n_clusters=NUM_TOPICS, random_state=42, n_init=10)
labels = kmeans.fit_predict(embeddings)

# 📄 Assegna i topic al DataFrame
df_result = df.loc[df["listingtitle"].notna()].copy()
df_result["cluster"] = labels

# 📄 Calcolo delle distanze dal centroide di ciascun cluster
distances = kmeans.transform(embeddings)

# Calcolo dei punteggi esponenziali inversi delle distanze
probabilities = np.exp(-distances)

# Normalizzazione affinché la somma delle probabilità per ogni titolo sia 1
probabilities_normalized = probabilities / probabilities.sum(axis=1, keepdims=True)

# 📄 Aggiungi le probabilità normalizzate al DataFrame
for i in range(NUM_TOPICS):
    df_result[f'prob_topic_{i}'] = probabilities_normalized[:, i]

# 📄 Parole chiave indicative per ciascun topic (facoltativo)
vectorizer = TfidfVectorizer(stop_words="english", max_features=1000)
X_text = vectorizer.fit_transform(titoli)
terms = vectorizer.get_feature_names_out()

top_keywords = []
for i in range(NUM_TOPICS):
    center = X_text[np.where(labels == i)].mean(axis=0)
    top_idx = np.argsort(center).flatten().argsort()[-10:][::-1]
    keywords = [terms[j] for j in top_idx]
    top_keywords.append(", ".join(keywords))

# 📄 Salva le parole chiave dei cluster
topic_info = pd.DataFrame({
    "cluster": list(range(NUM_TOPICS)),
    "keywords": top_keywords
})
topic_info.to_excel("KMeans_Topic_Keywords.xlsx", index=False)

# 📄 Salva il file con i titoli, i topic e le probabilità normalizzate
df_result.to_excel("Titoli_con_cluster_e_probabilita.xlsx", index=False)

print("Il processo è stato completato con successo e i file sono stati salvati.")
```

Fonte: OpenAI. (2025). ChatGPT (Versione GPT-4.5) [Modello linguistico]. <https://chat.openai.com>

Al fine di poter utilizzare l'algoritmo presente in figura 41 ([OpenAI, 2025](#)) realizzato da ChatGPT, si è ridotta la dimensione del campione a 9750 osservazioni mediante un campionamento casuale, mantenendo le proporzioni in termini della variabile "year" e della variabile "overallrating". Tale campione è sufficiente a descrivere la popolazione in modo rappresentativo: in base alla formula fornita da [Cochran, 1977](#), di fatti, al fine di ottenere un campione rappresentativo con un margine di errore all'1%, data una popolazione caratterizzata da 426.959 osservazioni, bastano circa 9.393 osservazioni. Per tale calcolo sono stati utilizzati i seguenti dati:

- Popolazione caratterizzata da 426.959 osservazioni;
- Livello di confidenza pari al 95%, dunque $z = 1,96$;
- Proporzione, che per motivi conservativi si è deciso di considerare pari a 0,5;
- Margine d'errore che in questo caso si è considerato pari a 1%.

L'algoritmo proponeva, inizialmente, un numero elevato di topic, i quali sono stati successivamente ridotti a sei e sono stati etichettati come:

- Tipologia e caratteristiche dell'alloggio: quando è prevalente questo argomento nel titolo dell'annuncio, vuol dire che il focus di quest'ultimo è sulla tipologia di alloggio offerto e/o sulle proprie caratteristiche. Le keywords associate a questo argomento sono: "room", "bedrooms", "double", "bedroom", "private", "camera", "bathroom", "view", "cozy", "bed". Alcuni esempi di titoli ad alta prevalenza di questo argomento sono:
 - "Beautiful double room";
 - "WFP/FAO/ICCROM/IFAD / Univ Roma3 /WFP single room".
- Personalizzazione del titolo: quando prevale questo argomento nel titolo dell'annuncio, vuol dire che in esso sono contenute informazioni prevalentemente circa appartamenti, suite oppure personalizzazioni dell'alloggio stesso (sia in termini di proprietà dell'alloggio mediante la definizione del nome, sia in termini di caratteristiche), come:
 - "Penthouse Apartment in Central Rome";

- “Kandinsky flat”.

Le keywords associate a questo argomento sono: “apartment”, “flat”, “appartamento”, “cozy”, “terrace”, “view”, “near”, “center”, “luxury”, “apartments”.

- Comfort: quando prevale questo argomento, significa che nel titolo dell’annuncio sono presenti caratteristiche circa i comfort offerti dall’alloggio e/o riferimenti ad alloggi di lusso come le ville, ad esempio:

- “Villa W POOL Siena/Casole 12 People”;
- “Tropea: Villa near sea in quiet area, pool garden”.

Le keywords associate a questo argomento sono: “villa”, “pool”, “sea”, “tuscany”, “garden”, “near”, “view”, “Tuscan”, “Florence”, “country”.

- Posizione: quando prevale questo argomento, significa che nel titolo dell’annuncio sono presenti indicazioni circa la localizzazione dell’alloggio, come:

- “Circo Massimo hystorical Rome apt sleeps 5 people”;
- “Rione Monti, next to Foro Romano”.

Le keywords associate a questo argomento sono: “studio”, “terrace”, “apt”, “view”, “loft”, “center”, “rome”, “heart”, “ca”, “suite”.

- Tipicità abitativa: quando prevale questo argomento, significa che il titolo dell’annuncio fa principalmente riferimento ad alloggi caratteristici, loft, trulli, appartamenti unici, definendo in modo differente le caratteristiche, soffermandosi in particolare su stile/atmosfera/esperienze, come:

- “IL "CASTELLO" SUL LAGO MAGGIORE”;
- “I SETTE CONI - TRULLO EDERA”.

Le keywords associate a questo argomento sono: “ca”, “apartment”, “wifi”, “rialto”, “apt”, “suite”, “marco”, “terrace”, “la”, “trullo”.

- Zone rurali: quando prevale questo argomento, significa che nel titolo dell'annuncio sono presenti indicazioni prevalentemente in merito a case con caratteristiche rurali e riferimenti al relax, come:

- “Casa Il Capanno”;
- “Casa montana per famiglie e amici”.

Le keywords associate a questo argomento sono: “house”, “casa”, “home”, “sea”, “farmhouse”, “view”, “country”, “cottage”, “holiday”, “beach”.

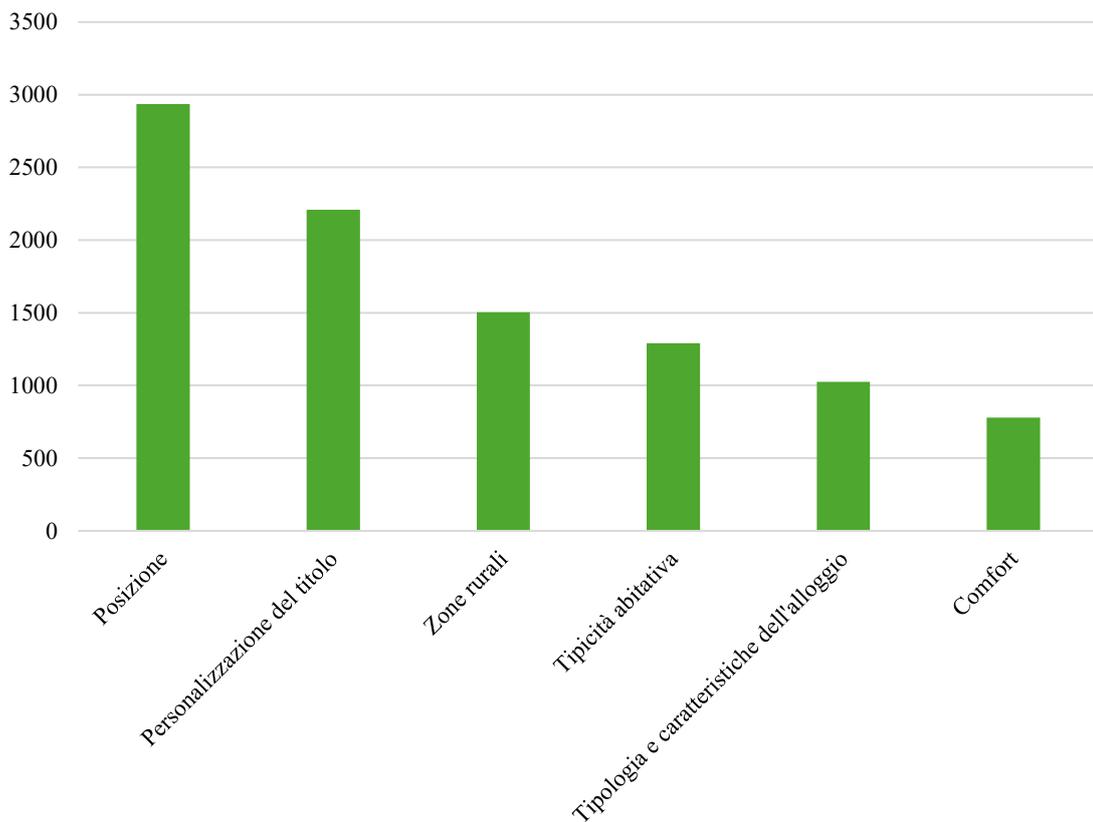
In genere, nei modelli K-Means, il valore associato ad ogni topic rappresenta la distanza dell'osservazione dal centroide di quel determinato topic ([Kavlakoglu & Winland, 2024](#)); dunque, per ottenere dei valori assimilabili a delle probabilità è stato normalizzato ad uno l'inverso esponenziale delle distanze ($e^{-\text{distanza}}$) ([OpenAI, 2025](#)). Si ottengono, in fine, due file di output:

- Un file Excel contenente le parole chiave associate ad ogni topic;
- Un file Excel in cui, per ogni titolo, sono presenti, oltre ai metadati del dataset originale, la probabilità di appartenere a ciascun topic e l'assegnazione dei titoli ad un determinato topic.

L'assegnazione delle etichette ad ogni topic è stata realizzata facendo riferimento congiuntamente a due elementi:

- Le parole chiave associate ad ogni topic;
- La lettura dei titoli aventi la maggior probabilità di contenere un determinato topic: ovvero, ad ogni titolo viene assegnato dall'algorithm il topic che mostra la probabilità maggiore di essere presente in quel determinato titolo, dopodiché si filtrano i titoli in funzione dei topic assegnati e, a seguito della loro lettura, si propone un'etichetta per il topic.

Figura 42 Ordinamento dei topic in termini di discussione all'interno del dataset



Fonte: Elaborazione dei dati del campione del dataset pulito, dataset fornito dalla relatrice Elisabetta Raguseo (Raguseo, 2025).

Associando ad ogni titolo del campione del dataset il topic avente il valore maggiore di probabilità, è stato possibile ottenere una classificazione degli argomenti maggiormente discussi. Come si evince dalla figura 42 (Raguseo, 2025) il topic maggiormente discusso è “Posizione”, mentre il topic meno discusso è quello intitolato “Comfort”.

3.5 Sentiment Analysis

Al fine di realizzare la Sentiment Analysis è stato utilizzato l’algoritmo sotto riportato realizzato da ChatGPT, un modello di intelligenza artificiale sviluppato da OpenAI (OpenAI, 2025), e inizializzato sul software Python.

Figura 43 Algoritmo di clusterizzazione dei titoli per lo svolgimento della Sentiment Analysis in termini di: Positive, Negative e Neutral

```
|from transformers import pipeline
import pandas as pd

# Carica il file Excel con i titoli
df = pd.read_excel("Titoli_con_cluster_e_probabilita.xlsx")
titles = df["listingtitle"].dropna().astype(str).tolist()

# Inizializza la pipeline (esempio con modello multilingue)
classifier = pipeline("sentiment-analysis", model="nlptown/bert-base-multilingual-uncased-sentiment")

# Passa la lista completa (batch) direttamente alla pipeline
results = classifier(titles)

# Estrai le info in liste
labels = []
scores = []
values = []

for result in results:
    label = result['label']
    score = round(result['score'], 3)
    star = int(label[0])

    if star <= 2:
        sentiment = "negative"
        value = -1
    elif star == 3:
        sentiment = "neutral"
        value = 0
    else:
        sentiment = "positive"
        value = 1

    labels.append(sentiment)
    scores.append(score)
    values.append(value)

# Aggiungi le colonne al DataFrame
df["sentiment_label"] = labels
df["sentiment_score"] = scores
df["sentiment_value"] = values

# Salva su Excel
df.to_excel("titoli_con_sentiment_batch.xlsx", index=False)

print("✅ Elaborazione completata, file salvato!")
```

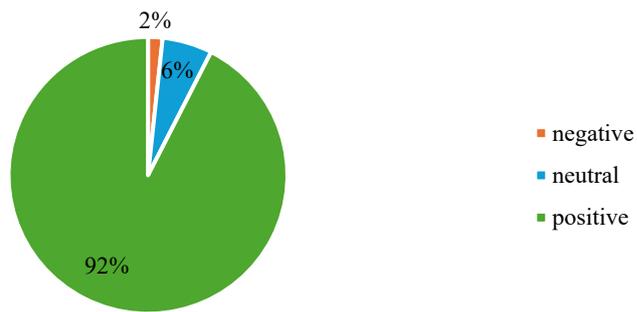
Fonte: OpenAI. (2025). ChatGPT (Versione GPT-4.5) [Modello linguistico]. <https://chat.openai.com>

L'algoritmo, analizzando i testi dei titoli degli annunci, assegna un punteggio in termini di stelle ad ogni annuncio, dopodiché il punteggio in stelle viene riconvertito in:

- Positivo: se l'algoritmo associa al titolo un numero di stelle maggiore di tre;
- Neutro: se l'algoritmo associa al titolo un numero di stelle pari a tre;
- Negativo: se l'algoritmo associa al titolo un numero di stelle inferiore a tre.

Dall'analisi del sottoinsieme di annunci, risulta che:

Figura 44 Composizione percentuale dei titoli del campione del dataset pulito in termini di sentiment

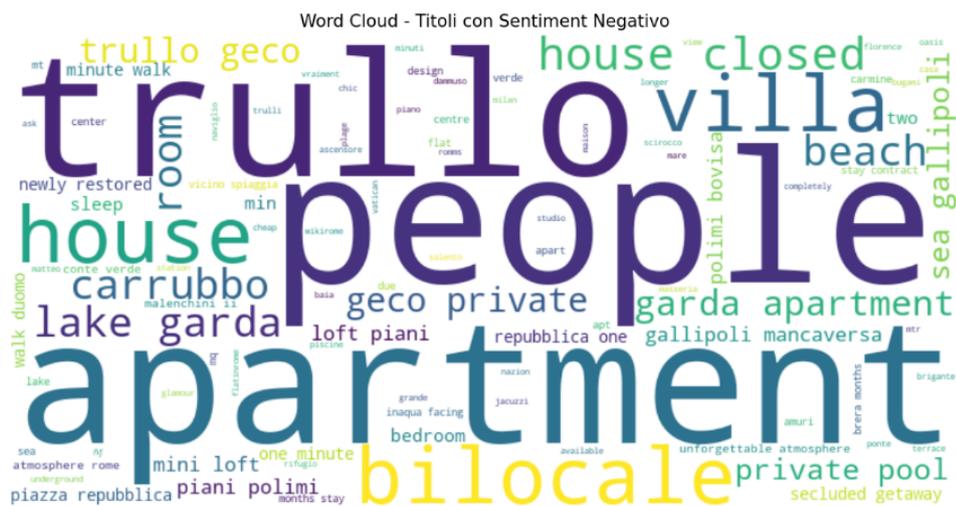


Fonte: Elaborazione dei dati del campione del dataset pulito, dataset fornito dalla relatrice Elisabetta Raguseo (Raguseo, 2025).

In accordo con quanto riscontrato da [Canziani e Nemati \(2021\)](#), la maggioranza dei titoli degli annunci presentano toni positivi, seguiti poi da titoli caratterizzati da toni neutri ed infine da titoli caratterizzati da toni negativi.

Risulta essere d'impatto la presenza, riscontrata dall'algoritmo, di titoli di annunci caratterizzati da un sentiment negativo; dunque, analizzandoli nel dettaglio si è deciso di definire la seguente nuvola di parole:

Figura 45 Nuvola di parole associata ai titoli degli annunci classificati come aventi sentiment negativo



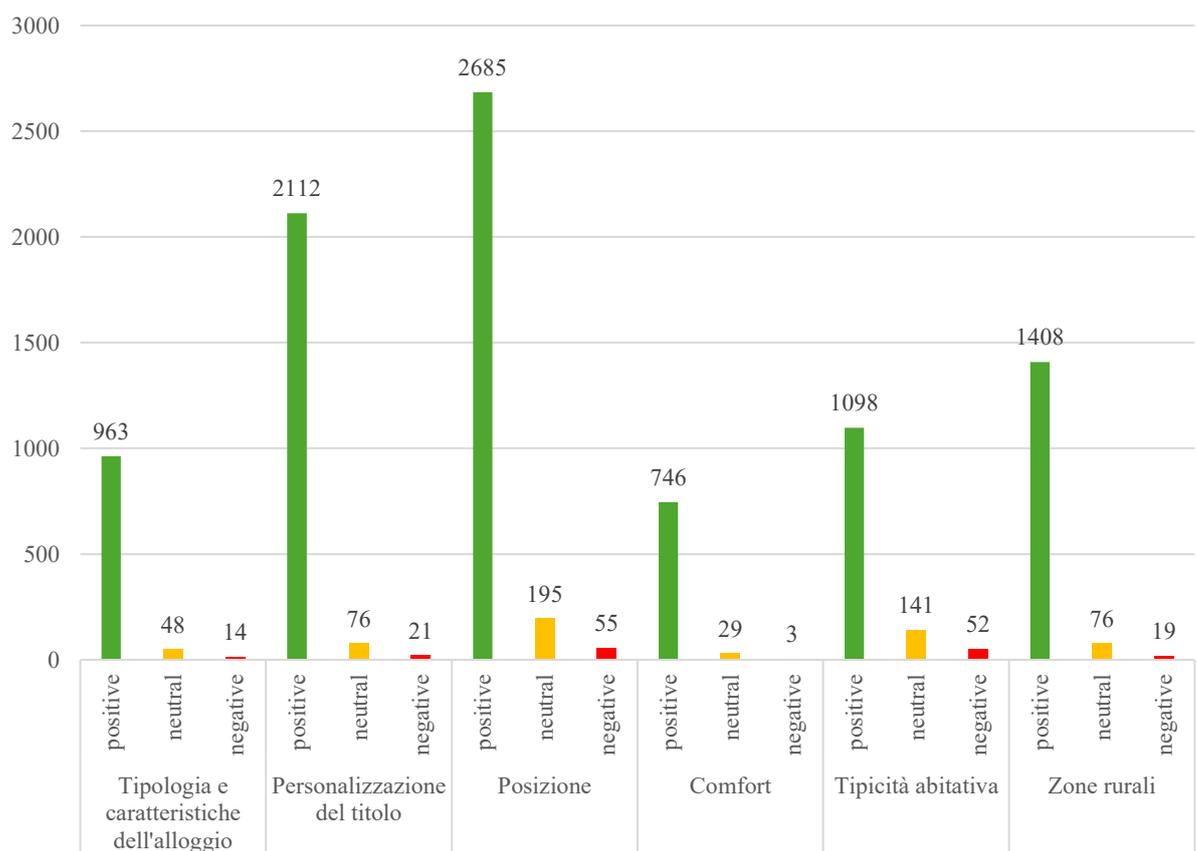
Fonte: Elaborazione dei dati del campione del dataset pulito, dataset fornito dalla relatrice Elisabetta Raguseo (Raguseo, 2025). Screen effettuato dal software Python.

È evidente dalla nuvola di parole che, sebbene 164 annunci siano stati classificati come caratterizzati da sentiment negativo, questi rappresentino un outlier in termini di accuratezza dell’algoritmo. Alcuni esempi di titoli classificati come negativi sono:

- “No longer available on this platform, ask me why!”;
- “Chic&Cheap only 12 min to S.Peter”;
- “Mini Loft 2 piani PoliMI Bovisa”.

Si può notare come, probabilmente, la classificazione errata derivi presumibilmente da una non corretta interpretazione delle parole, ad esempio: nel primo titolo “No” potrebbe essere stata percepita come una parola negativa o ancora “Cheap” nel secondo titolo o “Mini” nel terzo titolo.

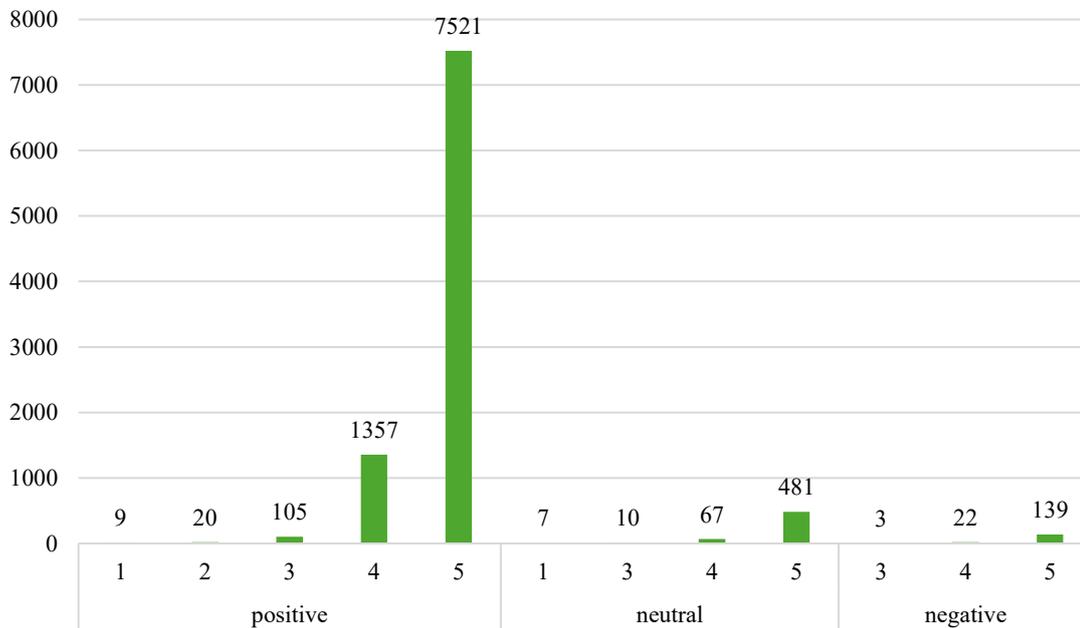
Figura 46 Distribuzione del sentiment dei titoli degli annunci suddivisi per topic



Fonte: Elaborazione dei dati del campione del dataset pulito, dataset fornito dalla relatrice Elisabetta Raguseo (Raguseo, 2025).

In linea con quanto definito in precedenza, i titoli associati ad ogni topic sono per lo più positivi, seguiti da una minoranza di titoli neutri.

Figura 47 Distribuzione del sentiment dei titoli degli annunci suddivisi per valutazione (overallrating)



Fonte: Elaborazione dei dati del campione del dataset pulito, dataset fornito dalla relatrice Elisabetta Raguseo (Raguseo, 2025).

Come si evince dalla figura 47 (Raguseo, 2025), la valutazione in termini di stelle (relativa alla colonna overallrating) non presenta vincoli in termini sentiment, ovvero: sia che l'annuncio sia caratterizzato da un sentiment positivo, sia che sia caratterizzato da un sentiment neutro, sono presenti annunci che godono di una valutazione che va da una a cinque stelle.

3.6 Analisi di regressione

L'analisi di regressione è stata svolta mediante l'utilizzo di modelli OLS (Ordinary Least Squares), in particolare si è deciso di utilizzare le seguenti funzioni:

$$ADR_{it} = \beta_0 + \beta_1 T1_{it} + \beta_2 T2_{it} + \beta_3 T3_{it} + \beta_4 T4_{it} + \beta_5 T5_{it} + \beta_6 X + \varepsilon_{it}$$

$$revPAN_{it} = \beta_0 + \beta_1 T1_{it} + \beta_2 T2_{it} + \beta_3 T3_{it} + \beta_4 T4_{it} + \beta_5 T5_{it} + \beta_6 X + \varepsilon_{it}$$

Dove:

- $ADR = \frac{revenue\ usd}{reservation\ days}$
- $revPAN = \frac{revenue\ usd}{reservation\ days + available\ days}$
- T1 = topic “Posizione”
- T2 = topic “Personalizzazione del titolo”
- T3 = topic “Zone rurali”
- T4 = topic “Tipologia e caratteristiche dell'alloggio”
- T5 = topic “Comfort”
- “X” è un vettore contenente tutte le variabili di controllo: $Entirehomeapt_{it}$, $Privateroom_{it}$, $Sharedroom_{it}$, $Bedrooms_{it}$, $Bathrooms_{it}$, $Urbana_{it}$, $Rurale_{it}$, $Airbnbsuperhost_{it}$, $Responcerate_{it}$, $Maxguests_{it}$, $Minimumstay_{it}$, $ln_numberofphotos_{it}$, $Overallrating_{it}$, $Instantbookenabled_{it}$, $ln_numberofreviews_{it}$, $numero_caratteri_{it}$.

Il pedice “i” fa riferimento all’i-esimo annuncio ed il pedice “t” fa riferimento alla variazione rispetto all’anno (indicato nel database con “years”); in quanto nel campione del dataset pulito sono presenti titoli di annunci che variano dal 2019 al 2024. Il termine “ β_0 ” rappresenta il termine costante della funzione di regressione, ovvero il valore che assume la variabile dipendente quando tutte le altre variabili assumono un valore nullo. Infine, il termine ε_{it} rappresenta il termine di errore del modello, ovvero tutto ciò che non viene spiegato dalla funzione di regressione.

I modelli si propongono come strumento di analisi della correlazione tra le caratteristiche semantiche del titolo dell’annuncio e le performance dell’annuncio stesso. La variabile dipendente “ADR” rappresenta il ricavo giornaliero medio (durante la definizione della variabile ADR, quando la variabile “revenue usd” assumeva valore nullo allora ADR è stata posta uguale a zero, in modo analogo quando “reservation days” assumeva valore nullo anche ADR è stata posta uguale a zero, infine, quando sia il numeratore sia il

denominatore hanno assunto un valore nullo allora anche ADR è stata posta uguale a zero). La variabile dipendente “revPAN” rappresenta il ricavo medio per giorni disponibili, in quanto i ricavi sono divisi per la somma dei giorni prenotati e disponibili, ovvero la totalità dei giorni messi a disposizione dall’host. I topic “Posizione”, “Personalizzazione del titolo”, “Zone rurali”, “Tipologia e caratteristiche dell’alloggio” e “Comfort” rappresentano le variabili indipendenti. Nei modelli sono presenti cinque dei sei topic individuati per evitare problemi di collinearità, in quanto la somma della probabilità della presenza dei topic nel titolo di un annuncio è pari a uno; dunque, sono legati da una relazione lineare.

Le altre variabili inserite nei modelli rappresentano le variabili di controllo e sono:

- Entirehomeapt, Privateroom e Sharedroom: sono delle variabili dummy utilizzate per comprendere se ci sono delle variazioni in funzione della tipologia di alloggio. Per evitare di cadere nella “trappola delle dummy” sono state inserite tre delle quattro categorie presenti nel campione del database, in modo tale che se tutte le dummy assumono un valore nullo allora vuol dire che la tipologia di annuncio fa riferimento alla categoria “Hotel”;
- Bedrooms: è una variabile che indica il numero di camere da letto;
- Bathrooms: è una variabile che indica il numero di bagni;
- Airbnbsuperhost: è una variabile dummy che assume un valore pari a “1” se l’host possiede il badge “superhost”, “0” altrimenti;
- Responce rate: rappresenta il tasso di risposta dell’host;
- Urbana, Rurale: sono delle variabili dummy utilizzate per comprendere se sono presenti delle variazioni in funzione della tipologia di zona in cui è situato l’alloggio. Per evitare di cadere nella “trappola delle dummy” sono state inserite due delle tre categorie presenti nel campione del database, in modo tale che se tutte le dummy assumono un valore nullo allora vuol dire che la tipologia di zona fa riferimento alla categoria “Intermedia”;
- Maxguests: rappresenta il numero massimo di ospiti accettati;
- Minimumstay: rappresenta il numero minimo di notti necessarie per effettuare la prenotazione;
- ln_numberofphotos: rappresenta il logaritmo naturale del numero di foto associate all’annuncio. Si è deciso di utilizzare il logaritmo naturale del numero di foto e

non direttamente il numero di foto, per ottenere una distribuzione dei valori quanto più possibile simile ad una gaussiana;

- `ln_numberofreviews`: rappresenta il logaritmo naturale del numero di recensioni associate all'annuncio. Si è deciso di utilizzare il logaritmo naturale del numero di recensioni e non direttamente il numero di recensioni, per ottenere una distribuzione dei valori quanto più possibile simile ad una gaussiana;
- `Overallrating`: rappresenta il rating associato ad un annuncio e varia tra un minimo di una stella ed un massimo di cinque stelle;
- `Instantbookenabled`: rappresenta una variabile dummy che assume un valore pari ad "1" se l'annuncio prevede la possibilità di effettuare una prenotazione istantanea, "0" altrimenti;
- `Numero_caratteri`: rappresenta il numero di caratteri che compongono il titolo dell'annuncio.

È possibile clusterizzare le variabili di controllo in tre macrogruppi per poter analizzare come le differenti caratteristiche controllino la relazione tra le performance dell'annuncio e le sue caratteristiche semantiche:

- `Entirehomeapt`, `Privateroom`, `Sharedroom`, `Bedrooms`, `Bathrooms`, `Urbana`, `Rurale`: sono delle variabili che permettono di comprendere l'impatto delle caratteristiche dell'offerta sulle performance;
- `Airbnbsuperhost`, `Responce rate`: sono variabili che permettono di comprendere l'impatto delle caratteristiche dell'host sulle performance;
- `Maxguests`, `Minimumstay`, `ln_numberofphotos`, `Overallrating`, `Instantbookenabled`, `ln_numberofreviews`, `numero_caratteri`: sono delle variabili che permettono di comprendere l'impatto delle caratteristiche dell'annuncio sulle performance.

La suddivisione delle zone in "Urbana", "Intermedia" e "Rurale" è stata realizzata mediante il download di dataset forniti dall'ISTAT circa la popolazione residente al primo gennaio 2025 per ciascun comune a cui si riferisce l'annuncio⁸, ed il download di un

⁸ Nella sezione "Riferimenti" sono contenuti tutti i riferimenti comprensivi di link dei dataset scaricati per ciascuna regione.

dataset fornito dall'ISTAT circa i dati sulla superficie di ogni comune ([ISTAT, 2024](#)) espressi in kmq. Una volta ottenuti i dati, è stata calcolata la densità di popolazione intesa come il rapporto tra il numero di residenti e la superficie. Sono stati, dunque, considerati:

- Rurali, gli annunci Airbnb riferiti ad alloggi situati in comuni con densità di popolazione inferiore al trentatreesimo percentile;
- Intermedi, gli annunci Airbnb riferiti ad alloggi situati in comuni con densità di popolazione compresa tra il trentatreesimo percentile e il sessantasettesimo percentile;
- Urbani, gli annunci Airbnb riferiti ad alloggi situati in comuni con densità di popolazione superiore al sessantasettesimo percentile.

Tabella 4 Sommario delle caratteristiche delle variabili

	obs	mean	sd	min	max
year	9741	2021.488	1.788907	2019	2024
numberofreservations	9741	22.74109	28.95093	0	245
revenueusd	9741	14614.36	28247.66	0	608651
reservationdays	9741	74.14198	82.84747	0	363
Urbana	9741	.4141259	.4925957	0	1
Rurale	9741	.3344626	.4718265	0	1
bedrooms	9741	1.883072	1.52022	0	50
bathrooms	9741	1.560261	1.114683	0	22
maxguests	9741	4.701468	2.733469	1	16
responserate	9741	93.4902	19.54625	0	100
airbnbsuperhost	9741	.2797454	.4488971	0	1
minimumstay	9741	4.156041	13.04194	1	365
numberofreviews	9741	60.23262	95.79272	0	1138
numberofphotos	9741	29.50323	18.00809	1	234
instantbookenabled	9741	.4109434	.4920303	0	1
overallrating	9741	4.687597	.3978403	1	5
ln_numberofreviews	9741	2.983908	1.65846	0	7.037028
ln_numberofphotos	9741	3.220119	.5859343	0	5.455321

Entirehomeapt	9741	.8849194	.3191356	0	1
Privateroom	9741	.1073812	.3096131	0	1
Sharedroom	9741	.0016425	.0404971	0	1
Tipologiaecaratteristiche	9741	.1666102	.0074932	.1509927	.2008773
Personalizzazione del titolo	9741	.1669646	.0119808	.1499137	.2058704
Posizione	9741	.1726341	.0053649	.1569643	.1893003
Comfort	9741	.1601451	.0103522	.1434394	.208367
Tipicità abitativa	9741	.1666241	.010506	.1492824	.2188609
Zonerurali	9741	.1670219	.0084129	.1555144	.2042405
numero_caratteri	9741	34.87373	10.88553	4	127
revPAN	9741	67.17186	122.4843	0	2876.527
ADR	9741	166.7238	203.6058	0	3414.319

Fonte: File di Stata dell'elaborazione dei dati del campione del dataset pulito, dataset fornito dalla relatrice Elisabetta Raguseo (Raguseo, 2025).

È stata realizzata inizialmente una regressione contenente solo le variabili di controllo, per poi inserire un topic alla volta e valutarne il singolo legame con la variabile dipendente. Verrà analizzato dapprima il modello avente come variabile dipendente ADR, dopodiché verrà analizzato il modello avente come variabile dipendente revPAN.

Tabella 5 Regressione caratterizzata da un modello con solo variabili di controllo con variabile dipendente ADR

ADR	Coef.	St.Err.	t-value	p-value	[95% Conf	Interval]	Sig
bedrooms	-9.439	10.864	-0.87	.385	-30.734	11.856	
bathrooms	79.484	9.628	8.26	0	60.611	98.357	***
maxguests	18.464	3.851	4.80	0	10.916	26.013	***
responserate	.035	.062	0.56	.573	-.087	.157	
airbnbsuperhost	23.763	4.289	5.54	0	15.356	32.17	***
minimumstay	-.04	.07	-0.57	.569	-.176	.097	
ln_numberofreviews	-2.045	1.181	-1.73	.083	-4.36	.27	*
ln_numberofphotos	6.011	3.431	1.75	.08	-.715	12.736	*
instantbookenabled	33.067	3.408	9.70	0	26.388	39.747	***
overallrating	23.838	6.821	3.49	0	10.468	37.208	***
Entirehomeapt	-25.671	33.184	-0.77	.439	-90.718	39.376	
Privateroom	-60.029	32.825	-1.83	.067	-124.373	4.315	*
Sharedroom	-62.868	39.849	-1.58	.115	-140.98	15.244	
Urbana	-31.007	4.304	-7.20	0	-39.444	-22.569	***

Rurale	-49.052	5.242	-9.36	0	-59.328	-38.777	***
numero_caratteri	.581	.167	3.47	.001	.253	.908	***
Constant	-136.664	47.944	-2.85	.004	-230.645	-42.683	***

Mean dependent var	166.487	SD dependent var	202.522
R-squared	0.388	Number of obs	9700
F-test	98.911	Prob > F	0.000
Akaike crit. (AIC)	125830.204	Bayesian crit. (BIC)	125952.262

*** $p < .01$, ** $p < .05$, * $p < .1$

Fonte: File di Stata dell'elaborazione dei dati del campione del dataset pulito, dataset fornito dalla relatrice Elisabetta Raguseo (Raguseo, 2025).

Tabella 6 Regressione caratterizzata da un modello con variabili di controllo ed il topic "Tipologia e caratteristiche dell'alloggio" con variabile dipendente ADR

ADR	Coef.	St.Err.	t-value	p-value	[95% Conf	Interval]	Sig
Tipologiae caratteristiche	74.042	217.359	0.34	.733	-352.028	500.111	
bedrooms	-9.453	10.866	-0.87	.384	-30.752	11.846	
bathrooms	79.503	9.641	8.25	0	60.605	98.401	***
maxguests	18.475	3.848	4.80	0	10.933	26.017	***
responserate	.034	.062	0.55	.58	-.088	.156	
airbnbsuperhost	23.76	4.288	5.54	0	15.354	32.166	***
minimumstay	-.04	.07	-0.57	.566	-.176	.096	
ln_numberofreviews	-2.058	1.184	-1.74	.082	-4.379	.263	*
ln_numberofphotos	6.088	3.434	1.77	.076	-.643	12.819	*
instantbookenabled	33.023	3.393	9.73	0	26.373	39.674	***
overallrating	23.876	6.824	3.50	0	10.499	37.253	***
Entirehomeapt	-24.766	33.73	-0.73	.463	-90.883	41.351	
Privateroom	-59.882	32.879	-1.82	.069	-124.332	4.567	*
Sharedroom	-62.582	39.929	-1.57	.117	-140.851	15.686	
Urbana	-31.148	4.303	-7.24	0	-39.583	-22.713	***
Rurale	-49.008	5.258	-9.32	0	-59.314	-38.701	***
numero_caratteri	.571	.169	3.37	.001	.239	.904	***
Constant	-149.82	65.44	-2.29	.022	-278.095	-21.544	**

Mean dependent var	166.487	SD dependent var	202.522
R-squared	0.388	Number of obs	9700
F-test	93.214	Prob > F	0.000
Akaike crit. (AIC)	125832.121	Bayesian crit. (BIC)	125961.359

*** $p < .01$, ** $p < .05$, * $p < .1$

Fonte: File di Stata dell'elaborazione dei dati del campione del dataset pulito, dataset fornito dalla relatrice Elisabetta Raguseo (Raguseo, 2025).

In tabella 6 (Raguseo, 2025) si evince che il topic “Tipologia e caratteristiche dell'alloggio” è correlato positivamente con la variabile dipendente (in quanto il coefficiente associato al topic è positivo), ma non è statisticamente significativo (in quanto $p\text{-value}=0,733$).

Tabella 7 Regressione caratterizzata da un modello con variabili di controllo ed il topic “Personalizzazione del titolo” con variabile dipendente ADR

ADR	Coef.	St.Err.	t-value	p-value	[95% Conf	Interval]	Sig
Personalizzazionedeltitolo	-218.695	123.81	-1.77	.077	-461.388	23.997	*
bedrooms	-9.389	10.836	-0.87	.386	-30.63	11.851	
bathrooms	79.342	9.631	8.24	0	60.464	98.22	***
maxguests	18.432	3.836	4.81	0	10.913	25.951	***
responserate	.037	.062	0.60	.55	-.085	.159	
airbnbsuperhost	23.744	4.287	5.54	0	15.339	32.148	***
minimumstay	-.041	.069	-0.60	.549	-.177	.094	
ln_numberofreviews	-2.006	1.181	-1.70	.089	-4.322	.309	*
ln_numberofphotos	5.467	3.421	1.60	.11	-1.239	12.174	
instantbookenabled	33.208	3.408	9.74	0	26.528	39.888	***
overallrating	23.868	6.815	3.50	0	10.509	37.228	***
Entirehomeapt	-24.034	33.234	-0.72	.47	-89.179	41.112	
Privateroom	-59.873	32.839	-1.82	.068	-124.245	4.499	*
Sharedroom	-62.582	40.026	-1.56	.118	-141.041	15.878	
Urbana	-30.38	4.278	-7.10	0	-38.765	-21.994	***
Rurale	-49.355	5.234	-9.43	0	-59.614	-39.095	***
numero_caratteri	.601	.167	3.60	0	.274	.929	***
Constant	-100.963	51.191	-1.97	.049	-201.308	-.618	**

Mean dependent var	166.487	SD dependent var	202.522
R-squared	0.388	Number of obs	9700
F-test	93.079	Prob > F	0.000
Akaike crit. (AIC)	125829.759	Bayesian crit. (BIC)	125958.997

*** $p < .01$, ** $p < .05$, * $p < .1$

Fonte: File di Stata dell'elaborazione dei dati del campione del dataset pulito, dataset fornito dalla relatrice Elisabetta Raguseo (Raguseo, 2025).

In tabella 7 ([Raguseo, 2025](#)) si evince che il topic “Personalizzazione del titolo” è correlato negativamente con la variabile dipendente in quanto il coefficiente associato al topic è negativo, ed è statisticamente significativo al 10%.

Tabella 8 Regressione caratterizzata da un modello con variabili di controllo ed il topic “Posizione” con variabile dipendente ADR

ADR	Coef.	St.Err.	t-value	p-value	[95% Conf	Interval]	Sig
Posizione	1023.445	327.01	3.13	.002	382.437	1664.454	***
bedrooms	-9.371	10.855	-0.86	.388	-30.649	11.907	
bathrooms	79.331	9.627	8.24	0	60.459	98.202	***
maxguests	18.735	3.841	4.88	0	11.207	26.264	***
responserate	.039	.063	0.62	.538	-.084	.161	
airbnbsuperhost	23.759	4.286	5.54	0	15.359	32.16	***
minimumstay	-.035	.069	-0.51	.612	-.169	.1	
ln_numberofreviews	-2.403	1.199	-2.00	.045	-4.754	-.052	**
ln_numberofphotos	6.202	3.419	1.81	.07	-.5	12.904	*
instantbookenabled	32.879	3.412	9.64	0	26.191	39.567	***
overallrating	24.478	6.862	3.57	0	11.026	37.929	***
Entirehomeapt	-27.308	33.272	-0.82	.412	-92.529	37.912	
Privateroom	-60.354	32.85	-1.84	.066	-124.746	4.039	*
Sharedroom	-64.173	39.925	-1.61	.108	-142.434	14.087	
Urbana	-32.096	4.252	-7.55	0	-40.432	-23.761	***
Rurale	-48.44	5.244	-9.24	0	-58.719	-38.161	***
numero_caratteri	.474	.169	2.81	.005	.143	.806	***
Constant	-311.883	72.623	-4.29	0	-454.24	-169.527	***
Mean dependent var		166.487	SD dependent var		202.522		
R-squared		0.389	Number of obs		9700		
F-test		94.379	Prob > F		0.000		
Akaike crit. (AIC)		125821.815	Bayesian crit. (BIC)		125951.053		

*** $p < .01$, ** $p < .05$, * $p < .1$

Fonte: File di Stata dell'elaborazione dei dati del campione del dataset pulito, dataset fornito dalla relatrice Elisabetta Raguseo (Raguseo, 2025).

In tabella 8 ([Raguseo, 2025](#)) si evince che il topic “Posizione” è correlato positivamente con la variabile dipendente, ed è statisticamente significativo all’1% in quanto p-value=0,002 e risulta essere, dunque, inferiore di 0,01.

Tabella 9 Regressione caratterizzata da un modello con variabili di controllo ed il topic “Comfort” con variabile dipendente ADR

ADR	Coef.	St.Err.	t-value	p-value	[95% Conf	Interval]	Sig
Comfort	1356.198	322.068	4.21	0	724.878	1987.518	***
bedrooms	-9.194	10.63	-0.86	.387	-30.031	11.643	
bathrooms	77.548	9.616	8.06	0	58.699	96.396	***
maxguests	17.861	3.713	4.81	0	10.583	25.138	***
responserate	.042	.062	0.67	.501	-.08	.163	
airbnbsuperhost	23.749	4.278	5.55	0	15.365	32.134	***
minimumstay	-.055	.073	-0.75	.451	-.198	.088	
ln_numberofreviews	-1.757	1.176	-1.49	.135	-4.061	.548	
ln_numberofphotos	2.639	3.27	0.81	.42	-3.771	9.05	
instantbookenabled	34.586	3.428	10.09	0	27.866	41.306	***
overallrating	22.348	6.786	3.29	.001	9.046	35.649	***
Entirehomeapt	-25.802	33.401	-0.77	.44	-91.274	39.67	
Privateroom	-60.189	33.046	-1.82	.069	-124.965	4.588	*
Sharedroom	-66.376	41.569	-1.60	.11	-147.86	15.109	
Urbana	-25.583	4.151	-6.16	0	-33.719	-17.446	***
Rurale	-52.242	5.347	-9.77	0	-62.724	-41.76	***
numero_caratteri	.496	.169	2.94	.003	.165	.826	***
Constant	-330.74	68.528	-4.83	0	-465.068	-196.411	***
Mean dependent var		166.487	SD dependent var		202.522		
R-squared		0.392	Number of obs		9700		
F-test		93.735	Prob > F		0.000		
Akaike crit. (AIC)		125772.921	Bayesian crit. (BIC)		125902.158		

*** $p < .01$, ** $p < .05$, * $p < .1$

Fonte: File di Stata dell’elaborazione dei dati del campione del dataset pulito, dataset fornito dalla relatrice Elisabetta Raguseo (Raguseo, 2025).

In tabella 9 (Raguseo, 2025) si evince che il topic “Comfort” è correlato positivamente con la variabile dipendente, ed è statisticamente significativo all’1%.

Tabella 10 Regressione caratterizzata da un modello con variabili di controllo ed il topic “Tipicità abitativa” con variabile dipendente ADR

ADR	Coef.	St.Err.	t-value	p-value	[95% Conf	Interval]	Sig
Tipicitàabitativa	-347.732	164.44	-2.11	.034	-670.069	-25.394	**

bedrooms	-9.524	10.844	-0.88	.38	-30.782	11.733	
bathrooms	79.423	9.621	8.26	0	60.565	98.282	***
maxguests	18.463	3.841	4.81	0	10.934	25.991	***
responserate	.033	.062	0.54	.59	-.088	.155	
airbnbsuperhost	23.774	4.289	5.54	0	15.367	32.181	***
minimumstay	-.039	.07	-0.56	.578	-.177	.099	
ln_numberofreviews	-2.143	1.184	-1.81	.07	-4.464	.179	*
ln_numberofphotos	6.084	3.441	1.77	.077	-.661	12.829	*
instantbookenabled	33.31	3.422	9.74	0	26.603	40.016	***
overallrating	23.325	6.81	3.43	.001	9.976	36.675	***
Entirehomeapt	-25.037	33.162	-0.76	.45	-90.041	39.966	
Privateroom	-59.408	32.806	-1.81	.07	-123.713	4.898	*
Sharedroom	-63.153	39.934	-1.58	.114	-141.432	15.127	
Urbana	-30.892	4.287	-7.21	0	-39.295	-22.488	***
Rurale	-49.783	5.317	-9.36	0	-60.205	-39.361	***
numero_caratteri	.477	.179	2.66	.008	.125	.829	***
Constant	-72.791	54.728	-1.33	.184	-180.07	34.487	

Mean dependent var	166.487	SD dependent var	202.522
R-squared	0.388	Number of obs	9700
F-test	93.877	Prob > F	0.000
Akaike crit. (AIC)	125827.685	Bayesian crit. (BIC)	125956.923

*** $p < .01$, ** $p < .05$, * $p < .1$

Fonte: File di Stata dell'elaborazione dei dati del campione del dataset pulito, dataset fornito dalla relatrice Elisabetta Raguseo (Raguseo, 2025).

In tabella 10 (Raguseo, 2025) si evince che il topic “Tipicità abitativa” è correlato negativamente con la variabile dipendente, ed è statisticamente significativo all’5%.

Tabella 11 Regressione caratterizzata da un modello con variabili di controllo ed il topic “Zone rurali” con variabile dipendente ADR

ADR	Coef.	St.Err.	t-value	p-value	[95% Conf	Interval]	Sig
Zonerurali	-1249.067	183.347	-6.81	0	-1608.465	-889.668	***
bedrooms	-9.342	10.899	-0.86	.391	-30.706	12.022	
bathrooms	78.861	9.654	8.17	0	59.937	97.784	***
maxguests	18.611	3.856	4.83	0	11.053	26.169	***
responserate	.029	.062	0.47	.64	-.093	.151	

airbnbsuperhost	23.771	4.276	5.56	0	15.39	32.152	***
minimumstay	-.044	.07	-0.62	.533	-.18	.093	
ln_numberofreviews	-2.306	1.184	-1.95	.051	-4.626	.014	*
ln_numberofphotos	7.284	3.447	2.11	.035	.527	14.041	**
instantbookenabled	31.814	3.408	9.34	0	25.134	38.495	***
overallrating	25.558	6.793	3.76	0	12.243	38.873	***
Entirehomeapt	-24.158	33.165	-0.73	.466	-89.169	40.853	
Privateroom	-61.218	32.781	-1.87	.062	-125.476	3.041	*
Sharedroom	-63.49	39.509	-1.61	.108	-140.935	13.956	
Urbana	-33.717	4.295	-7.85	0	-42.136	-25.298	***
Rurale	-46.134	5.252	-8.78	0	-56.43	-35.838	***
numero_caratteri	.47	.165	2.85	.004	.147	.793	***
Constant	64.54	54.751	1.18	.239	-42.782	171.863	

Mean dependent var	166.487	SD dependent var	202.522
R-squared	0.390	Number of obs	9700
F-test	93.513	Prob > F	0.000
Akaike crit. (AIC)	125793.806	Bayesian crit. (BIC)	125923.044

*** $p < .01$, ** $p < .05$, * $p < .1$

Fonte: File di Stata dell'elaborazione dei dati del campione del dataset pulito, dataset fornito dalla relatrice Elisabetta Raguseo (Raguseo, 2025).

In tabella 11 (Raguseo, 2025) si evince che il topic “Zone rurali” è correlato negativamente con la variabile dipendente, ed è statisticamente significativo all'1%.

Terminata l'analisi delle singole correlazioni tra i topic e la variabile dipendente mediante il modello con i controlli, è stato realizzato il modello contenente cinque dei sei topic analizzati (per evitare problemi di collinearità non sono stati inseriti tutti i topic individuati), e si è deciso di omettere il topic “Tipicità abitativa”.

Tabella 12 Regressione caratterizzata da un modello con variabili di controllo ed i topic con variabile dipendente ADR

ADR	Coef.	St.Err.	t-value	p-value	[95% Conf Interval]	Sig
Tipologiaecaratteristiche	416.076	251.649	1.65	.098	-77.21 909.361	*
Personalizzazione del titolo	194.09	167.837	1.16	.248	-134.905 523.086	
Posizione	1353.573	423.706	3.19	.001	523.022 2184.125	***

Comfort	1773.48	334.201	5.31	0	1118.377	2428.584	***
Zonerurali	-1059.5	258.149	-4.10	0	-1565.526	-553.474	***
bedrooms	-9.068	10.613	-0.85	.393	-29.871	11.735	
bathrooms	76.452	9.646	7.93	0	57.543	95.36	***
maxguests	18.247	3.699	4.93	0	10.996	25.499	***
responserate	.038	.063	0.61	.54	-.085	.161	
airbnbsuperhost	23.749	4.257	5.58	0	15.404	32.094	***
minimumstay	-.056	.073	-0.77	.441	-.2	.087	
ln_numberofreviews	-2.471	1.204	-2.05	.04	-4.832	-.11	**
ln_numberofphotos	3.852	3.31	1.16	.245	-2.635	10.339	
instantbookenabled	33.367	3.407	9.79	0	26.689	40.045	***
overallrating	24.382	6.776	3.60	0	11.099	37.665	***
Entirehomeapt	-23.096	34.125	-0.68	.499	-89.989	43.798	
Privateroom	-60.988	33.064	-1.84	.065	-125.799	3.824	*
Sharedroom	-68.361	41.586	-1.64	.1	-149.878	13.157	
Urbana	-29.004	4.161	-6.97	0	-37.159	-20.848	***
Rurale	-49.418	5.329	-9.27	0	-59.863	-38.973	***
numero_caratteri	.165	.182	0.90	.367	-.193	.522	
Constant	-557.137	167.555	-3.33	.001	-885.579	-228.694	***

Mean dependent var	166.487	SD dependent var	202.522
R-squared	0.396	Number of obs	9700
F-test	77.565	Prob > F	0.000
Akaike crit. (AIC)	125717.606	Bayesian crit. (BIC)	125875.563

*** $p < .01$, ** $p < .05$, * $p < .1$

Fonte: File di Stata dell'elaborazione dei dati del campione del dataset pulito, dataset fornito dalla relatrice Elisabetta Raguseo (Raguseo, 2025).

In tabella 12 ([Raguseo, 2025](#)) si evince che dei cinque topic presenti nel modello, quattro risultano statisticamente significativi e sono:

- “Tipologia e caratteristiche dell'alloggio”, che risulta significativo al 10% (in quanto possiede un $p\text{-value}=0,098$) e correlato positivamente con la variabile dipendente;
- “Posizione”, che risulta significativo all'1% (in quanto possiede un $p\text{-value}=0,001$) e correlato positivamente con la variabile dipendente;

- “Comfort”, che risulta significativo all’1% (in quanto possiede un p-value=0,000) e correlato positivamente con la variabile dipendente;
- “Zone rurali”, che risulta significativo all’1% (in quanto possiede un p-value=0,000) e correlato negativamente con la variabile dipendente.

In merito alle variabili di controllo rappresentative delle caratteristiche dell’alloggio offerto, ovvero “Entirehomeapt”, “Privateroom”, “Sharedroom”, “Bedrooms”, “Bathrooms”, “Urbana” e “Rurale”, la variabile “Privateroom” è statisticamente significativa al 10%, mentre, solo le variabili “Urbana”, “Rurale” e “Bathrooms” sono risultate statisticamente significative all’1% nel modello in tabella 12 ([Raguseo, 2025](#)).

Relativamente alle variabili di controllo rappresentative delle caratteristiche dell’host, ovvero “Airbnbsuperhost” e “Responce rate”, solo la variabile “Airbnbsuperhost” risulta essere statisticamente significativa all’1% nel modello in tabella 12 ([Raguseo, 2025](#)).

Delle variabili di controllo riferite alle caratteristiche dell’annuncio, ovvero “Maxguests”, “Minimumstay”, “ln_numberofphotos”, “Overallrating”, “Instantbookenabled”, “numero_caratteri” e “ln_numberofreviews”, solo “ln_numberofreviews” risulta essere statisticamente significativo al 5% mentre “Maxguests”, “Overallrating” e “Instantbookenabled” sono statisticamente significativi all’1% nel modello in tabella 12 ([Raguseo, 2025](#)).

Tabella 13 Tabella riassuntiva dei modelli con i topic ed i controlli con variabile dipendente ADR

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	ADR	ADR	ADR	ADR	ADR	ADR	ADR	ADR
Controlli	Si	Si	Si	Si	Si	Si	Si	Si
Tipologia e caratteristiche dell'alloggio		74.04						416.1 [*]
		(217.4)						(251.6)
Personalizzazione del titolo			-218.7 [*]					194.1
			(123.8)					(167.8)
Posizione				1023.4 ^{***}				1353.6 ^{***}
				(327.0)				(423.7)
Comfort					1356.2 ^{***}			1773.5 ^{***}
					(322.1)			(334.2)

Tipicità abitativa						-347.7 ^{**}		
							(164.4)	
Zone rurali							-1249.1 ^{***}	-1059.5 ^{***}
							(183.3)	(258.1)
_cons	-136.7 ^{***}	-149.8 ^{**}	-101.0 ^{**}	-311.9 ^{***}	-330.7 ^{***}	-72.79	64.54	-557.1 ^{***}
	(47.94)	(65.44)	(51.19)	(72.62)	(68.53)	(54.73)	(54.75)	(167.6)
N	9700	9700	9700	9700	9700	9700	9700	9700

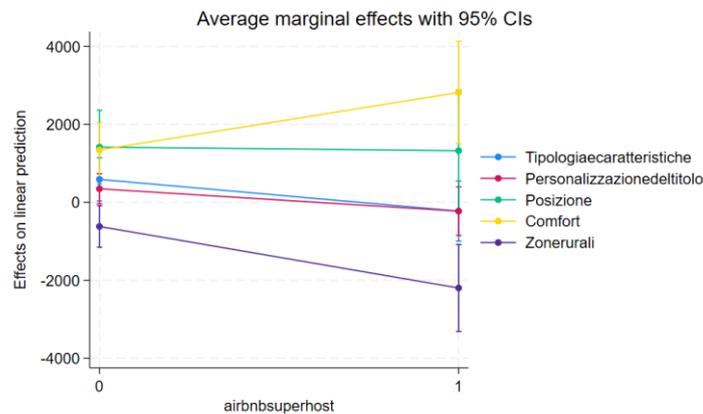
Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Fonte: File di Stata dell'elaborazione dei dati del campione del dataset pulito, dataset fornito dalla relatrice Elisabetta Raguseo (Raguseo, 2025).

Dopo aver analizzato le correlazioni tra la variabile dipendente ed i topic, sono state analizzate le moderazioni per ogni topic, delle variabili: “Airbnb Superhost”, “Overall rating”, “Urbana” e “Rurale”.

Figura 41 Grafico degli effetti marginali derivanti dalla variabile “Airbnb Superhost” per il modello con variabile dipendente ADR



Fonte: Screen di Stata dell'elaborazione dei dati del campione del dataset pulito, dataset fornito dalla relatrice Elisabetta Raguseo (Raguseo, 2025).

Tabella 14 Effetti marginali derivanti dalla variabile “Airbnb Superhost” per il modello con variabile dipendente ADR

Average marginal effects
Model VCE: Robust

Number of obs = 9,700

Expression: Linear prediction, predict()

dy/dx wrt: Tipologiaeacaratteristiche Personalizzazione del titolo Posizione Comfort Zonerurali

		Delta-method				
		dy/dx	std.err.	t	P>t	[95% conf. interval]
Tipologiaeacaratteristiche						
airnbnsuperhost						
0		588.308	283.554	2.070	0.038	32.483 1144.132
1		-224.166	393.399	-0.570	0.569	-995.310 546.979
Personalizzazione del titolo						
airnbnsuperhost						
0		346.854	196.772	1.760	0.078	-38.861 732.569
1		-226.587	317.776	-0.710	0.476	-849.495 396.321
Posizione						
airnbnsuperhost						
0		1416.944	482.943	2.930	0.003	470.275 2363.614
1		1324.324	794.146	1.670	0.095	-232.368 2881.015
Comfort						
airnbnsuperhost						
0		1342.032	357.297	3.760	0.000	641.655 2042.409
1		2818.030	670.214	4.200	0.000	1504.271 4131.789
Zonerurali						
airnbnsuperhost						
0		-620.441	272.453	-2.280	0.023	-1154.506 -86.375
1		-2199.391	568.802	-3.870	0.000	-3314.361 -1084.421

Fonte: File di Stata dell'elaborazione dei dati del campione del dataset pulito, dataset fornito dalla relatrice Elisabetta Raguseo (Raguseo, 2025).

Dall'analisi degli effetti marginali relativi alla moderazione della variabile "Airnbnsuperhost", si evince che la moderazione è statisticamente significativa solo per:

- Il topic "Posizione", dove il topic viene moderato positivamente e in modo statisticamente significativo al 10%;
- Il topic "Comfort", dove il topic viene moderato positivamente e in modo statisticamente significativo all'1%;

- Il topic “Zone rurali”, dove il topic viene moderato negativamente e in modo statisticamente significativo all’1%.

Tabella 15 Regressione con moderazione “Airbnbsuperhost” per il modello con variabile dipendente ADR

ADR	Coef.	St.Err.	t-value	p-value	[95% Conf	Interval]	Sig
Tipologiaeacaratteristiche	588.308	283.554	2.07	.038	32.483	1144.132	**
Personalizzazione del titolo	346.854	196.772	1.76	.078	-38.861	732.569	*
Posizione	1416.944	482.943	2.93	.003	470.275	2363.614	***
Comfort	1342.032	357.297	3.76	0	641.656	2042.409	***
Zonerurali	-620.441	272.453	-2.28	.023	-1154.506	-86.375	**
bedrooms	-8.788	10.654	-0.82	.409	-29.672	12.096	
bathrooms	76.059	9.698	7.84	0	57.05	95.068	***
maxguests	18.15	3.697	4.91	0	10.904	25.397	***
responserate	.037	.062	0.60	.551	-.085	.16	
airbnbsuperhost	297.929	316.158	0.94	.346	-321.806	917.665	
minimumstay	-.053	.071	-0.74	.459	-.192	.087	
ln_numeroofreviews	-2.322	1.197	-1.94	.052	-4.669	.025	*
ln_numerooffotos	3.512	3.303	1.06	.288	-2.963	9.987	
instantbookenabled	33.325	3.414	9.76	0	26.633	40.016	***
overallrating	24.866	6.77	3.67	0	11.596	38.136	***
Entirehomeapt	-21.459	34.173	-0.63	.53	-88.444	45.527	
Privateroom	-58.634	33.139	-1.77	.077	-123.595	6.326	*
Sharedroom	-66.324	40.922	-1.62	.105	-146.539	13.891	
Urbana	-29.665	4.149	-7.15	0	-37.797	-21.532	***
Rurale	-50.158	5.292	-9.48	0	-60.531	-39.785	***
numero_caratteri	.179	.183	0.98	.329	-.18	.537	
airbnbsuperhost*tipologiaeacaratteristiche	-812.474	435.822	-1.86	.062	-1666.777	41.83	*
airbnbsuperhost*personalizzazione del titolo	-573.441	374.306	-1.53	.126	-1307.159	160.277	
airbnbsuperhost*posizione	-92.621	905.764	-0.10	.919	-1868.108	1682.866	
airbnbsuperhost*comfort	1475.998	723.029	2.04	.041	58.71	2893.287	**
airbnbsuperhost*zone rurali	-1578.95	618.705	-2.55	.011	-2791.742	-366.158	**
Constant	-629.105	191.154	-3.29	.001	-1003.808	-254.403	***
Mean dependent var		166.487	SD dependent var			202.522	
R-squared		0.398	Number of obs			9700	
F-test		65.224	Prob > F			0.000	

*** $p < .01$, ** $p < .05$, * $p < .1$

Fonte: File di Stata dell'elaborazione dei dati del campione del dataset pulito, dataset fornito dalla relatrice Elisabetta Raguseo (Raguseo, 2025).

Analizzando la moderazione della variabile “Airbnbsuperhost” nel modello in cui sono presenti anche i controlli, si verifica che la moderazione è statisticamente significativa al 10% per il topic “Tipologia e caratteristiche dell'alloggio” e tale moderazione è negativa; dunque, se l'host possiede il badge superhost, tale topic ha un effetto minore sui ricavi per giorni prenotati. Il topic “Comfort” è moderato in modo statisticamente significativo al 5% e positivamente dalla variabile “Airbnbsuperhost”, dunque, il possesso del badge Superhost determina un incremento degli effetti positivi del topic relativo ai comfort sulla variabile dipendente. La variabile “Airbnbsuperhost”, infine, modera negativamente ed in modo statisticamente significativo al 5% il topic “Zone rurali”, andandone a ridurre, di conseguenza, gli effetti negativi sulla variabile dipendente. Si evince, dunque, che gli effetti della moderazione variano a seconda che si tenga conto solo della singola relazione come in tabella 14 ([Raguseo, 2025](#)) o che vengano inserite anche le variabili di controllo nel modello, come in tabella 15 ([Raguseo, 2025](#)).

Oltre all'analisi della moderazione della variabile “Airbnbsuperhost”, si è deciso di analizzare anche la moderazione relativa alla variabile “Overallrating”; a differenza della variabile “Airbnbsuperhost”, la variabile “Overallrating” non è una variabile binaria, bensì una variabile continua. Per poter analizzare, dunque, l'effetto moderatore del rating, si definisce una nuova variabile: “mediana_rating”, la quale assume un valore pari a “1” se il rating associato all'annuncio è maggiore o uguale del valore della mediana relativa alla variabile “Overallrating”, “0” altrimenti.

Tabella 16 Analisi delle caratteristiche della variabile “Overallrating”

Statistiche	Valori				
Osservazioni	9,741	Kurtosis	21.66932	Percentile 75%	4.95
Media	4.687597	Percentile 1%	3	Percentile 90%	5
Deviazione standard	0.3978403	Percentile 5%	4	Percentile 95%	5
Varianza	0.1582769	Percentile 10%	4.2	Percentile 99%	5
Skewness	-3.302597	Percentile 25%	4.55	Valore minimo	1
		Mediana (50%)	4.8	Valore massimo	5

Fonte: File di Stata dell’elaborazione dei dati del campione del dataset pulito, dataset fornito dalla relatrice Elisabetta Raguseo (Raguseo, 2025). La conversione da screen di Stata a file modificabile è avvenuta tramite ChatGPT (OpenAI, 2025).

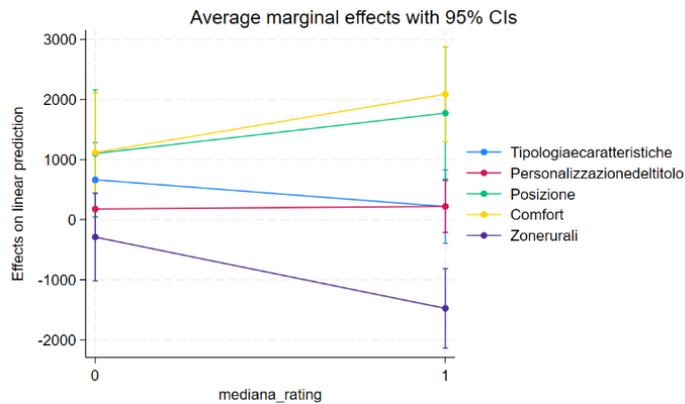
Il valore della mediana della variabile “Overallrating” è stato ottenuto mediante l’analisi, sul software Stata, delle caratteristiche della variabile stessa; dalla tabella 16 (Raguseo, 2025) emerge che il valore della mediana associata al rating è pari a 4,8.

Tabella 17 Caratteristiche della variabile “mediana_rating”

Variable	Obs	Mean	Std. Dev.	Min	Max
Mediana_rating	9741	.562	.496	0	1

Fonte: File di Stata dell’elaborazione dei dati del campione del dataset pulito, dataset fornito dalla relatrice Elisabetta Raguseo (Raguseo, 2025).

Figura 49 Grafico degli effetti marginali derivanti dalla variabile “Mediana_rating” per il modello con variabile dipendente ADR



Fonte: Screen di Stata dell'elaborazione dei dati del campione del dataset pulito, dataset fornito dalla relatrice Elisabetta Raguseo (Raguseo, 2025).

Tabella 18 Effetti marginali derivanti dalla variabile “Mediana_rating” per il modello con variabile dipendente ADR

Average marginal effects
 Model VCE: Robust
 Expression: Linear prediction, predict()
 dy/dx wrt: Tipologiaecaratteristiche Personalizzazione del titolo Posizione Comfort Zonerurali

Number of obs = 9,700

	Delta-method				
	dy/dx	std. err.	t	P>t	[95% conf. interval]
Tipologiaecaratteristiche					
mediana rating					
0	664.635	315.552	2.110	0.035	46.087 1283.182
1	217.290	312.409	0.700	0.487	-395.098 829.678
Personalizzazione del titolo					
mediana rating					
0	177.574	226.863	0.780	0.434	-267.125 622.272
1	220.259	221.087	1.000	0.319	-213.118 653.635
Posizione					
mediana rating					
0	1101.098	541.000	2.040	0.042	40.624 2161.571
1	1772.719	561.592	3.160	0.002	671.881 2873.556
Comfort					
mediana rating					

0	1116.739	507.898	2.200	0.028	121.152	2112.326
1	2088.779	403.425	5.180	0.000	1297.982	2879.577
Zonerurali						
mediana rating						
0	-290.680	372.091	-0.780	0.435	-1020.057	438.698
1	-1476.727	337.211	-4.380	0.000	-2137.730	-815.723

Fonte: File di Stata dell'elaborazione dei dati del campione del dataset pulito, dataset fornito dalla relatrice Elisabetta Raguseo (Raguseo, 2025).

Dall'analisi degli effetti marginali della variabile "Mediana_rating", si può notare che la moderazione è statisticamente significativa solo per:

- Il topic "Posizione", dove il topic viene moderato positivamente e in modo statisticamente significativo all'1%;
- Il topic "Comfort", dove il topic viene moderato positivamente e in modo statisticamente significativo all'1%;
- Il Topic "Zone rurali", dove il topic viene moderato negativamente e in modo statisticamente significativo all'1%.

Tabella 19 Regressione con moderazione "Mediana_rating" per il modello con variabile dipendente ADR

ADR	Coef.	St.Err.	t-value	p-value	[95% Conf	Interval]	Sig
Tipologiae caratteristiche	664.635	315.552	2.11	.035	46.087	1283.182	**
Personalizzazione del titolo	177.574	226.863	0.78	.434	-267.125	622.272	
Posizione	1101.098	541	2.04	.042	40.624	2161.571	**
Comfort	1116.739	507.898	2.20	.028	121.152	2112.326	**
Zonerurali	-290.68	372.092	-0.78	.435	-1020.057	438.698	
bedrooms	-8.678	10.536	-0.82	.41	-29.33	11.975	
bathrooms	75.593	9.647	7.84	0	56.683	94.502	***
maxguests	18.44	3.673	5.02	0	11.241	25.639	***
response rate	.044	.063	0.70	.481	-.079	.167	
airbnb superhost	19.54	4.274	4.57	0	11.161	27.918	***
minimum stay	-.051	.076	-0.67	.504	-.201	.098	
ln_number_of_reviews	-1.203	1.217	-0.99	.323	-3.589	1.183	
ln_number_of_photos	2.741	3.307	0.83	.407	-3.742	9.224	
instant_bookable	33.314	3.41	9.77	0	26.629	39.998	***
overall rating	5.601	8.494	0.66	.51	-11.05	22.251	

Entirehomeapt	-21.24	34.153	-0.62	.534	-88.186	45.706	
Privateroom	-59.502	33.148	-1.80	.073	-124.478	5.475	*
Sharedroom	-64.038	41.836	-1.53	.126	-146.045	17.97	
Urbana	-28.261	4.139	-6.83	0	-36.374	-20.149	***
Rurale	-49.611	5.312	-9.34	0	-60.024	-39.199	***
numero_caratteri	.153	.183	0.84	.403	-.205	.51	
mediana_rating*tipologiaecaratteristiche	-447.345	386.235	-1.16	.247	-1204.447	309.758	
mediana_rating*personalizzazione del titolo	42.685	300.97	0.14	.887	-547.279	632.648	
mediana_rating*posizione	671.621	713.882	0.94	.347	-727.737	2070.979	
mediana_rating*comfort	972.04	616.031	1.58	.115	-235.509	2179.589	
mediana_rating*zone rurali	-1186.047	496.931	-2.39	.017	-2160.136	-211.958	**
mediana_rating	17.734	263.553	0.07	.946	-498.885	534.352	
Constant	-502.257	225.14	-2.23	.026	-943.578	-60.936	**
Mean dependent var		166.487	SD dependent var			202.522	
R-squared		0.398	Number of obs			9700	
F-test		61.503	Prob > F			0.000	
Akaike crit. (AIC)		125684.657	Bayesian crit. (BIC)			125885.694	

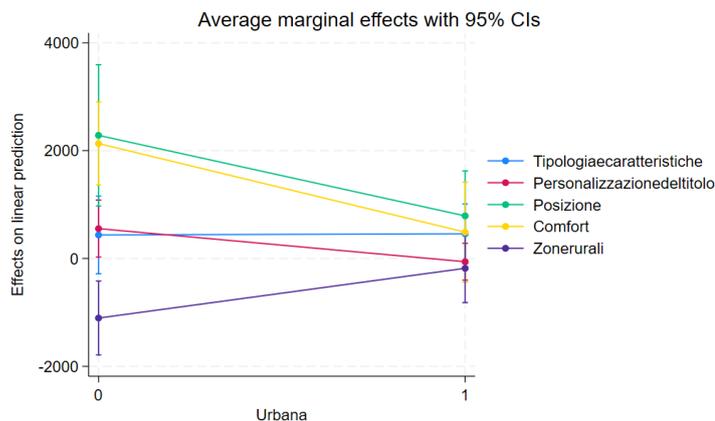
*** $p < .01$, ** $p < .05$, * $p < .1$

Fonte: File di Stata dell'elaborazione dei dati del campione del dataset pulito, dataset fornito dalla relatrice Elisabetta Raguseo (Raguseo, 2025).

Analizzando la moderazione della variabile “Mediana_rating” nel modello in cui sono presenti anche i controlli, si verifica che la moderazione è statisticamente significativa al 5% solo per il topic “Zone rurali” e tale moderazione è negativa; dunque, se l’host presenta un rating maggiore o uguale a 4,8 stelle, l’effetto negativo della presenza del topic “Zone rurali” sui ricavi per giorni prenotati, diminuisce. Anche in questo caso, come nel caso dell’analisi della moderazione della variabile “Airbnbsuperhost”, si evince che gli effetti della moderazione variano a seconda che si consideri solo la singola relazione, come in tabella 18 (Raguseo, 2025), o che vengano inserite anche le variabili di controllo nel modello, come in tabella 19 (Raguseo, 2025).

Sono state analizzate, infine, le moderazioni circa la localizzazione dell’alloggio.

Figura 50 Grafico degli effetti marginali derivanti dalla variabile "Urbana" per il modello con variabile dipendente ADR



Fonte: Screen di Stata dell'elaborazione dei dati del campione del dataset pulito, dataset fornito dalla relatrice Elisabetta Raguseo (Raguseo, 2025).

Tabella 20 Effetti marginali derivanti dalla variabile "Urbana" per il modello con variabile dipendente ADR

Average marginal effects

Number of obs = 9,700

Model VCE: Robust

Expression: Linear prediction, predict()

dy/dx wrt: Tipologiaeacaratteristiche Personalizzazioneedeltitolo Posizione Comfort Zonerurali

	Delta-method				
	dy/dx	std. err.	t	P>t	[95% conf. interval]
Tipologiaeacaratteristiche					
Urbana					
0	436.951	367.275	1.190	0.234	-282.985 1156.887
1	457.988	282.503	1.620	0.105	-95.777 1011.754
Personalizzazioneedeltitolo					
Urbana					
0	554.376	269.492	2.060	0.040	26.117 1082.636
1	-59.591	174.189	-0.340	0.732	-401.039 281.856
Posizione					
Urbana					
0	2283.768	668.544	3.420	0.001	973.283 3594.254
1	790.592	424.982	1.860	0.063	-42.461 1623.645
Comfort					
Urbana					

0	2132.137	391.934	5.440	0.000	1363.865	2900.409
1	489.133	472.589	1.040	0.301	-437.240	1415.507
Zonerurali						
Urbana						
0	-1103.942	349.261	-3.160	0.002	-1788.567	-419.318
1	-181.336	324.187	-0.560	0.576	-816.809	454.138

Fonte: File di Stata dell'elaborazione dei dati del campione del dataset pulito, dataset fornito dalla relatrice Elisabetta Raguseo (Raguseo, 2025).

La variabile “Urbana” è una dummy che assume un valore pari a “1” se l’annuncio fa riferimento ad un alloggio situato in una zona urbana, “0” altrimenti. Dall’analisi degli effetti marginali della variabile “Urbana”, si può notare che la moderazione è statisticamente significativa al 10% solo per il topic “Posizione”, il quale viene moderato positivamente.

Tabella 21 Regressione con moderazione “Urbana” per il modello con variabile dipendente ADR

ADR	Coef.	St.Err.	t-value	p-value	[95% Conf	Interval]	Sig
Tipologiaeaccharatteristiche	436.951	367.275	1.19	.234	-282.985	1156.887	
Personalizzazione dellalloggio	554.376	269.492	2.06	.04	26.117	1082.636	**
Posizione	2283.768	668.544	3.42	.001	973.283	3594.254	***
Comfort	2132.137	391.934	5.44	0	1363.865	2900.409	***
Zonerurali	-1103.942	349.261	-3.16	.002	-1788.567	-419.318	***
bedrooms	-9.204	10.635	-0.87	.387	-30.05	11.642	
bathrooms	76.674	9.671	7.93	0	57.717	95.631	***
maxguests	18.303	3.706	4.94	0	11.039	25.567	***
responserate	.034	.063	0.54	.587	-.089	.157	
airbnbsuperhost	23.549	4.268	5.52	0	15.183	31.916	***
minimumstay	-.056	.072	-0.77	.442	-.197	.086	
ln_numberofreviews	-2.363	1.194	-1.98	.048	-4.704	-.022	**
ln_numberofphotos	3.386	3.305	1.02	.306	-3.092	9.864	
instantbookenable	33.401	3.433	9.73	0	26.671	40.132	***
overallrating	24.005	6.761	3.55	0	10.752	37.258	***
Entirehomeapt	-21.484	34.338	-0.63	.532	-88.794	45.826	

Privateroom	-58.761	33.28	-1.77	.077	-123.996	6.475	*
Sharedroom	-70.971	41.882	-1.69	.09	-153.068	11.126	*
Urbana	434.639	272.211	1.60	.11	-98.951	968.229	
Rurale	-48.682	5.381	-9.05	0	-59.23	-38.133	***
numero_caratteri	.155	.182	0.85	.396	-.203	.512	
Urbana*tipologiae caratteristiche	21.037	417.02	0.05	.96	-796.41	838.484	
Urbana*personalizzazione del titolo	-613.968	307.397	-2.00	.046	-1216.53	-11.405	**
Urbana*posizione	-1493.176	735.004	-2.03	.042	-2933.937	-52.415	**
Urbana*comfort	-1643.003	579.374	-2.84	.005	-2778.698	-507.309	***
Urbana*zonerurali	922.606	468.638	1.97	.049	3.978	1841.235	**
Constant	-829.663	262.828	-3.16	.002	-1344.862	-314.465	***

Mean dependent var	166.487	SD dependent var	202.522
R-squared	0.397	Number of obs	9700
F-test	66.096	Prob > F	0.000
Akaike crit. (AIC)	125708.572	Bayesian crit. (BIC)	125902.429

*** $p < .01$, ** $p < .05$, * $p < .1$

Fonte: File di Stata dell'elaborazione dei dati del campione del dataset pulito, dataset fornito dalla relatrice Elisabetta Raguseo (Raguseo, 2025).

Dall'analisi del modello nel suo complesso, in cui sono compresi anche i controlli oltre che la moderazione della variabile “Urbana”, si evince che:

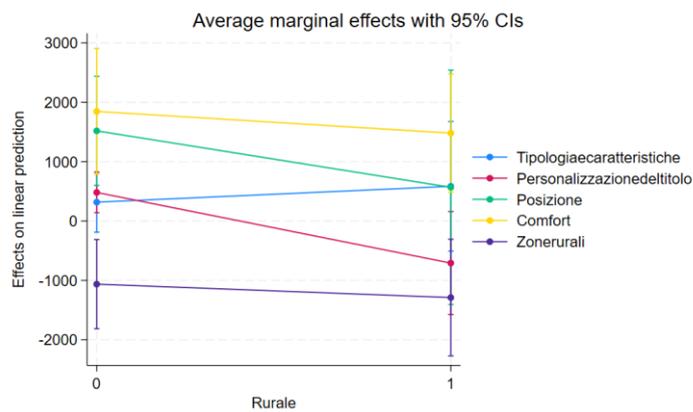
- Il topic “Personalizzazione del titolo” è moderato negativamente dall'essere in una zona urbana dell'alloggio (moderazione statisticamente significativa al 5%), dunque, l'effetto positivo del topic sulla variabile dipendente, diminuisce quando l'alloggio è situato in una zona urbana;
- Il topic “Posizione” è moderato negativamente dall'essere in una zona urbana dell'alloggio (moderazione statisticamente significativa al 5%), dunque, l'effetto positivo del topic sulla variabile dipendente, diminuisce quando l'alloggio è situato in una zona urbana;
- Il topic “Comfort” è moderato negativamente dall'essere in una zona urbana dell'alloggio (moderazione statisticamente significativa all'1%), dunque, l'effetto

positivo del topic sulla variabile dipendente, diminuisce quando l'alloggio è situato in una zona urbana;

- Il topic "Zone rurali" è moderato positivamente dall'essere in una zona urbana dell'alloggio (moderazione statisticamente significativa al 5%), dunque, l'effetto negativo del topic, aumenta se l'alloggio è situato in una zona urbana.

L'analisi è proseguita con lo studio della moderazione della variabile "Rurale". La variabile "Rurale" è una variabile dummy che assume un valore pari a "1" se l'annuncio fa riferimento ad un alloggio situato in una zona rurale, "0" altrimenti.

Figura 51 Grafico degli effetti marginali derivanti dalla variabile "Rurale" per il modello con variabile dipendente ADR



Fonte: Screen di Stata dell'elaborazione dei dati del campione del dataset pulito, dataset fornito dalla relatrice Elisabetta Raguseo (Raguseo, 2025).

Tabella 22 Effetti marginali derivanti dalla variabile "Rurale" per il modello con variabile dipendente ADR

Average marginal effects
 Model VCE: Robust
 Expression: Linear prediction, predict()
 dy/dx wrt: Tipologiaeacaratteristiche Personalizzazione del titolo Posizione Comfort Zonerurali
 Number of obs = 9,700

	Delta-method				
	dy/dx	std. err.	t	P>t	[95% conf. interval]
Tipologiaeacaratteristiche					
Rurale					
0	317.628	258.002	1.230	0.218	-188.110 823.366
1	584.999	557.126	1.050	0.294	-507.086 1677.083
Personalizzazione de					

Ititolo						
Rurale						
0	481.795	173.737	2.770	0.006	141.234	822.356
1	-707.821	442.256	-1.600	0.110	-1574.736	159.094
Posizione						
Rurale						
0	1518.823	468.987	3.240	0.001	599.510	2438.136
1	567.882	1006.557	0.560	0.573	-1405.181	2540.945
Comfort						
Rurale						
0	1845.694	539.915	3.420	0.001	787.347	2904.041
1	1480.700	507.218	2.920	0.004	486.446	2474.953
Zonerurali						
Rurale						
0	-1062.794	383.146	-2.770	0.006	-1813.839	-311.748
1	-1288.551	500.810	-2.570	0.010	-2270.243	-306.859

Fonte: File di Stata dell'elaborazione dei dati del campione del dataset pulito, dataset fornito dalla relatrice
Elisabetta Raguseo (Raguseo, 2025).

Dall'analisi degli effetti marginali, si può notare che la moderazione della variabile "Rurale" è statisticamente significativa solo per:

- Il topic "Comfort", dove il topic viene moderato positivamente ed in modo statisticamente significativo all'1%.
- Il topic "Zone rurali", dove il topic viene moderato negativamente ed in modo statisticamente significativo al 5%.

Tabella 23 Regressione con moderazione "Rurale" per il modello con variabile dipendente ADR

ADR	Coef.	St.Err.	t-value	p-value	[95% Conf	Interval]	Sig
Tipologiaeacatteristiche	317.628	258.002	1.23	.218	-188.11	823.366	
Personalizzazione del titolo	481.795	173.737	2.77	.006	141.234	822.356	***
Posizione	1518.823	468.987	3.24	.001	599.51	2438.136	***
Comfort	1845.694	539.915	3.42	.001	787.347	2904.041	***
Zonerurali	-1062.794	383.146	-2.77	.006	-1813.839	-311.748	***
bedrooms	-8.925	10.584	-0.84	.399	-29.673	11.823	
bathrooms	76.115	9.644	7.89	0	57.212	95.019	***

maxguests	18.181	3.673	4.95	0	10.982	25.381	***
responserate	.041	.063	0.65	.513	-.083	.166	
airbnbsuperhost	23.736	4.231	5.61	0	15.442	32.029	***
minimumstay	-.056	.075	-0.74	.459	-.203	.092	
ln_numberofreviews	-2.358	1.202	-1.96	.05	-4.714	-.003	**
ln_numberofphotos	3.883	3.307	1.17	.24	-2.6	10.365	
instantbookenable	32.933	3.428	9.61	0	26.212	39.653	***
overallrating	23.896	6.737	3.55	0	10.689	37.102	***
Entirehomeapt	-22.405	34.279	-0.65	.513	-89.599	44.79	
Privateroom	-60.096	33.229	-1.81	.071	-125.231	5.04	*
Sharedroom	-66.467	41.805	-1.59	.112	-148.412	15.479	
Urbana	-29.591	4.119	-7.18	0	-37.666	-21.516	***
Rurale	364.071	434.942	0.84	.403	-488.507	1216.65	
numero_caratteri	.199	.183	1.09	.276	-.159	.558	
Rurale*tipologiae caratteristiche	267.371	570.764	0.47	.639	-851.447	1386.188	
Rurale*personalizzazione del titolo	-1189.616	467.835	-2.54	.011	-2106.67	-272.562	**
Rurale*posizione	-950.941	1067.041	-0.89	.373	-3042.564	1140.682	
Rurale*comfort	-364.994	716.81	-0.51	.611	-1770.091	1040.103	
Rurale*zonerurali	-225.758	630.331	-0.36	.72	-1461.339	1009.824	
Constant	-627.823	154.189	-4.07	0	-930.066	-325.581	***

Mean dependent var	166.487	SD dependent var	202.522
R-squared	0.396	Number of obs	9700
F-test	68.246	Prob > F	0.000
Akaike crit. (AIC)	125715.657	Bayesian crit. (BIC)	125909.514

*** $p < .01$, ** $p < .05$, * $p < .1$

Fonte: File di Stata dell'elaborazione dei dati del campione del dataset pulito, dataset fornito dalla relatrice Elisabetta Raguseo (Raguseo, 2025).

Analizzando il modello nel suo complesso, in cui sono compresi anche i controlli, si può notare che solo il topic “Personalizzazione del titolo” risulta essere moderato negativamente ed in modo statisticamente significativo al 5% dall’essere in una zona rurale dell’alloggio; dunque, se l’annuncio fa riferimento ad un alloggio situato in una

zona rurale, il topic “Personalizzazione del titolo”, ha un effetto positivo minore sui ricavi per giorni prenotati.

Tabella 24 Tabella riassuntiva dei modelli con controlli e moderazioni con variabile dipendente ADR

	(1)	(2)	(3)	(4)
	ADR	ADR	ADR	ADR
Tipologia e caratteristiche dell'alloggio	588.3**	664.6**	317.6	437.0
	(283.6)	(315.6)	(258.0)	(367.3)
Personalizzazione del titolo	346.9*	177.6	481.8***	554.4**
	(196.8)	(226.9)	(173.7)	(269.5)
Posizione	1416.9***	1101.1**	1518.8***	2283.8***
	(482.9)	(541.0)	(469.0)	(668.5)
Comfort	1342.0***	1116.7**	1845.7***	2132.1***
	(357.3)	(507.9)	(539.9)	(391.9)
Zone rurali	-620.4**	-290.7	-1062.8***	-1103.9***
	(272.5)	(372.1)	(383.1)	(349.3)
1.airbnbsuperhost#c.Tipologiaeacaratteristiche	-812.5*			
	(435.8)			
1.airbnbsuperhost#c.Personalizzazionedeltitolo	-573.4			
	(374.3)			
1.airbnbsuperhost#c.Posizione	-92.62			
	(905.8)			
1.airbnbsuperhost#c.Comfort	1476.0**			
	(723.0)			
1.airbnbsuperhost#c.Zonerurali	-1579.0**			
	(618.7)			
1.mediana_rating#c.Tipologiaeacaratteristiche		-447.3		
		(386.2)		
1.mediana_rating#c.Personalizzazionedeltitolo		42.68		
		(301.0)		
1.mediana_rating#c.Posizione		671.6		
		(713.9)		
1.mediana_rating#c.Comfort		972.0		
		(616.0)		
1.mediana_rating#c.Zonerurali		-1186.0**		
		(496.9)		
1.Rurale#c.Tipologiaeacaratteristiche			267.4	

			(570.8)	
1.Rurale#c.Personalizzazionedeltitolo			-1189.6**	
			(467.8)	
1.Rurale#c.Posizione			-950.9	
			(1067.0)	
1.Rurale#c.Comfort			-365.0	
			(716.8)	
1.Rurale#c.Zonerurali			-225.8	
			(630.3)	
1.Urbana#c.Tipologiaeacaratteristiche				21.04
				(417.0)
1.Urbana#c.Personalizzazionedeltitolo				-614.0**
				(307.4)
1.Urbana#c.Posizione				-1493.2**
				(735.0)
1.Urbana#c.Comfort				-1643.0***
				(579.4)
1.Urbana#c.Zonerurali				922.6**
				(468.6)
Controlli	Si	Si	Si	Si
cons	-629.1***	-502.3**	-627.8***	-829.7***
	(191.2)	(225.1)	(154.2)	(262.8)
<i>N</i>	9700	9700	9700	9700

Standard errors in parentheses * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Fonte: File di Stata dell'elaborazione dei dati del campione del dataset pulito, dataset fornito dalla relatrice Elisabetta Raguseo (Raguseo, 2025).

Dopo aver effettuato l'analisi della correlazione tra i topic ed i ricavi medi per giorni prenotati, è stata analizzata la correlazione tra i topic ed i ricavi medi per giorni disponibili totali (ovvero, sia quelli prenotati sia i restanti disponibili), cioè si prende in considerazione come variabile dipendente “revPAN”.

Tabella 25 Regressione caratterizzata da un modello con solo variabili di controllo con variabile dipendente revPAN

revPAN	Coef.	St.Err.	t-value	p-value	[95% Conf	Interval]	Sig
bedrooms	-7.682	4.834	-1.59	.112	-17.157	1.793	

bathrooms	38.76	6.49	5.97	0	26.037	51.482	***
maxguests	9.186	1.695	5.42	0	5.864	12.508	***
responserate	.179	.035	5.18	0	.111	.247	***
airbnbsuperhost	15.136	2.963	5.11	0	9.328	20.944	***
minimumstay	-.021	.023	-0.89	.372	-.066	.025	
ln_numberofreviews	6.674	.78	8.56	0	5.145	8.203	***
ln_numberofphotos	2.005	2.228	0.90	.368	-2.361	6.371	
instantbookenabled	29.789	2.468	12.07	0	24.951	34.627	***
overallrating	18.656	3.917	4.76	0	10.979	26.333	***
Entirehomeapt	-29.434	15.019	-1.96	.05	-58.874	.006	*
Privateroom	-40.294	14.962	-2.69	.007	-69.623	-10.966	***
Sharedroom	-37.294	18.716	-1.99	.046	-73.98	-.607	**
Urbana	-2.498	2.813	-0.89	.375	-8.012	3.016	
Rurale	-26.561	3.636	-7.31	0	-33.688	-19.434	***
numero_caratteri	.175	.125	1.41	.16	-.069	.419	
Constant	-134.715	23.916	-5.63	0	-181.596	-87.835	***

Mean dependent var	67.197	SD dependent var	122.619
R-squared	0.253	Number of obs	9700
F-test	87.288	Prob > F	0.000
Akaike crit. (AIC)	118021.706	Bayesian crit. (BIC)	118143.764

*** $p < .01$, ** $p < .05$, * $p < .1$

Fonte: File di Stata dell'elaborazione dei dati del campione del dataset pulito, dataset fornito dalla relatrice Elisabetta Raguseo (Raguseo, 2025).

Tabella 26 Regressione caratterizzata da un modello con variabili di controllo ed il topic "Tipologia e caratteristiche dell'alloggio" con variabile dipendente revPAN

revPAN	Coef.	St.Err.	t-value	p-value	[95% Conf	Interval]	Sig
Tipologiaeacaratteristiche	207.576	148.063	1.40	.161	-82.658	497.811	
bedrooms	-7.72	4.837	-1.60	.111	-17.203	1.762	
bathrooms	38.812	6.504	5.97	0	26.064	51.561	***
maxguests	9.217	1.694	5.44	0	5.897	12.537	***
responserate	.177	.035	5.12	0	.109	.245	***
airbnbsuperhost	15.129	2.961	5.11	0	9.324	20.933	***
minimumstay	-.021	.023	-0.92	.356	-.067	.024	
ln_numberofreviews	6.637	.781	8.50	0	5.106	8.168	***
ln_numberofphotos	2.222	2.193	1.01	.311	-2.078	6.521	

instantbookenabled	29.664	2.454	12.09	0	24.855	34.474	***
overallrating	18.763	3.916	4.79	0	11.087	26.44	***
Entirehomeapt	-26.898	14.844	-1.81	.07	-55.997	2.2	*
Privateroom	-39.882	14.799	-2.69	.007	-68.891	-10.874	***
Sharedroom	-36.494	18.585	-1.96	.05	-72.924	-.065	**
Urbana	-2.894	2.783	-1.04	.299	-8.35	2.562	
Rurale	-26.435	3.657	-7.23	0	-33.602	-19.267	***
numero_caratteri	.149	.127	1.17	.241	-.1	.398	
Constant	-171.597	34.028	-5.04	0	-238.3	-104.894	***

Mean dependent var	67.197	SD dependent var	122.619
R-squared	0.254	Number of obs	9700
F-test	82.227	Prob > F	0.000
Akaike crit. (AIC)	118022.239	Bayesian crit. (BIC)	118151.477

*** $p < .01$, ** $p < .05$, * $p < .1$

Fonte: File di Stata dell'elaborazione dei dati del campione del dataset pulito, dataset fornito dalla relatrice Elisabetta Raguseo (Raguseo, 2025).

In tabella 26 (Raguseo, 2025) si evince che il topic “Tipologia e caratteristiche dell'alloggio” è correlato positivamente con la variabile dipendente (in quanto il coefficiente associato al topic è positivo), ma non è statisticamente significativo (in quanto $p\text{-value}=0,161$).

Tabella 27 Regressione caratterizzata da un modello con variabili di controllo ed il topic “Personalizzazione del titolo” con variabile dipendente revPAN

revPAN	Coef.	St.Err.	t-value	p-value	[95% Conf	Interval]	Sig
Personalizzationedeltitolo	3.37	86.228	0.04	.969	-165.656	172.395	
bedrooms	-7.683	4.834	-1.59	.112	-17.158	1.793	
bathrooms	38.762	6.503	5.96	0	26.014	51.509	***
maxguests	9.187	1.693	5.42	0	5.867	12.506	***
responserate	.179	.034	5.19	0	.111	.246	***
airbnbsuperhost	15.136	2.962	5.11	0	9.33	20.943	***
minimumstay	-.021	.023	-0.89	.371	-.066	.025	
ln_numberofreviews	6.673	.781	8.54	0	5.142	8.205	***
ln_numberofphotos	2.013	2.209	0.91	.362	-2.317	6.344	
instantbookenabled	29.787	2.462	12.10	0	24.96	34.613	***
overallrating	18.656	3.916	4.76	0	10.98	26.331	***

Entirehomeapt	-29.459	15.039	-1.96	.05	-58.939	.021	*
Privateroom	-40.297	14.964	-2.69	.007	-69.629	-10.965	***
Sharedroom	-37.298	18.715	-1.99	.046	-73.984	-.612	**
Urbana	-2.508	2.819	-0.89	.374	-8.033	3.017	
Rurale	-26.556	3.622	-7.33	0	-33.656	-19.455	***
numero_caratteri	.175	.123	1.42	.155	-.066	.415	
Constant	-135.265	29.193	-4.63	0	-192.489	-78.041	***

Mean dependent var	67.197	SD dependent var	122.619
R-squared	0.253	Number of obs	9700
F-test	83.416	Prob > F	0.000
Akaike crit. (AIC)	118023.705	Bayesian crit. (BIC)	118152.943

*** $p < .01$, ** $p < .05$, * $p < .1$

Fonte: File di Stata dell'elaborazione dei dati del campione del dataset pulito, dataset fornito dalla relatrice Elisabetta Raguseo (Raguseo, 2025).

In tabella 27 (Raguseo, 2025) si evince che il topic “Personalizzazione del titolo” è correlato positivamente con la variabile dipendente, in quanto il coefficiente associato al topic è positivo, ma non è statisticamente significativo.

Tabella 28 Regressione caratterizzata da un modello con variabili di controllo ed il topic “Posizione” con variabile dipendente revPAN

revPAN	Coef.	St.Err.	t-value	p-value	[95% Conf Interval]	Sig
Posizione	350.659	196.428	1.79	.074	-34.381 735.7	*
bedrooms	-7.659	4.836	-1.58	.113	-17.139 1.822	
bathrooms	38.707	6.494	5.96	0	25.978 51.436	***
maxguests	9.279	1.69	5.49	0	5.965 12.593	***
responserate	.18	.035	5.20	0	.112 .248	***
airbnbsuperhost	15.135	2.961	5.11	0	9.33 20.94	***
minimumstay	-.019	.023	-0.83	.407	-.064 .026	
ln_numberofreviews	6.551	.782	8.37	0	5.017 8.085	***
ln_numberofphotos	2.071	2.222	0.93	.351	-2.285 6.426	
instantbookenabed	29.724	2.469	12.04	0	24.885 34.563	***
overallrating	18.875	3.936	4.80	0	11.161 26.59	***
Entirehomeapt	-29.995	15.095	-1.99	.047	-59.584 -.406	**
Privateroom	-40.405	15.033	-2.69	.007	-69.873 -10.938	***
Sharedroom	-37.741	18.743	-2.01	.044	-74.481 -1.001	**

Urbana	-2.871	2.776	-1.03	.301	-8.313	2.57	
Rurale	-26.351	3.649	-7.22	0	-33.504	-19.197	***
numero_caratteri	.139	.127	1.09	.275	-.11	.387	
Constant	-194.75	42.079	-4.63	0	-277.233	-112.267	***

Mean dependent var	67.197	SD dependent var	122.619
R-squared	0.254	Number of obs	9700
F-test	82.400	Prob > F	0.000
Akaike crit. (AIC)	118020.979	Bayesian crit. (BIC)	118150.217

*** $p < .01$, ** $p < .05$, * $p < .1$

Fonte: File di Stata dell'elaborazione dei dati del campione del dataset pulito, dataset fornito dalla relatrice Elisabetta Raguseo (Raguseo, 2025).

In tabella 28 (Raguseo, 2025) si evince che il topic “Posizione” è correlato positivamente con la variabile dipendente, ed è statisticamente significativo all’10% in quanto $p\text{-value}=0,074$ e risulta essere, dunque, inferiore di 0,100.

Tabella 29 Regressione caratterizzata da un modello con variabili di controllo ed il topic “Comfort” con variabile dipendente revPAN

revPAN	Coef.	St.Err.	t-value	p-value	[95% Conf	Interval]	Sig
Comfort	370.486	224.405	1.65	.099	-69.394	810.367	*
bedrooms	-7.615	4.783	-1.59	.111	-16.992	1.762	
bathrooms	38.231	6.581	5.81	0	25.33	51.131	***
maxguests	9.021	1.663	5.43	0	5.763	12.28	***
responserate	.181	.035	5.22	0	.113	.249	***
airbnbsuperhost	15.132	2.963	5.11	0	9.325	20.94	***
minimumstay	-.025	.023	-1.08	.282	-.07	.02	
ln_numberofreviews	6.753	.778	8.68	0	5.227	8.278	***
ln_numberofphotos	1.084	2.06	0.53	.599	-2.954	5.122	
instantbookenabled	30.203	2.455	12.30	0	25.391	35.016	***
overallrating	18.249	3.95	4.62	0	10.507	25.992	***
Entirehomeapt	-29.47	15.07	-1.96	.051	-59.01	.069	*
Privateroom	-40.338	15.014	-2.69	.007	-69.769	-10.907	***
Sharedroom	-38.252	19.11	-2.00	.045	-75.712	-.792	**
Urbana	-1.016	2.772	-0.37	.714	-6.45	4.418	
Rurale	-27.432	3.741	-7.33	0	-34.765	-20.098	***
numero_caratteri	.152	.126	1.21	.227	-.095	.398	

Constant	-187.733	38.838	-4.83	0	-263.864	-111.602	***
Mean dependent var		67.197	SD dependent var		122.619		
R-squared		0.254	Number of obs		9700		
F-test		82.898	Prob > F		0.000		
Akaike crit. (AIC)		118013.836	Bayesian crit. (BIC)		118143.074		

*** $p < .01$, ** $p < .05$, * $p < .1$

Fonte: File di Stata dell'elaborazione dei dati del campione del dataset pulito, dataset fornito dalla relatrice Elisabetta Raguseo (Raguseo, 2025).

In tabella 29 (Raguseo, 2025) si evince che il topic “Comfort” è correlato positivamente con la variabile dipendente, ed è statisticamente significativo all'10%.

Tabella 30 Regressione caratterizzata da un modello con variabili di controllo ed il topic “Tipicità abitativa” con variabile dipendente revPAN

revPAN	Coef.	St.Err.	t-value	p-value	[95% Conf	Interval]	Sig
Tipicitàabitativa	-52.176	108.68	-0.48	.631	-265.212	160.861	
bedrooms	-7.695	4.834	-1.59	.111	-17.17	1.78	
bathrooms	38.751	6.491	5.97	0	26.027	51.474	***
maxguests	9.186	1.693	5.43	0	5.867	12.505	***
responserate	.179	.035	5.18	0	.111	.246	***
airbnbsuperhost	15.138	2.963	5.11	0	9.33	20.946	***
minimumstay	-.021	.023	-0.89	.374	-.066	.025	
ln_numberofreviews	6.659	.783	8.51	0	5.125	8.193	***
ln_numberofphotos	2.016	2.236	0.90	.367	-2.368	6.4	
instantbookenabled	29.825	2.469	12.08	0	24.986	34.664	***
overallrating	18.579	3.932	4.73	0	10.872	26.286	***
Entirehomeapt	-29.339	15.012	-1.95	.051	-58.766	.088	*
Privateroom	-40.201	14.956	-2.69	.007	-69.517	-10.885	***
Sharedroom	-37.337	18.736	-1.99	.046	-74.062	-.611	**
Urbana	-2.481	2.805	-0.88	.376	-7.979	3.017	
Rurale	-26.67	3.689	-7.23	0	-33.901	-19.44	***
numero_caratteri	.159	.13	1.23	.219	-.095	.414	
Constant	-125.131	31.434	-3.98	0	-186.749	-63.514	***
Mean dependent var		67.197	SD dependent var		122.619		
R-squared		0.253	Number of obs		9700		

F-test	82.700	Prob > F	0.000
Akaike crit. (AIC)	118023.479	Bayesian crit. (BIC)	118152.717

*** $p < .01$, ** $p < .05$, * $p < .1$

Fonte: File di Stata dell'elaborazione dei dati del campione del dataset pulito, dataset fornito dalla relatrice Elisabetta Raguseo (Raguseo, 2025).

In tabella 30 ([Raguseo, 2025](#)) si evince che il topic “Tipicità abitativa” è correlato negativamente con la variabile dipendente, ma non risulta essere statisticamente significativo.

Tabella 31 Regressione caratterizzata da un modello con variabili di controllo ed il topic “Zone rurali” con variabile dipendente revPAN

revPAN	Coef.	St.Err.	t-value	p-value	[95% Conf	Interval]	Sig
Zonerurali	-682.533	122.981	-5.55	0	-923.601	-441.465	***
bedrooms	-7.629	4.86	-1.57	.117	-17.155	1.898	
bathrooms	38.419	6.512	5.90	0	25.655	51.183	***
maxguests	9.266	1.697	5.46	0	5.941	12.592	***
responserate	.176	.035	5.07	0	.108	.244	***
airbnbsuperhost	15.141	2.956	5.12	0	9.345	20.936	***
minimumstay	-.023	.024	-0.97	.332	-.069	.023	
ln_numberofreviews	6.531	.783	8.34	0	4.996	8.067	***
ln_numberofphotos	2.701	2.225	1.21	.225	-1.661	7.062	
instantbookenabled	29.104	2.46	11.83	0	24.283	33.925	***
overallrating	19.596	3.908	5.01	0	11.936	27.257	***
Entirehomeapt	-28.607	14.897	-1.92	.055	-57.808	.593	*
Privateroom	-40.944	14.839	-2.76	.006	-70.032	-11.855	***
Sharedroom	-37.634	18.463	-2.04	.042	-73.826	-1.442	**
Urbana	-3.979	2.801	-1.42	.156	-9.47	1.512	
Rurale	-24.966	3.637	-6.86	0	-32.095	-17.837	***
numero_caratteri	.115	.122	0.94	.347	-.124	.354	
Constant	-24.77	30.503	-0.81	.417	-84.562	35.021	
Mean dependent var		67.197	SD dependent var		122.619		
R-squared		0.255	Number of obs		9700		
F-test		84.089	Prob > F		0.000		
Akaike crit. (AIC)		117998.079	Bayesian crit. (BIC)		118127.317		

*** $p < .01$, ** $p < .05$, * $p < .1$

Fonte: File di Stata dell'elaborazione dei dati del campione del dataset pulito, dataset fornito dalla relatrice Elisabetta Raguseo (Raguseo, 2025).

In tabella 31 ([Raguseo, 2025](#)) si evince che il topic “Zone rurali” è correlato negativamente con la variabile dipendente, ed è statisticamente significativo all’1%.

Terminata l’analisi delle singole correlazioni tra i topic e la variabile dipendente mediante il modello con i controlli, è stato realizzato il modello contenente cinque dei sei topic analizzati (per evitare problemi di collinearità non sono stati inseriti tutti i topic individuati), e si è deciso di omettere il topic “Tipicità abitativa” in modo coerente con quanto è stato effettuato per la variabile dipendente ADR.

Tabella 32 Regressione caratterizzata da un modello con variabili di controllo ed i topic con variabile dipendente revPAN

revPAN	Coef.	St.Err.	t-value	p-value	[95% Conf	Interval]	Sig
Tipologiaacaratteistiche	190.416	158.899	1.20	.231	-121.058	501.891	
Personalizzazione del titolo	60.358	106.898	0.56	.572	-149.185	269.901	
Posizione	305.112	241.551	1.26	.207	-168.378	778.602	
Comfort	529.649	229.787	2.30	.021	79.218	980.081	**
Zona rurali	-626.206	168.094	-3.73	0	-955.705	-296.706	***
bedrooms	-7.566	4.803	-1.58	.115	-16.981	1.849	
bathrooms	37.733	6.623	5.70	0	24.751	50.714	***
max guests	9.142	1.66	5.51	0	5.887	12.396	***
response rate	.178	.035	5.10	0	.109	.246	***
airbnb superhost	15.132	2.953	5.12	0	9.343	20.921	***
minimum stay	-.027	.023	-1.17	.243	-.073	.019	
ln_number of reviews	6.504	.788	8.26	0	4.96	8.049	***
ln_number of photos	1.733	2.08	0.83	.405	-2.345	5.81	
instant book enabled	29.544	2.442	12.10	0	24.758	34.331	***
overall rating	19.217	3.936	4.88	0	11.501	26.933	***
Entire home apt	-27.341	14.885	-1.84	.066	-56.518	1.836	*
Private room	-40.714	14.83	-2.75	.006	-69.785	-11.644	***
Shared room	-38.71	18.874	-2.05	.04	-75.707	-1.714	**
Urbana	-2.599	2.765	-0.94	.347	-8.02	2.822	
Rurale	-25.962	3.717	-6.99	0	-33.248	-18.676	***
numero_caratteri	.025	.132	0.19	.849	-.235	.285	
Constant	-205.561	98.371	-2.09	.037	-398.389	-12.733	**

Mean dependent var	67.197	SD dependent var	122.619
R-squared	0.257	Number of obs	9700
F-test	69.720	Prob > F	0.000
Akaike crit. (AIC)	117990.291	Bayesian crit. (BIC)	118148.248

*** $p < .01$, ** $p < .05$, * $p < .1$

Fonte: File di Stata dell'elaborazione dei dati del campione del dataset pulito, dataset fornito dalla relatrice Elisabetta Raguseo (Raguseo, 2025).

In tabella 32 ([Raguseo, 2025](#)) si evince che dei cinque topic presenti nel modello, due risultano statisticamente significativi e sono:

- “Comfort”, che risulta significativo all’5% (in quanto possiede un p-value=0,021) e correlato positivamente con la variabile dipendente;
- “Zone rurali”, che risulta significativo all’1% (in quanto possiede un p-value=0,000) e correlato negativamente con la variabile dipendente.

In merito alle variabili di controllo rappresentative delle caratteristiche dell’alloggio offerto, ovvero “Entirehomeapt”, “Privateroom”, “Sharedroom”, “Bedrooms”, “Bathrooms”, “Urbana” e “Rurale”, la variabile “Entirehomeapt” è statisticamente significativa al 10%, la variabile “Sharedroom” è statisticamente significativa al 5%, mentre le variabili “Privateroom”, “Bathrooms” e “Rurale” sono statisticamente significative all’1% nel modello in tabella 32 ([Raguseo, 2025](#)).

Relativamente alle variabili di controllo rappresentative delle caratteristiche dell’host, ovvero “Airbnbsuperhost” e “Responce rate”, entrambe risultano essere statisticamente significative all’1% nel modello in tabella 32 ([Raguseo, 2025](#)).

Delle variabili di controllo riferite alle caratteristiche dell’annuncio, ovvero “Maxguests”, “Minimumstay”, “ln_numberofphotos”, “Overallrating”, “Instantbookenabled”, “numero_caratteri” e “ln_numberofreviews”, solo le variabili “Maxguests”, “Overallrating”, “Instantbookenabled” e “ln_numberofreviews” risultano essere statisticamente significative all’1% nel modello in tabella 32 ([Raguseo, 2025](#)).

Tabella 33 Tabella riassuntiva dei modelli con i topic ed i controlli con variabile dipendente revPAN

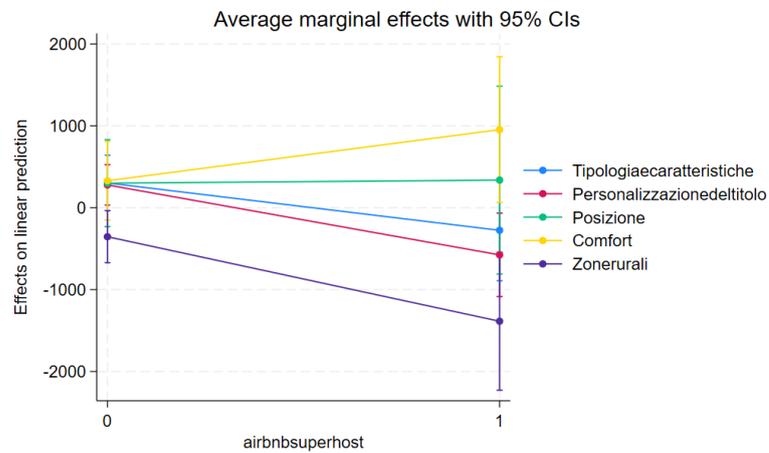
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	revPAN							
Controlli	Si							
Tipologia e caratteristiche dell'alloggio		207.6						190.4
		(148.1)						(158.9)
Personalizzazione del titolo			3.370					60.36
			(86.23)					(106.9)
Posizione				350.7*				305.1
				(196.4)				(241.6)
Comfort					370.5*			529.6**
					(224.4)			(229.8)
Tipicità abitativa						-52.18		
						(108.7)		
Zone rurali							-682.5***	-626.2***
							(123.0)	(168.1)
_cons	-134.7***	-171.6***	-135.3***	-194.8***	-187.7***	-125.1***	-24.77	-205.6**
	(23.92)	(34.03)	(29.19)	(42.08)	(38.84)	(31.43)	(30.50)	(98.37)
N	9700	9700	9700	9700	9700	9700	9700	9700

Standard errors in parentheses * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Fonte: File di Stata dell'elaborazione dei dati del campione del dataset pulito, dataset fornito dalla relatrice Elisabetta Raguseo (Raguseo, 2025).

Dopo aver analizzato le correlazioni tra la variabile dipendente ed i topic, sono state analizzate le moderazioni per ogni topic, delle variabili “Airbnbsuperhost”, “Overallrating”, “Urbana” e “Rurale”, in analogia con quanto studiato per la variabile dipendente ADR.

Figura 52 Grafico degli effetti marginali derivanti dalla variabile “Airbnbsuperhost” per il modello con variabile dipendente revPAN



Fonte: Screen di Stata dell'elaborazione dei dati del campione del dataset, dataset fornito dalla relatrice Elisabetta Raguseo (Raguseo, 2025).

Tabella 34 Effetti marginali derivanti dalla variabile “Airbnbsuperhost” per il modello con variabile dipendente revPAN

Average marginal effects
 Model VCE: Robust
 Expression: Linear prediction, predict()
 dy/dx wrt: Tipologiaecaratteristiche Personalizzazione del titolo Posizione Comfort Zonerurali
 Number of obs = 9,700

	Delta-method				
	dy/dx	std. err.	t	P>t	[95% conf. interval]
Tipologiaecaratteristiche					
che					
airbnbsuperhost					
0	303.022	172.488	1.760	0.079	-35.091 641.135

1	-275.184	314.747	-0.870	0.382	-892.155	341.786
Personalizzazione del titolo						
airbnbsuperhost						
0	278.662	126.223	2.210	0.027	31.238	526.086
1	-575.181	260.203	-2.210	0.027	-1085.234	-65.129
Posizione						
airbnbsuperhost						
0	299.787	270.963	1.110	0.269	-231.357	830.931
1	337.783	585.618	0.580	0.564	-810.151	1485.717
Comfort						
airbnbsuperhost						
0	329.833	244.728	1.350	0.178	-149.885	809.551
1	953.921	454.244	2.100	0.036	63.507	1844.335
Zone rurali						
airbnbsuperhost						
0	-353.668	162.071	-2.180	0.029	-671.361	-35.974
1	-1385.924	429.721	-3.230	0.001	-2228.266	-543.582

Fonte: File di Stata dell'elaborazione dei dati del campione del dataset pulito, dataset fornito dalla relatrice Elisabetta Raguseo (Raguseo, 2025).

Dall'analisi degli effetti marginali relativi alla moderazione della variabile "Airbnbsuperhost", si evince che la moderazione è statisticamente significativa solo per:

- Il topic "Personalizzazione del titolo", dove il topic viene moderato negativamente e in modo statisticamente significativo al 5%;
- Il topic "Comfort", dove il topic viene moderato positivamente e in modo statisticamente significativo al 5%;
- Il topic "Zone rurali", dove il topic viene moderato negativamente e in modo statisticamente significativo all'1%.

Tabella 35 Regressione con moderazione "Airbnbsuperhost" per il modello con variabile dipendente revPAN

revPAN	Coef.	St.Err.	t-value	p-value	[95% Conf Interval]	Sig
Tipologiae caratteristiche	303.022	172.488	1.76	.079	-35.091 641.135	*
Personalizzazione del titolo	278.662	126.223	2.21	.027	31.238 526.086	**

Posizione	299.787	270.963	1.11	.269	-231.357	830.931	
Comfort	329.833	244.728	1.35	.178	-149.885	809.551	
Zonerurali	-353.668	162.071	-2.18	.029	-671.361	-35.974	**
bedrooms	-7.461	4.843	-1.54	.123	-16.954	2.032	
bathrooms	37.591	6.681	5.63	0	24.496	50.686	***
maxguests	9.098	1.658	5.49	0	5.848	12.349	***
responserate	.176	.035	5.06	0	.108	.244	***
airnbnsuperhost	319.831	257.881	1.24	.215	-185.669	825.331	
minimumstay	-.028	.023	-1.23	.218	-.073	.017	
ln_numberofreviews	6.597	.788	8.37	0	5.052	8.142	***
ln_numberofphotos	1.434	2.087	0.69	.492	-2.656	5.525	
instantbookenabled	29.576	2.452	12.06	0	24.769	34.383	***
overallrating	19.417	3.942	4.93	0	11.689	27.145	***
Entirehomeapt	-26.097	14.825	-1.76	.078	-55.157	2.962	*
Privateroom	-38.996	14.768	-2.64	.008	-67.945	-10.048	***
Sharedroom	-36.798	18.502	-1.99	.047	-73.066	-.529	**
Urbana	-3.27	2.76	-1.18	.236	-8.679	2.14	
Rurale	-26.44	3.671	-7.20	0	-33.636	-19.244	***
numero_caratteri	.039	.133	0.29	.768	-.222	.3	
Airnbnsuperhost*tipologiaecaratteristiche	-578.206	337.819	-1.71	.087	-1240.402	83.99	*
Airnbnsuperhost*personalizzazione del titolo	-853.843	303.532	-2.81	.005	-1448.829	-258.858	***
Airnbnsuperhost*posizione	37.996	661.937	0.06	.954	-1259.539	1335.531	
Airnbnsuperhost*comfort	624.088	480.822	1.30	.194	-318.423	1566.6	
Airnbnsuperhost*zonerurali	-1032.257	449.853	-2.29	.022	-1914.063	-150.45	**
Constant	-274.564	107.942	-2.54	.011	-486.153	-62.975	**

Mean dependent var	67.197	SD dependent var	122.619
R-squared	0.259	Number of obs	9700
F-test	57.245	Prob > F	0.000
Akaike crit. (AIC)	117967.052	Bayesian crit. (BIC)	118160.909

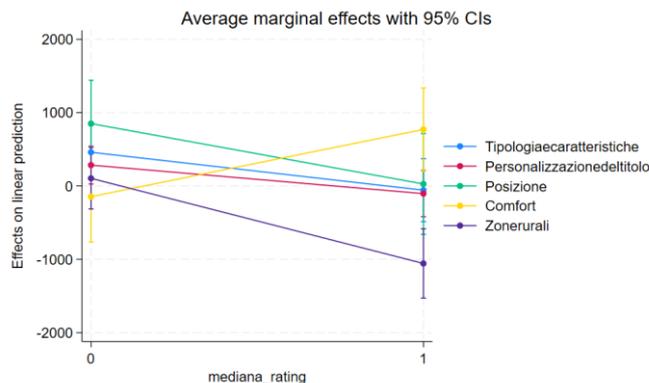
*** $p < .01$, ** $p < .05$, * $p < .1$

Fonte: File di Stata dell'elaborazione dei dati del campione del dataset pulito, dataset fornito dalla relatrice Elisabetta Raguseo (Raguseo, 2025).

Analizzando la moderazione della variabile “Airbnbsuperhost” nel modello in cui sono presenti anche i controlli, si verifica che la moderazione è statisticamente significativa per:

- Il topic “Tipologia e caratteristiche dell’alloggio”, dove il topic è moderato negativamente ed in modo statisticamente significativo al 10%. La presenza del badge superhost, dunque, riduce gli effetti positivi sulla variabile dipendente dovuti alla definizione della tipologia e delle caratteristiche dell’alloggio nel titolo di un annuncio;
- Il topic “Personalizzazione del titolo”, dove il topic è moderato negativamente ed in modo statisticamente significativo all’1%. La presenza del badge superhost, dunque, riduce gli effetti positivi sulla variabile dipendente dovuti alla personalizzazione del titolo degli annunci;
- Il topic “Zone rurali”, dove il topic è moderato negativamente ed in modo statisticamente significativo al 5%. La presenza del badge superhost, dunque, riduce gli effetti negativi sulla variabile dipendente dovuti ai riferimenti alle zone rurali presenti nei titoli degli annunci.

Figura 51 Grafico degli effetti marginali derivanti dalla variabile “Mediana_rating” per il modello con variabile dipendente revPAN



Fonte: Screen di Stata dell’elaborazione dei dati del campione del dataset pulito, dataset fornito dalla relatrice Elisabetta Raguseo (Raguseo, 2025).

Tabella 36 Effetti marginali derivanti dalla variabile “Mediana_rating” per il modello con variabile dipendente revPAN

Average marginal effects
Model VCE: Robust

Number of obs = 9,700

Expression: Linear prediction, predict()

dy/dx wrt: Tipologiaeacaratteristiche Personalizzazione del titolo Posizione Comfort Zonerurali

		Delta-method				
		dy/dx	std. err.	t	P>t	[95% conf. interval]
Tipologiaeacaratteristiche						
mediana_rating						
0		459.682	185.214	2.480	0.013	96.624 822.739
1		-57.146	219.268	-0.260	0.794	-486.956 372.665
Personalizzazione del titolo						
mediana_rating						
0		284.380	131.836	2.160	0.031	25.955 542.806
1		-104.029	160.922	-0.650	0.518	-419.470 211.412
Posizione						
mediana_rating						
0		850.966	300.779	2.830	0.005	261.377 1440.555
1		27.152	350.603	0.080	0.938	-660.103 714.407
Comfort						
mediana_rating						
0		-146.229	315.637	-0.460	0.643	-764.944 472.486
1		770.987	288.920	2.670	0.008	204.642 1337.331
Zonerurali						
mediana_rating						
0		104.288	213.641	0.490	0.625	-314.494 523.070
1		-1056.669	240.895	-4.390	0.000	-1528.874 -584.465

Fonte: File di Stata dell'elaborazione dei dati del campione del dataset pulito, dataset fornito dalla relatrice Elisabetta Raguseo (Raguseo, 2025).

Dall'analisi degli effetti marginali della variabile "Mediana_rating", si può notare che la moderazione è statisticamente significativa solo per:

- Il topic "Comfort", il quale è moderato positivamente ed in modo statisticamente significativo all'1%;
- Il topic "Zone rurali", il quale è moderato negativamente ed in modo statisticamente significativo all'1%.

Tabella 37 Regressione con moderazione "Mediana_rating" per il modello con variabile dipendente revPAN

revPAN	Coef.	St.Err.	t-value	p-value	[95% Conf	Interval]	Sig
Tipologiaeacaratteristiche	459.682	185.214	2.48	.013	96.624	822.739	**
Personalizzazione del titolo	284.38	131.836	2.16	.031	25.955	542.806	**
Posizione	850.966	300.779	2.83	.005	261.377	1440.555	***
Comfort	-146.229	315.637	-0.46	.643	-764.944	472.486	
Zonerurali	104.288	213.641	0.49	.625	-314.494	523.07	
bedrooms	-7.219	4.772	-1.51	.13	-16.574	2.135	
bathrooms	37.121	6.615	5.61	0	24.154	50.088	***
maxguests	9.233	1.644	5.62	0	6.011	12.455	***
responserate	.18	.035	5.20	0	.112	.248	***
airbnbsuperhost	12.978	2.997	4.33	0	7.104	18.853	***
minimumstay	-.025	.025	-1.01	.315	-.074	.024	
ln_numberofreviews	7.234	.781	9.26	0	5.702	8.766	***
ln_numberofphotos	1.145	2.08	0.55	.582	-2.933	5.223	
instantbookenabled	29.378	2.444	12.02	0	24.587	34.17	***
overallrating	6.962	4.756	1.46	.143	-2.361	16.284	
Entirehomeapt	-26.664	14.85	-1.80	.073	-55.774	2.445	*
Privateroom	-39.816	14.866	-2.68	.007	-68.957	-10.676	***
Sharedroom	-37.814	18.881	-2.00	.045	-74.825	-.803	**
Urbana	-2.82	2.745	-1.03	.304	-8.201	2.561	
Rurale	-26.352	3.706	-7.11	0	-33.617	-19.088	***
numero_caratteri	.031	.132	0.23	.816	-.228	.29	
mediana_rating*tipologiaeacaratteristiche	-516.827	257.071	-2.01	.044	-1020.741	-12.914	**
mediana_rating*personalizzazione del titolo	-388.409	204.358	-1.90	.057	-788.994	12.176	*
mediana_rating*posizione	-823.814	444.052	-1.86	.064	-1694.25	46.621	*
mediana_rating*comfort	917.216	393.313	2.33	.02	146.239	1688.192	**
mediana_rating*zonerurali	-1160.957	320.911	-3.62	0	-1790.011	-531.904	***
mediana_rating	355.134	171.203	2.07	.038	19.54	690.728	**
Constant	-347.73	119.518	-2.91	.004	-582.011	-113.45	***

Mean dependent var	67.197	SD dependent var	122.619
R-squared	0.261	Number of obs	9700
F-test	55.647	Prob > F	0.000
Akaike crit. (AIC)	117944.891	Bayesian crit. (BIC)	118145.928

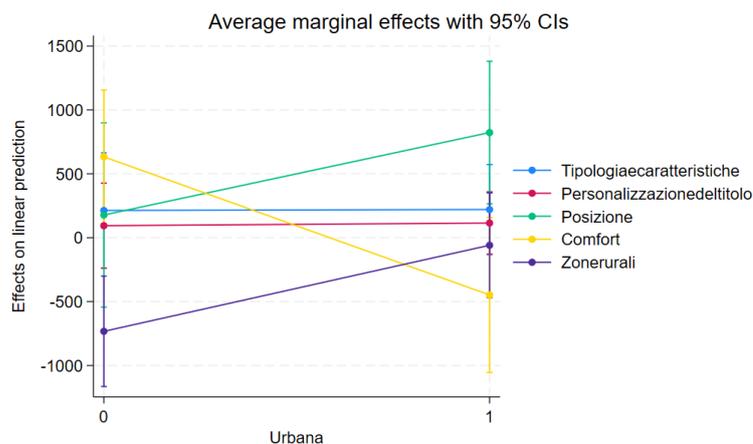
*** $p < .01$, ** $p < .05$, * $p < .1$

Fonte: File di Stata dell'elaborazione dei dati del campione del dataset pulito, dataset fornito dalla relatrice Elisabetta Raguseo (Raguseo, 2025).

Analizzando la moderazione della variabile “Mediana_rating” nel modello in cui sono presenti anche i controlli, si verifica che la moderazione è statisticamente significativa per:

- Il topic “Tipologia e caratteristiche dell'alloggio”, il quale è moderato negativamente ed in modo statisticamente significativo al 5%. La moderazione negativa riduce gli effetti positivi del topic in questione;
- Il topic “Personalizzazione del titolo”, il quale è moderato negativamente ed in modo statisticamente significativo al 10%. La moderazione negativa riduce gli effetti positivi del topic in questione;
- Il topic “Posizione”, il quale è moderato negativamente ed in modo statisticamente significativo al 10%. La moderazione negativa riduce gli effetti positivi del topic in questione;
- Il topic “Comfort”, il quale è moderato positivamente ed in modo statisticamente significativo al 5%. Il topic “Comfort”, in questo caso, inverte il segno della correlazione: ovvero passa dall'essere correlato positivamente alla variabile dipendente all'essere correlato negativamente. Essendo, dunque, il topic correlato negativamente con la variabile dipendente, la moderazione positiva incrementa gli effetti negativi associati al topic;
- Il topic “Zone rurali”, il quale è moderato negativamente ed in modo statisticamente significativo all'1%. Il topic “Zone rurali”, in questo caso, inverte il segno della correlazione: ovvero passa dall'essere correlato negativamente alla variabile dipendente all'essere correlato positivamente. Essendo, dunque, il topic correlato positivamente con la variabile dipendente, la moderazione negativa diminuisce gli effetti positivi associati al topic.

Figura 54 Grafico degli effetti marginali derivanti dalla variabile "Urbana" per il modello con variabile dipendente revPAN



Fonte: Screen di Stata dell'elaborazione dei dati del campione del dataset pulito, dataset fornito dalla relatrice Elisabetta Raguseo (Raguseo, 2025).

Tabella 38 Effetti marginali derivanti dalla variabile "Urbana" per il modello con variabile dipendente revPAN

Average marginal effects Number of obs = 9,700
 Model VCE: Robust
 Expression: Linear prediction, predict()
 dy/dx wrt: Tipologiaecaratteristiche Personalizzazioneedeltitolo Posizione Comfort Zonerurali

	Delta-method				
	dy/dx	std. err.	t	P>t	[95% conf. interval]
Tipologiaecaratteristiche					
Urbana					
0	212.989	230.041	0.930	0.355	-237.939 663.916
1	220.388	179.751	1.230	0.220	-131.960 572.737
Personalizzazioneedeltitolo					
Urbana					
0	93.801	169.797	0.550	0.581	-239.036 426.637
1	114.962	124.543	0.920	0.356	-129.169 359.093
Posizione					
Urbana					
0	177.371	367.814	0.480	0.630	-543.622 898.364
1	822.806	284.679	2.890	0.004	264.776 1380.836
Comfort					

Urbana						
0	632.905	266.409	2.380	0.018	110.689	1155.122
1	-448.033	309.305	-1.450	0.148	-1054.336	158.270
Zonerurali						
Urbana						
0	-732.910	220.351	-3.330	0.001	-1164.844	-300.976
1	-59.136	209.248	-0.280	0.777	-469.305	351.034

Fonte: File di Stata dell'elaborazione dei dati del campione del dataset pulito, dataset fornito dalla relatrice Elisabetta Raguseo (Raguseo, 2025).

Dall'analisi degli effetti marginali della variabile "Urbana", si può notare che la moderazione è statisticamente significativa solo per il topic "Posizione", il quale risulta essere moderato positivamente ed in modo statisticamente significativo all'1%.

Tabella 39 Regressione con moderazione "Urbana" per il modello con variabile dipendente revPAN

revPAN	Coef.	St.Err.	t-value	p-value	[95% Conf	Interval]	Sig
Tipologiaeaccharatteristiche	212.989	230.041	0.93	.355	-237.939	663.916	
Personalizzazione del titolo	93.801	169.797	0.55	.581	-239.036	426.637	
Posizione	177.371	367.814	0.48	.63	-543.622	898.364	
Comfort	632.905	266.408	2.38	.018	110.689	1155.122	**
Zonerurali	-732.91	220.351	-3.33	.001	-1164.844	-300.976	***
bedrooms	-7.528	4.828	-1.56	.119	-16.992	1.937	
bathrooms	37.75	6.627	5.70	0	24.759	50.74	***
maxguests	9.118	1.67	5.46	0	5.844	12.392	***
responserate	.175	.035	5.04	0	.107	.243	***
airbnbsuperhost	15.148	2.961	5.12	0	9.344	20.952	***
minimumstay	-.028	.023	-1.22	.223	-.074	.017	
ln_numberofreviews	6.46	.788	8.20	0	4.915	8.004	***
ln_numberofphotos	1.454	2.084	0.70	.485	-2.63	5.538	
instantbookenabled	29.682	2.486	11.94	0	24.81	34.555	***
overallrating	19.21	3.925	4.89	0	11.516	26.904	***
Entirehomeapt	-27.029	14.818	-1.82	.068	-56.076	2.017	*
Privateroom	-39.617	14.741	-2.69	.007	-68.512	-10.721	***

Sharedroom	-39.953	18.839	-2.12	.034	-76.882	-3.024	**
Urbana	-61.542	167.941	-0.37	.714	-390.741	267.657	
Rurale	-25.984	3.742	-6.94	0	-33.319	-18.648	***
numero_caratteri	.03	.132	0.23	.819	-.229	.289	
Urbana*tipologiae caratteristiche	7.4	262.074	0.03	.977	-506.32	521.12	
Urbana*personalizzazione del titolo	21.162	208.609	0.10	.919	-387.755	430.078	
Urbana*posizione	645.436	444.65	1.45	.147	-226.172	1517.043	
Urbana*comfort	-1080.938	379.956	-2.84	.004	-1825.731	-336.145	***
Urbana*zone rurali	673.775	298.554	2.26	.024	88.546	1259.003	**
Constant	-190.994	146.514	-1.30	.192	-478.193	96.205	
<hr/>							
Mean dependent var		67.197	SD dependent var			122.619	
R-squared		0.257	Number of obs			9700	
F-test		57.304	Prob > F			0.000	
Akaike crit. (AIC)		117989.072	Bayesian crit. (BIC)			118182.929	

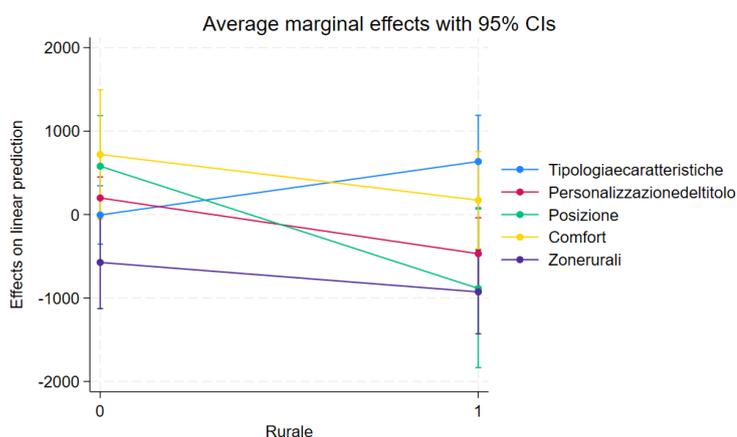
*** $p < .01$, ** $p < .05$, * $p < .1$

Fonte: File di Stata dell'elaborazione dei dati del campione del dataset pulito, dataset fornito dalla relatrice Elisabetta Raguseo (Raguseo, 2025).

Analizzando la moderazione della variabile “Urbana” nel modello in cui sono presenti anche i controlli, si verifica che la moderazione è statisticamente significativa per:

- Il topic “Comfort”, il quale risulta essere moderato negativamente ed in modo statisticamente significativo all’1%. La moderazione negativa riduce gli effetti positivi del topic in questione;
- Il topic “Zone rurali”, il quale risulta essere moderato positivamente ed in modo statisticamente significativo al 5%. La moderazione positiva incrementa gli effetti negativi del topic in questione.

Figura 55 Grafico degli effetti marginali derivanti dalla variabile "Rurale" per il modello con variabile dipendente revPAN



Fonte: Screen di Stata dell'elaborazione dei dati del campione del dataset pulito, dataset fornito dalla relatrice Elisabetta Raguseo (Raguseo, 2025).

Tabella 40 Effetti marginali derivanti dalla variabile "Rurale" per il modello con variabile dipendente revPAN

Average marginal effects Number of obs = 9,700
 Model VCE: Robust
 Expression: Linear prediction, predict()
 dy/dx wrt: Tipologiaeacaratteristiche Personalizzazione del titolo Posizione Comfort Zonerurali

	Delta-method				
	dy/dx	std. err.	t	P>t	[95% conf. interval]
Tipologiaeacaratteristiche					
Rurale					
0	-4.825	178.228	-0.030	0.978	-354.189 344.539
1	635.658	282.690	2.250	0.025	81.527 1189.790
Personalizzazione del titolo					
Rurale					
0	198.693	128.283	1.550	0.121	-52.768 450.154
1	-469.287	220.051	-2.130	0.033	-900.633 -37.940
Posizione					
Rurale					
0	579.382	309.058	1.870	0.061	-26.436 1185.200
1	-884.575	484.323	-1.830	0.068	-1833.950 64.800
Comfort					
Rurale					

0	720.098	395.370	1.820	0.069	-54.910	1495.107
1	172.070	297.016	0.580	0.562	-410.143	754.283
Zonerurali						
Rurale						
0	-573.716	282.441	-2.030	0.042	-1127.359	-20.074
1	-926.734	255.950	-3.620	0.000	-1428.448	-425.019

Fonte: File di Stata dell'elaborazione dei dati del campione del dataset pulito, dataset fornito dalla relatrice Elisabetta Raguseo (Raguseo, 2025).

Dall'analisi degli effetti marginali della variabile "Rurale", si può notare che la moderazione è statisticamente significativa per:

- Il topic "Tipologia e caratteristiche dell'alloggio", il quale risulta essere moderato positivamente ed in modo statisticamente significativo al 5%;
- Il topic "Personalizzazione del titolo", il quale risulta essere moderato negativamente ed in modo statisticamente significativo al 5%;
- Il topic "Posizione", il quale risulta essere moderato negativamente ed in modo statisticamente significativo al 10%;
- Il topic "Zone rurali", il quale risulta essere moderato negativamente ed in modo statisticamente significativo all'1%.

Tabella 41 Regressione con moderazione "Rurale" per il modello con variabile dipendente revPAN

revPAN	Coef.	St.Err.	t-value	p-value	[95% Conf	Interval]	Sig
Tipologiaeacaratteristiche	-4.825	178.228	-0.03	.978	-354.189	344.539	
Personalizzazione del titolo	198.693	128.283	1.55	.121	-52.768	450.154	
Posizione	579.382	309.058	1.87	.061	-26.436	1185.2	*
Comfort	720.098	395.37	1.82	.069	-54.91	1495.107	*
Zonerurali	-573.716	282.441	-2.03	.042	-1127.359	-20.074	**
bedrooms	-7.463	4.821	-1.55	.122	-16.913	1.987	
bathrooms	37.562	6.63	5.67	0	24.565	50.559	***
maxguests	9.161	1.664	5.51	0	5.9	12.422	***
responserate	.179	.035	5.09	0	.11	.248	***
airbnbsuperhost	14.857	2.918	5.09	0	9.137	20.578	***
minimumstay	-.022	.025	-0.90	.37	-.071	.026	

ln_numberofreviews	6.625	.77	8.60	0	5.115	8.135	***
ln_numberofphotos	1.585	2.066	0.77	.443	-2.466	5.635	
instantbookenable	29.476	2.469	11.94	0	24.636	34.315	***
overallrating	19.042	3.94	4.83	0	11.318	26.765	***
Entirehomeapt	-26.403	15.172	-1.74	.082	-56.144	3.338	*
Privateroom	-39.66	15.141	-2.62	.009	-69.339	-9.981	***
Sharedroom	-36.054	19.407	-1.86	.063	-74.095	1.987	*
Urbana	-1.824	2.782	-0.66	.512	-7.276	3.629	
Rurale	379.05	230.022	1.65	.099	-71.841	829.941	*
numero_caratteri	.064	.131	0.49	.624	-.193	.321	
Rurale*tipologiae caratteristiche	640.484	311.332	2.06	.04	30.208	1250.76	**
Rurale*personalizzazione del titolo	-667.98	255.603	-2.61	.009	-1169.015	-166.944	***
Rurale*posizione	-1463.957	569.559	-2.57	.01	-2580.411	-347.503	**
Rurale*comfort	-548.028	469.992	-1.17	.244	-1469.31	373.254	
Rurale*zonerurali	-353.017	378.21	-0.93	.351	-1094.387	388.353	
Constant	-284.346	105.67	-2.69	.007	-491.481	-77.211	***

Mean dependent var	67.197	SD dependent var	122.619
R-squared	0.258	Number of obs	9700
F-test	60.826	Prob > F	0.000
Akaike crit. (AIC)	117982.525	Bayesian crit. (BIC)	118176.382

*** $p < .01$, ** $p < .05$, * $p < .1$

Fonte: File di Stata dell'elaborazione dei dati del campione del dataset pulito, dataset fornito dalla relatrice Elisabetta Raguseo (Raguseo, 2025).

Analizzando la moderazione della variabile “Rurale” nel modello in cui sono presenti anche i controlli, si verifica che la moderazione è statisticamente significativa per:

- Il topic “Tipologia e caratteristiche dell'alloggio”, il quale risulta essere moderato positivamente ed in modo statisticamente significativo al 5%. Il topic “Tipologia e caratteristiche dell'alloggio”, in questo caso, inverte il segno della correlazione: ovvero passa dall'essere correlato positivamente alla variabile dipendente all'essere correlato negativamente. Essendo, dunque, il topic correlato

negativamente con la variabile dipendente, la moderazione positiva incrementa gli effetti negativi associati al topic;

- Il topic “Personalizzazione del titolo”, il quale risulta essere moderato negativamente ed in modo statisticamente significativo all’1%. La moderazione negativa riduce gli effetti positivi del topic in questione;
- Il topic “Posizione”, il quale risulta essere moderato negativamente ed in modo statisticamente significativo al 5%. La moderazione negativa riduce gli effetti positivi del topic in questione.

Tabella 42 Tabella riassuntiva dei modelli con controlli e moderazioni con variabile dipendente revPAN

	(1)	(2)	(3)	(4)
	revPAN	revPAN	revPAN	revPAN
Tipologia e caratteristiche dell'alloggio	303.0*	459.7**	-4.825	213.0
	(172.5)	(185.2)	(178.2)	(230.0)
Personalizzazione del titolo	278.7**	284.4**	198.7	93.80
	(126.2)	(131.8)	(128.3)	(169.8)
Posizione	299.8	851.0***	579.4*	177.4
	(271.0)	(300.8)	(309.1)	(367.8)
Comfort	329.8	-146.2	720.1*	632.9**
	(244.7)	(315.6)	(395.4)	(266.4)
Zone rurali	-353.7**	104.3	-573.7**	-732.9***
	(162.1)	(213.6)	(282.4)	(220.4)
1.airbnbsuperhost#c.Tipologiaeacaratteristiche	-578.2*			
	(337.8)			
1.airbnbsuperhost#c.Personalizzazionedeltitolo	-853.8***			
	(303.5)			
1.airbnbsuperhost#c.Posizione	38.00			
	(661.9)			
1.airbnbsuperhost#c.Comfort	624.1			
	(480.8)			
1.airbnbsuperhost#c.Zonerurali	-1032.3**			
	(449.9)			
1.mediana_rating#c.Tipologiaeacaratteristiche		-516.8**		
		(257.1)		
1.mediana_rating#c.Personalizzazionedeltitolo		-388.4*		
		(204.4)		
1.mediana_rating#c.Posizione		-823.8*		

		(444.1)		
1.mediana rating#c.Comfort		917.2**		
		(393.3)		
1.mediana rating#c.Zonerurali		-1161.0***		
		(320.9)		
1.Rurale#c.Tipologiaeacaratteristiche			640.5**	
			(311.3)	
1.Rurale#c.Personalizzazionedeltitolo			-668.0***	
			(255.6)	
1.Rurale#c.Posizione			-1464.0**	
			(569.6)	
1.Rurale#c.Comfort			-548.0	
			(470.0)	
1.Rurale#c.Zonerurali			-353.0	
			(378.2)	
1.Urbana#c.Tipologiaeacaratteristiche				7.400
				(262.1)
1.Urbana#c.Personalizzazionedeltitolo				21.16
				(208.6)
1.Urbana#c.Posizione				645.4
				(444.7)
1.Urbana#c.Comfort				-1080.9***
				(380.0)
1.Urbana#c.Zonerurali				673.8**
				(298.6)
Controlli	Si	Si	Si	Si
_Cons	-274.6**	-347.7***	-284.3***	-191.0
	(107.9)	(119.5)	(105.7)	(146.5)
N	9700	9700	9700	9700

Standard errors in parentheses * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Fonte: File di Stata dell'elaborazione dei dati del campione del dataset pulito, dataset fornito dalla relatrice Elisabetta Raguseo (Raguseo, 2025).

Capitolo 4 – Discussione dei risultati

L'analisi delle caratteristiche semantiche del titolo, in relazione alle variabili prestazionali ADR e revPAN, ha portato ad una serie di risultati interessanti. Innanzitutto, sono stati individuati sei topic mediante un algoritmo di clustering che combina modelli di SBERT e K-Means: “Posizione”, “Comfort”, “Zone rurali”, “Tipicità abitativa”, “Personalizzazione del titolo” e “Tipologia e caratteristiche dell'alloggio”. Di tali topic sono state dapprima analizzate le singole correlazioni con le variabili dipendenti, per poi generare un modello complessivo con cinque dei sei topic per evitare problemi di collinearità. Dall'analisi delle singole correlazioni:

- Per la variabile ADR risultano statisticamente significativi i seguenti topic: “Personalizzazione del titolo” (correlazione negativa), “Posizione” (correlazione positiva), “Comfort” (correlazione positiva), “Tipicità abitativa” (correlazione negativa) e “Zone rurali” (correlazione negativa);
- Per la variabile revPAN risultano statisticamente significativi i seguenti topic: “Posizione” (correlazione positiva), “Comfort” (correlazione positiva) e “Zone rurali” (correlazione negativa).

Considerando, invece, il modello complessivo che comprende tutti i topic escluso “Tipicità abitativa”, si verifica che:

- Per la variabile ADR risultano statisticamente significativi i seguenti topic: “Tipologia e caratteristiche dell'alloggio” (correlazione positiva), “Posizione” (correlazione positiva), “Comfort” (correlazione positiva) e “Zone rurali” (correlazione negativa);
- Per la variabile revPAN risultano statisticamente significativi i seguenti topic: “Comfort” (correlazione positiva) e “Zone rurali” (correlazione negativa).

Ne consegue che, per migliorare i ricavi medi per giorni prenotati, risulta consigliabile inserire nel titolo di un annuncio Airbnb informazioni come: la tipologia e le caratteristiche dell'alloggio, il posizionamento dell'alloggio ed i comfort ad esso associati. Il riferimento alle caratteristiche dell'alloggio, all'interno del titolo di un

annuncio, è un'informazione che impatta positivamente la variabile ADR in quanto, tali informazioni potrebbero, presumibilmente, catturare l'attenzione del cliente spingendolo a visionare l'annuncio stesso. La localizzazione dell'alloggio, sebbene possa essere inserita come filtro di ricerca degli annunci, risulta essere un argomento che ha un impatto significativo sulla variabile dipendente ADR; tale risultato potrebbe essere spiegato in quanto la posizione precisa dell'alloggio viene fornita dalla piattaforma solo in seguito alla conferma della prenotazione, dunque, avere maggiori informazioni circa il luogo in cui è situato l'alloggio, potrebbe essere percepito come qualcosa di positivo. All'aumentare di informazioni circa la presenza di comfort aumenta il guadagno rispetto ai giorni prenotati. La permanenza della significatività dell'argomento "Comfort" nei modelli, potrebbe essere spiegata dalla motivazione di utilizzo di Airbnb, in quanto [Airbnb \(n.d.\)](#) si offre come *"un mondo in cui tutti possano sentirsi a casa, ovunque"* ([Airbnb, n.d.](#)); si potrebbe pensare, dunque, che all'aumentare di comfort offerti possa aumentare l'attrattiva dell'annuncio per il cliente che ricerca un'esperienza tale per cui possa sentirsi a casa anche quando è in viaggio. Risulta, inoltre, alla luce dei risultati ottenuti, consigliabile ridurre informazioni circa la ruralità dell'alloggio, in quanto il rispettivo topic presenta una correlazione negativa e statisticamente significativa rispetto ai ricavi medi per giorni prenotati. La negatività della correlazione tra il topic "Zone rurali" e la variabile dipendente, potrebbe essere dovuta alla superfluità delle informazioni, ovvero: informazioni circa la ruralità od il relax potrebbero essere considerate dai clienti come informazioni non necessarie per la scelta dell'alloggio o ancora come informazioni che potrebbero essere presenti in altre parti dell'annuncio, come la descrizione, e, quindi, non essere considerate necessarie nel titolo.

Per migliorare i ricavi medi per giorni disponibili, alla luce dei risultati ottenuti, potrebbe essere utile inserire nel titolo degli annunci informazioni circa i comfort offerti dall'alloggio e ridurre informazioni circa la ruralità dell'alloggio. Si può notare come, il riferimento ad informazioni circa i comfort, risulti essere impattate positivamente sia in termini di ricavi medi per giorni prenotati, sia in termini di ricavi medi per giorni disponibili nel complesso, determinando, dunque, un duplice effetto positivo. Il riferimento alla ruralità, nei titoli degli annunci, in analogia con quanto accade per i comfort, determina un duplice effetto negativo in termini di prestazioni degli annunci Airbnb, in quanto il topic "Zone rurali" è correlato negativamente ed in modo statisticamente significativo, sia alla variabile ADR sia alla variabile revPAN.

In termini di moderazioni si verifica che:

- La variabile “Airbnb Superhost” modera:
 - Nel caso di variabile dipendente ADR:
 - Il topic “Tipologia e caratteristiche dell’alloggio”. La moderazione è negativa, il che determina una riduzione degli effetti positivi;
 - Il topic “Comfort”. La moderazione è positiva, il che determina un incremento degli effetti positivi;
 - Il topic “Zone rurali”. La moderazione è negativa, il che determina una riduzione degli effetti negativi.
 - Nel caso di variabile dipendente revPAN:
 - Il topic “Tipologia e caratteristiche dell’alloggio”. La moderazione è negativa, il che determina una riduzione degli effetti positivi;
 - Il topic “Personalizzazione del titolo”. La moderazione è negativa, il che determina una riduzione degli effetti positivi;
 - Il topic “Zone rurali”. La moderazione è negativa, il che determina una riduzione degli effetti negativi.
- La variabile “mediana_rating” modera:
 - Nel caso di variabile dipendente ADR:
 - Il topic “Zone rurali”. La moderazione è negativa, il che determina una riduzione degli effetti negativi.
 - Nel caso di variabile dipendente revPAN:
 - Il topic “Tipologia e caratteristiche dell’alloggio”. La moderazione è negativa, il che determina una riduzione degli effetti positivi;
 - Il topic “Personalizzazione del titolo”. La moderazione è negativa, il che determina una riduzione degli effetti positivi;
 - Il topic “Posizione”. La moderazione è negativa, il che determina una riduzione degli effetti positivi;
 - Il topic “Comfort”. La moderazione è positiva, il che determina un incremento degli effetti negativi. Il topic “Comfort” che nel modello in tabella 32 ([Raguseo, 2025](#)), ovvero il modello in cui sono presenti sia i topic sia i controlli, presentava una correlazione positiva, nel modello in tabella 37 ([Raguseo, 2025](#)) dove viene inserita la moderazione del rating, la correlazione diventa negativa.

Il cambiamento del segno della correlazione del topic con la variabile dipendente potrebbe indicare il forte effetto moderatore che il rating ha sul topic “Comfort”;

- Il topic “Zone rurali”. La moderazione è negativa, il che determina una riduzione degli effetti positivi. Il topic “Zone rurali” che nel modello in tabella 32 ([Raguseo, 2025](#)), ovvero il modello in cui sono presenti sia i topic sia i controlli, presentava una correlazione negativa, nel modello in tabella 37 ([Raguseo, 2025](#)) dove viene inserita la moderazione del rating, la correlazione diventa positiva. Il cambiamento del segno della correlazione del topic con la variabile dipendente potrebbe indicare il forte effetto moderatore che il rating ha sul topic “Zone rurali”.
- La variabile “Urbana” modera:
 - Nel caso di variabile dipendente ADR:
 - Il topic “Personalizzazione del titolo”. La moderazione è negativa, il che determina una riduzione degli effetti positivi;
 - Il topic “Posizione”. La moderazione è negativa, il che determina una riduzione degli effetti positivi;
 - Il topic “Comfort”. La moderazione è negativa, il che determina una riduzione degli effetti positivi;
 - Il topic “Zone rurali”. La moderazione è positiva, il che determina un incremento degli effetti negativi.
 - Nel caso di variabile dipendente revPAN:
 - Il topic “Comfort”. La moderazione è negativa, il che determina una riduzione degli effetti positivi;
 - Il topic “Zone rurali”. La moderazione è positiva, il che determina un incremento degli effetti negativi.
- La variabile “Rurale” modera:
 - Nel caso di variabile dipendente ADR:
 - Il topic “Personalizzazione del titolo”. La moderazione è negativa, il che determina una riduzione degli effetti positivi.
 - Nel caso di variabile dipendente revPAN:

- Il topic “Tipologia e caratteristiche dell’alloggio”. La moderazione è positiva, il che determina un incremento degli effetti negativi. Il topic “Tipologia e caratteristiche dell’alloggio” che nel modello in tabella 32 ([Raguseo, 2025](#)), ovvero il modello in cui sono presenti sia i topic sia i controlli, presentava una correlazione positiva, nel modello in tabella 41 ([Raguseo, 2025](#)) dove viene inserita la moderazione della variabile “Rurale”, la correlazione diventa negativa. Il cambiamento del segno della correlazione del topic con la variabile dipendente potrebbe indicare il forte effetto moderatore che la variabile “Rurale” ha sul topic “Tipologia e caratteristiche dell’alloggio”;
- Il topic “Personalizzazione del titolo”. La moderazione è negativa, il che determina una riduzione degli effetti positivi;
- Il topic “Posizione”. La moderazione è negativa, il che determina una riduzione degli effetti positivi.

Il possesso del badge superhost potrebbe essere considerato come un segnale credibile di qualità, in quanto, per poterlo ottenere, bisognerebbe soddisfare diversi criteri in termini di performance, che comprendono: il superamento di un determinato numero di soggiorni offerti, il rispetto di soglie in merito al tasso di risposta, al tasso di cancellazione e alla valutazione complessiva ([Airbnb, n.d.](#)). Le informazioni trasmesse dall’host che possiede il badge “Superhost”, di conseguenza, potrebbero essere non più percepite come “cheap talk” dai clienti, ma come informazioni credibili. La variabile associata al possesso del badge superhost, di fatto, determina un incremento degli effetti positivi del topic “Comfort” sulla variabile dipendente ADR ed una riduzione degli effetti negativi associati al topic “Zone rurali”, sia rispetto alla variabile dipendente ADR sia rispetto alla variabile dipendente revPAN. La riduzione degli effetti positivi del topic “Personalizzazione del titolo” sulla variabile revPAN e la riduzione degli effetti positivi del topic “Tipologia e caratteristiche dell’alloggio” sulle variabili ADR e revPAN, causate dalla moderazione della variabile Airbnbsuperhost, potrebbero essere spiegate, in quanto, in presenza di un segnale di qualità come il possesso del badge superhost, informazioni come la tipologia dell’alloggio, le caratteristiche dell’alloggio e la personalizzazione, potrebbero risultare superflue agli occhi dei clienti.

Un ragionamento analogo potrebbe essere applicato alla moderazione della variabile “mediana_rating”, in quanto, il possesso di un rating maggiore o uguale a 4,8 stelle, potrebbe essere considerato come un segnale di qualità. Informazioni circa le caratteristiche e la tipologia di alloggio, la posizione, la ruralità e la personalizzazione del titolo, potrebbero risultare superflue agli occhi degli utenti, spiegando il perché della moderazione negativa rispetto alla variabile revPAN. Un rating elevato, inoltre, determina una riduzione degli effetti negativi della variabile “Zone rurali” rispetto alla variabile ADR. Le variabili rappresentanti la qualità dell’host, ovvero il possesso del badge superhost ed il possesso di un rating elevato, sembrano poter incrementare la credibilità, percepita dagli utenti, delle informazioni fornite nell’annuncio. Tali segnali di qualità, potrebbero, dunque, rendere superflue alcune informazioni presenti nel titolo dell’annuncio, determinando una valutazione differente delle informazioni da parte dei clienti. Le variabili di moderazione riguardanti la qualità dell’host, di conseguenza, sembrano poter esercitare un effetto di filtro o di ridefinizione dell’importanza di alcune informazioni presenti nel titolo di un annuncio.

Alla luce dei risultati ottenuti, si ritiene opportuno inserire nei titoli degli annunci di Airbnb informazioni di impatto che risultino essere informative e capaci di catturare l’attenzione dei clienti, evitando però l’introduzione di elementi ridondanti che potrebbero essere inglobati in altre sezioni dell’annuncio. In termini di studio delle variabili di moderazione, si evince che, per la definizione del titolo di un annuncio, è consigliabile applicare un approccio olistico, ovvero, un approccio tale per cui si tengano in considerazione tutti gli elementi che caratterizzano l’annuncio stesso.

Capitolo 5 – Impatti pratici, limitazioni e ricerca futura

I risultati ottenuti dallo svolgimento di tale tesi possono essere molto utili da un punto di vista pratico, soprattutto nell'ambito del place marketing. Diversi sono stati gli studi in relazione alla composizione dell'annuncio, come [Canziani e Nemati \(2021\)](#) che si soffermano sullo studio degli spunti tonali degli annunci, [Liang, Zhang e Li \(2020\)](#) che attestano la correlazione positiva tra il dettaglio delle descrizioni di un annuncio e l'informatività delle rispettive recensioni, [Xie e Mao \(2017\)](#) che si soffermano sul legame tra le caratteristiche di qualità dell'host e le prestazioni del relativo annuncio, [Kirkos \(2022\)](#) che determina come elementi principali dell'annuncio, la descrizione dell'host e le sue caratteristiche come il possesso del badge superhost ed il tasso di risposta, caratterizzandoli come le principali determinanti dell'annuncio che ne influenzano le performance, o ancora come [Chung e Sarnikar \(2022\)](#) che si soffermano sullo studio dei legami tra le descrizioni degli annunci e le rispettive performance. I risultati della tesi qui proposta, si offrono come elemento di congiunzione con gli studi passati, individuando legami di correlazione tra i principali argomenti individuati nei titoli degli annunci di Airbnb presenti nel campione del dataset elaborato ([Raguseo, 2025](#)) e le variabili prestazionali ADR e revPAN. In particolare, potrebbe essere utile in ambito pratico, per i marketer o chiunque debba definire il titolo di un annuncio, conoscere i legami che coniugano le performance di un annuncio con i rispettivi argomenti trattati nel titolo. In tal senso, alla luce dei risultati ottenuti, per migliorare i ricavi medi per giorni prenotati, risulta consigliabile inserire nel titolo di un annuncio Airbnb informazioni in merito alla tipologia e le caratteristiche dell'alloggio, il posizionamento dell'alloggio ed i comfort ad esso associati e ridurre informazioni circa la ruralità dell'alloggio. Al fine di migliorare i ricavi medi per giorni disponibili, invece, risulta consigliabile inserire nel titolo di un annuncio informazioni circa i comfort offerti dall'alloggio e ridurre informazioni circa la ruralità dell'alloggio. L'annuncio, inoltre, dovrebbe essere considerato come un organismo i cui organi agiscono in modo sinergico, dunque, è essenziale considerare cosa inserire e cosa no all'interno del titolo dell'annuncio in relazione a quanto già definito nelle altre sue componenti ed in base alle caratteristiche dell'inserzione stessa. Se si possedesse, ad esempio, un rating elevato (inteso come una valutazione maggiore o uguale a 4,8 stelle) effetti negativi derivanti dall'inserimento di informazioni circa la

ruralità diminuirebbero in relazione ai ricavi medi per giorni prenotati, pertanto, se fosse possibile, sarebbe meglio evitare di inserire tali elementi, se invece risultasse necessario il loro inserimento, sarebbe meglio farlo in relazione ad annunci dotati di un rating elevato in modo tale da ammortizzare gli effetti negativi. In modo analogo, il possesso del badge “Superhost” potrebbe essere considerato come un segnale credibile di qualità di un host, in quanto il suo possesso richiede dei costi in termini di: superamento di un determinato numero di soggiorni offerti, il rispetto di soglie in merito al tasso di risposta, al tasso di cancellazione e alla valutazione complessiva ([Airbnb, n.d.](#)). Essere in possesso del badge superhost, di fatti, incrementa l’effetto positivo del topic “Comfort” sui ricavi medi per giorni prenotati; di conseguenza, i possessori di tale badge, potrebbero sfruttare la moderazione positiva per incrementare l’effetto positivo sulle prestazioni dell’annuncio del topic riguardante i comfort. Il possesso del badge superhost, inoltre, riduce gli effetti negativi associati al topic “Zone rurali” sia in termini di ricavi medi per giorni prenotati sia in termini di ricavi medi per giorni disponibili. La localizzazione dell’alloggio offerto nell’inserzione, infine, risulta un punto cruciale da tenere in considerazione durante la redazione del titolo di un annuncio, in quanto, dalle analisi svolte, risulta moderare differenti topic sia in relazione alla variabile prestazionale ADR sia in relazione alla variabile prestazionale revPAN.

Di seguito, le tabelle riassuntive dei risultati ottenuti che potrebbero essere utilizzate come strumento per la definizione dei titoli.

Legenda:

- L’asterisco rappresenta un cambiamento nel segno della correlazione tra il topic e la variabile dipendente;
-  Correlazione positiva e statisticamente significativa;
-  Correlazione positiva ma non statisticamente significativa;
-  Statisticamente significativo nel modello con moderazioni;
-  Correlazione negativa e statisticamente significativa;
-  Correlazione negativa ma non statisticamente significativa;
-  Non statisticamente significativo nel modello con moderazioni;
-  Incremento degli effetti negativi;
-  Riduzione degli effetti positivi;

-  Incremento degli effetti positivi;
-  Riduzione degli effetti negativi.

Tabella 43 Tabella riassuntiva delle correlazioni in tabella 12 e moderazioni con variabile dipendente ADR

Topic	Segno	Moderazione Superhost	Moderazione rating $\geq 4,8$ stelle	Moderazione Rurale	Moderazione Urbana
Tipologia e caratteristiche dell'alloggio		 			
Posizione					 
Comfort		 			 
Personalizzazione del titolo				 	 
Zone rurali		 	 		 

Fonte: Elaborazione dei dati del campione del dataset pulito, dataset fornito dalla relatrice Elisabetta Raguseo (Raguseo, 2025).

Tabella 44 Tabella riassuntiva delle correlazioni in tabella 32 e moderazioni con variabile dipendente revPAN

Topic	Segno	Moderazione Superhost	Moderazione rating $\geq 4,8$ stelle	Moderazione Rurale	Moderazione Urbana
Tipologia e caratteristiche dell'alloggio		 	 	*  	
Posizione			 	 	
Comfort			*  		 
Personalizzazione del titolo		 	 	 	
Zone rurali		 	*  		 

Fonte: Elaborazione dei dati del campione del dataset pulito, dataset fornito dalla relatrice Elisabetta Raguseo (Raguseo, 2025).

Il lavoro di tesi redatto presenta, però, differenti limitazioni: innanzitutto, il dataset fa riferimento solo ai dati di Airbnb riscontrati in Italia in un arco temporale compreso tra il 2019 ed il 2024, studi futuri potrebbero far riferimento a dataset con un arco temporale più ampio o facenti riferimento a paesi differenti. Lo studio, inoltre, si è concentrato su due indicatori di performance, ovvero ADR e revPAN, studi futuri potrebbero concentrarsi sull'analisi di differenti indicatori di performance come, ad esempio, il tasso di occupazione. I topic, inoltre, sono stati individuati mediante un algoritmo generato dall'intelligenza artificiale ([OpenAI, 2025](#)) e facente riferimento a modelli combinati di K-Means ed SBERT, ricerche future potrebbero utilizzare metodologie differenti per l'individuazione degli argomenti principali nei titoli degli annunci. In più, l'individuazione dei topic è avvenuta tramite l'utilizzo di un campione del dataset pulito, studi futuri potrebbero cercare di elaborare i dati di un'intera popolazione (ovvero di un intero dataset).

Capitolo 6 - Conclusioni

La prospettiva di crescita di Airbnb anche dopo che, secondo [Vinogradov, Leick e Kivedal \(2020\)](#), la domanda sarà uguale all'offerta e le numerose ricerche presenti in merito a tale piattaforma di economia collaborativa ([Andreu et al., 2020](#)), hanno reso interessante lo studio delle caratteristiche semantiche dei titoli degli annunci di Airbnb, poste in relazione al ricavo medio per giorni prenotati (ADR) e al ricavo medio per giorni disponibili (revPAN). È stato possibile realizzare lo studio grazie all'utilizzo di un algoritmo che ha combinato modelli di K-Means ed SBERT realizzato dall'intelligenza artificiale ([OpenAI, 2025](#)); tramite tale algoritmo sono stati individuati, nel campione rappresentativo del dataset, sei argomenti principali presenti nei titoli degli annunci Airbnb in Italia, in un arco temporale compreso tra il 2019 ed il 2024. Da una preliminare analisi del sentiment dei titoli, come ci si aspetterebbe, il tono è per lo più positivo in accordo con quanto riscontrato da [Canziani e Nemati \(2021\)](#). I topic individuati, in ordine decrescente di presenza all'interno dei titoli nel campione del dataset pulito, sono: "Posizione", "Personalizzazione del titolo", "Zone rurali", "Tipicità abitativa", "Tipologia e caratteristiche dell'alloggio" e "Comfort". In merito all'argomento "Posizione", diversi sono stati gli studi che hanno individuato il luogo come elemento rilevante; [Santos et al. \(2022\)](#) definiscono il luogo come uno dei temi maggiormente trattati nelle recensioni di Airbnb, in accordo con [Demirdelen et al. \(2020\)](#) che lo considerano come uno dei topic maggiormente discussi nell'analisi delle recensioni positive, [Sthapit, Björk, Coudounaris e Stone \(2022\)](#) definiscono la località come una delle tematiche che suscitano emozioni memorabili, mentre [Cavique et al. \(2022\)](#) considerano il luogo come uno degli argomenti core che influenzano la soddisfazione del cliente, in accordo anche con [Klegr \(2024\)](#) che definisce il quartiere come uno degli elementi distintivi di Airbnb. Secondo [Amat-Lefort, Barravecchia e Mastrogiacomo \(2023\)](#), il luogo è una delle determinanti di qualità e, secondo [Qiu et al. \(2023\)](#), è uno dei fattori che influenzano positivamente l'intimità (componente fondamentale delle recensioni Airbnb). Alla luce, dunque, di tutti i riscontri precedentemente elencati, il luogo viene spesso citato all'interno degli studi relativi alla piattaforma di economia collaborativa Airbnb. [Li, Chen e Huang \(2020\)](#) hanno studiato le tematiche che suscitano emozioni nei clienti Airbnb, tra le diverse tematiche spicca quella relativa all'atmosfera,

che risulta suscitare emozioni positive; nel presente studio, in merito all'atmosfera, è stato individuato il topic "Tipicità abitativa" che, sebbene risulti rilevante come in accordo con [Li, Chen e Huang \(2020\)](#), nel presente lavoro di tesi, il topic non è stato inserito né nel modello in tabella 12 ([Raguseo, 2025](#)) né nel modello in tabella 32 ([Raguseo, 2025](#)). Nell'analisi delle singole correlazioni, il topic "Tipicità abitativa" è risultato essere statisticamente significativo al 5% e correlato negativamente alla variabile ADR e negativamente correlato alla variabile revPAN ma in modo non statisticamente significativo. In merito alle caratteristiche, argomento trattato nel topic "Tipologia e caratteristiche dell'alloggio", diversi sono stati gli studi che le definiscono come elemento incisivo nelle analisi circa la piattaforma Airbnb: [Santos et al. \(2022\)](#) individuano le caratteristiche come argomento altamente presente nelle recensioni, in modo analogo [Demirdelen et al. \(2020\)](#) lo riscontrano nell'analisi delle recensioni positive, [Sthapit, Björk, Coudounaris e Stone \(2022\)](#) definiscono le caratteristiche come tematiche che suscitano emozioni memorabili, [Cavique et al. \(2022\)](#) le definiscono come argomenti core che influenzano la soddisfazione del cliente, [Klegr \(2024\)](#) considera le caratteristiche fisiche dell'appartamento come caratteristiche distintive degli hotel che però rappresentano un punto di debolezza per Airbnb, secondo [Amat-Lefort, Barravecchia e Mastrogiacomo \(2023\)](#), invece, sono una determinante di qualità. Sebbene l'argomento "Tipologia e caratteristiche dell'alloggio" risulti essere altamente discusso all'interno della letteratura precedente, soprattutto in merito ai riferimenti delle caratteristiche dell'alloggio offerto, nell'analisi realizzata in questa tesi, tale argomento risulta essere debolmente significativo; di fatti, tale topic, in tabella 12 ([Raguseo, 2025](#)) è statisticamente significativo al 10% mentre in tabella 32 ([Raguseo, 2025](#)) non è statisticamente significativo. Le caratteristiche e la tipologia di alloggio, di conseguenza, influenzano debolmente i ricavi medi per giorni prenotati ma non hanno un impatto statisticamente significativo sui ricavi medi per giorni disponibili totali. L'argomento "Posizione", oltre ad essere altamente discusso nella letteratura precedente in merito ai riferimenti al luogo in cui è situato l'alloggio, risulta essere statisticamente significativo nel modello completo in tabella 12 ([Raguseo, 2025](#)) e positivamente correlato con la variabile dipendente ADR, mentre, nel modello in tabella 32 ([Raguseo, 2025](#)), non è statisticamente significativo sebbene sia positivamente correlato con la variabile dipendente revPAN. Il topic "Personalizzazione del titolo" risulta essere positivamente correlato con la variabile prestazionale intesa come il rapporto tra i ricavi espressi in

dollari ed il numero di giorni prenotati, ma tale correlazione non è statisticamente significativa nel modello presente in tabella 12 ([Raguseo, 2025](#)), tali risultati si verificano in modo analogo nel modello in tabella 32 ([Raguseo, 2025](#)). L'argomento "Zone rurali", in relazione al modello in tabella 12 ([Raguseo, 2025](#)), risulta essere statisticamente significativo ma negativamente correlato con la variabile dipendente ADR, tali risultati si verificano in modo analogo nel modello in tabella 32 ([Raguseo, 2025](#)) in relazione alla variabile prestazionale revPAN. Il topic "Comfort", infine, è positivamente correlato con le variabili dipendenti ADR e revPAN, tali correlazioni sono statisticamente significative sia nel modello presente in tabella 12 ([Raguseo, 2025](#)) sia nel modello presente in tabella 32 ([Raguseo, 2025](#)).

Dal confronto tra gli argomenti maggiormente discussi all'interno del campione del dataset e gli argomenti risultati statisticamente significativi, emergono delle discrepanze: il topic "Comfort" che risulta essere l'argomento meno discusso all'interno del campione rappresentativo del dataset, è invece statisticamente significativo all'1% nel modello in tabella 12 ([Raguseo, 2025](#)) ed al 5% nel modello in tabella 32 ([Raguseo, 2025](#)), il topic "Personalizzazione del titolo" che in termini di discussione risulta essere il secondo argomento maggiormente discusso nel campione del dataset pulito, sia nel modello in tabella 12 ([Raguseo, 2025](#)) sia nel modello in tabella 32 ([Raguseo, 2025](#)), non risulta essere statisticamente significativo. Si evince, dunque, che la presenza in termini di topic nel campione rappresentativo del dataset non coincide con l'ordine dell'impatto che questi hanno sulle variabili dipendenti, determinando di fatti la necessità di un cambiamento nella stesura dei titoli degli annunci. La differenza tra gli argomenti maggiormente discussi e quelli che avrebbero maggiore impatto è stata riscontrata anche da [Chung e Sarnikar \(2022\)](#) nel rispettivo studio sulle descrizioni degli annunci di Airbnb.

In termini di moderazione, si riscontra che la variabile associata al possesso del badge superhost determina un incremento degli effetti positivi del topic "Comfort" sulla variabile dipendente ADR ed una riduzione degli effetti negativi associati al topic "Zone rurali", sia rispetto alla variabile dipendente ADR sia rispetto alla variabile dipendente revPAN. La moderazione della variabile "Airbnb Superhost" determina anche la riduzione degli effetti positivi del topic "Personalizzazione del titolo" sulla variabile revPAN e la riduzione degli effetti positivi del topic "Tipologia e caratteristiche dell'alloggio" sulle variabili ADR e revPAN, in quanto, probabilmente, in presenza di un

segnale di qualità come il possesso del badge superhost, informazioni come la tipologia dell'alloggio, le caratteristiche dell'alloggio e la personalizzazione potrebbero risultare superflue agli occhi dei clienti. I risultati evidenziano, dunque, l'impatto delle caratteristiche di qualità dell'host sulle variabili prestazionali, in accordo con quanto riscontrato da [Xie e Mao \(2017\)](#), i quali scoprono che gli host che possiedono qualità migliori riscontrano anche migliori prestazioni dell'annuncio, ed in accordo con [Kirkos \(2022\)](#), il quale individua il badge superhost come una tra le determinanti principali che influenzano le performance dell'annuncio. Un ulteriore indicatore di qualità, oltre al possesso del badge superhost, potrebbe essere il possesso di un rating elevato associato all'annuncio, dove per rating elevato si intende, in questo studio, una valutazione maggiore o uguale a 4,8 stelle. In termini di moderazione della variabile "mediana_rating", le informazioni circa le caratteristiche e la tipologia di alloggio, la posizione, i riferimenti alla ruralità e la personalizzazione del titolo, possiedono una moderazione negativa che ne determina una riduzione degli effetti positivi rispetto alla variabile revPAN e le informazioni circa i comfort possiedono una moderazione positiva rispetto alla variabile revPAN che ne determina un incremento degli effetti negativi, probabilmente ciò accade in quanto tali informazioni potrebbero essere considerate indicazioni superflue agli occhi degli utenti. Un rating elevato, inoltre, determina una riduzione degli effetti negativi della variabile "Zone rurali" rispetto alla variabile ADR. Il valore della qualità sugli atteggiamenti è stato ritenuto da [So, Kim e Min \(2022\)](#) determinante per i comportamenti di acquisto futuri dei clienti, di fatti, se si intende la qualità come il possesso del badge superhost ed il possesso di un rating maggiore o uguale a 4,8 stelle, risulta evidente l'impatto di tali indicatori di qualità sulle variabili prestazionali. Le variabili "Urbana" e "Rurale" ricoprono, anch'esse, un ruolo importante in termini di moderazione degli effetti dei topic sulle variabili prestazionali, sottolineando l'effetto significativo che il contesto spaziale possiede in termini di performance degli annunci.

In conclusione, per migliorare le prestazioni di un annuncio, risulta, dallo studio presente in questa tesi, consigliabile inserire nel titolo di un annuncio sulla piattaforma Airbnb, argomenti che risultino impattanti dal punto di vista del cliente. Inserire, ad esempio, informazioni circa la tipologia e le caratteristiche dell'alloggio, l'ubicazione dell'alloggio ed i comfort presenti, ha un impatto positivo sui ricavi medi per giorni prenotati. L'omissione di informazioni circa la ruralità dell'alloggio, riduce gli effetti negativi del

topic “Zone rurali” sia sui ricavi medi per giorni prenotati sia sui ricavi medi per giorni disponibili. Inserire, infine, informazioni nel titolo circa i comfort, oltre che ad avere un impatto significativo su ADR, ha un impatto positivo e statisticamente significativo su revPAN. L’effetto significativo delle variabili di moderazione suggerisce, inoltre, che la definizione di un annuncio dovrebbe avvenire in modo olistico; ovvero, nel definire il titolo, bisognerebbe considerare anche quanto presentato nelle altre parti dell’annuncio come se ogni parte lavorasse in modo sinergico con le altre.

Riferimenti

- Airbnb, Inc. (2020). *Annual report* [Annual report]. https://www.annualreports.com/HostedData/AnnualReportArchive/a/NASDAQ_ABNB_2020.pdf
- Airbnb, Inc. (2021). *Annual report* [Annual report]. https://www.annualreports.com/HostedData/AnnualReportArchive/a/NASDAQ_ABNB_2021.pdf
- Airbnb, Inc. (2022). *Annual report* [Annual report]. https://www.annualreports.com/HostedData/AnnualReportArchive/a/NASDAQ_ABNB_2022.pdf
- Airbnb, Inc. (2023). *Annual report* [Annual report]. https://www.annualreports.com/HostedData/AnnualReports/PDF/NASDAQ_ABNB_2023.pdf
- Airbnb. (n.d.). *Informazioni su Airbnb: cos'è e come funziona*. Airbnb. Recuperato il 29 marzo 2025, da <https://www.airbnb.it/help/article/2503>
- Airbnb. (n.d.). *La notte più prenotata di sempre*. Airbnb News. Recuperato il 2 giugno 2025, da <https://news.airbnb.com/it/la-notte-piu-prenotata-di-sempre/>
- Airbnb. (n.d.). *Requisiti per diventare Superhost*. <https://www.airbnb.it/help/article/829>
- Airbnb. (n.d.). *Termini di cancellazione per il tuo annuncio*. Airbnb. <https://www.airbnb.it/help/article/475#section-heading-0>
- Ajzen, I. (1991). The theory of planned behavior. *Organizational Behavior and Human Decision Processes*, 50(2), 179–211. [https://doi.org/10.1016/0749-5978\(91\)90020-T](https://doi.org/10.1016/0749-5978(91)90020-T)
- Amat-Lefort, N., Barravecchia, F., & Mastrogiacomo, L. (2023). Quality 4.0: Big data analytics to explore service quality attributes and their relation to user sentiment in Airbnb reviews. *International Journal of Quality & Reliability Management*, 40(4), 990–1008. <https://doi.org/10.1108/IJQRM-01-2022-0024>
- Andreu, L., Bigne, E., Amaro, S. & Palomo, J. (2020). Airbnb research: An analysis in tourism and hospitality journals. *International Journal of Culture, Tourism and Hospitality Research*, 14(1), 2–20. <https://doi.org/10.1108/IJCTHR-06-2019-0113>

- Barravecchia, F. (2024). *Come analizzare la digital VoC: Analisi dei risultati* [Presentazione non pubblicata]. Corso “Qualità nei servizi”, Politecnico di Torino.
- Barravecchia, F. (2024). *Come analizzare la digital VoC: Topic Modelling – Applicazione STM* [Presentazione non pubblicata]. Corso “Qualità nei servizi”, Politecnico di Torino.
- Barravecchia, F. (2024). *Come analizzare la digital VoC: Topic Modelling – Selezione del numero di Topic da estrarre* [Presentazione non pubblicata]. Corso “Qualità nei servizi”, Politecnico di Torino.
- Barravecchia, F. (2024). *Come analizzare la digital VoC: Topic Modelling – Labelling* [Presentazione non pubblicata]. Corso “Qualità nei servizi”, Politecnico di Torino.
- Canziani, B., & Nemati, H. R. (2021). Core and supplemental elements of hospitality in the sharing economy: Insights from semantic and tonal cues in Airbnb property listings. *Tourism Management*, 87, 104377. <https://doi.org/10.1016/j.tourman.2021.104377>
- Cavique, M., Ribeiro, R., Batista, F., & Correia, A. (2022). Examining Airbnb guest satisfaction tendencies: a text mining approach. *Current Issues in Tourism*, 25(22), 3607–3622. <https://doi.org/10.1080/13683500.2022.2115877>
- Chung, Y., & Sarnikar, S. (2022). Understanding host marketing strategies on Airbnb and their impact on listing performance: A text analytics approach. *Information Technology & People*, 35(7), 2075–2097. <https://doi.org/10.1108/ITP-10-2020-0718>
- Cochran, W. G. (1977). *Sampling Techniques* (3rd ed.). Wiley.
- Commissione Europea. (2016). *Un'agenda europea per l'economia collaborativa* (COM/2016/356). EUR-Lex. <https://eur-lex.europa.eu/legal-content/IT/TXT/PDF/?uri=CELEX%3A52016DC0356>
- CompaniesMarketCap.com. (2025, febbraio 25). *Market capitalization of Airbnb from 2020 to 2024 (in billion U.S. dollars)* [Data set]. In Statista. Recuperato il 24 marzo 2025, da <https://www.statista.com/statistics/339845/company-value-and-equity-funding-of-airbnb/>
- Demirdelen, D., Dinçer, M. Z., & İstanbullu Dinçer, F. (2020). Airbnb aracılığıyla konaklama yapan misafirlerin memnuniyet düzeylerini etkileyen faktörlerin belirlenmesine yönelik netnografik bir araştırma. *International Journal of*

Contemporary Tourism Research, 4(1), 1–14.
<https://doi.org/10.30625/ijctr.623218>

- Dogru, T., Majid, K., Laroche, M., Mody, M., & Suess, C. (2021). Communicating quality while evoking loss – How consumers assess extra charges in the Airbnb marketplace. *Tourism Management*, 87, 104376. <https://doi.org/10.1016/j.tourman.2021.104376>
- Edizioni Simone. (n.d.). *Adverse selection*. In *Dizionari Simone*. <https://dizionari.simone.it/6/adverse-selection>
- Eurostat. (2024). *Guest nights spent at short-stay accommodation offered via collaborative economy platforms by mode of accommodation - experimental statistics (tour_ce_oam)* [Data set]. Eurostat. Recuperato il 24 marzo 2025, da [https://ec.europa.eu/eurostat/databrowser/view/tour_ce_oam\\$defaultview/default/table](https://ec.europa.eu/eurostat/databrowser/view/tour_ce_oam$defaultview/default/table)
- Eurostat. (2024). *Guest nights spent at short-stay accommodation offered via collaborative economy platforms by NUTS 3 region - experimental statistics (tour_ce_oan3)* [Data set]. Eurostat. Recuperato il 24 marzo 2025, da [https://ec.europa.eu/eurostat/databrowser/view/tour_ce_oan3\\$defaultview/default/table](https://ec.europa.eu/eurostat/databrowser/view/tour_ce_oan3$defaultview/default/table)
- Fernando, J. (2024, agosto 17). *Free cash flow (FCF): Formula to Calculate and Interpret It*. Investopedia. Reviewed by D. Kindness. Fact checked by V. Velasquez. Recuperato il 29 marzo 2025, da <https://www.investopedia.com/terms/f/freecashflow.asp>
- Garay-Tamajón, L. A., & Morales-Pérez, S. (2023). ‘Belong anywhere’: Focusing on authenticity and the role of Airbnb in the projected destination image. *International Journal of Tourism Research*, 25(1), 63–78. <https://doi.org/10.1002/jtr.2551>
- Giovanardi, M. (2012). Haft and sord factors in place branding: Between functionalism and representationalism. *Place Branding and Public Diplomacy*, 8(1), 30–45. <https://doi.org/10.1057/pb.2012.1>
- Godovykh, M., & Baker, C. (2022). Hospitality art: A research note on the main factors affecting attractiveness of visual art in Airbnb settings. *Journal of Hospitality and Tourism Management*, 51, 401–405. <https://doi.org/10.1016/j.jhtm.2022.04.014>

- Gurran, N., Zhang, Y., & Shrestha, P. (2020). 'Pop-up' tourism or 'invasion'? Airbnb in coastal Australia. *Annals of Tourism Research*, 81, 102845. <https://doi.org/10.1016/j.annals.2019.102845>
- IBISWorld. (2024, maggio 15). *Market size of the tourism sector worldwide from 2011 to 2023, with a forecast for 2024 (in trillion U.S. dollars)* [Data set]. In Statista. Recuperato il 24 marzo 2025, da <https://www.statista.com/statistics/1220218/tourism-industry-market-size-global/>
- IBM. (2024, luglio 15). *EMEA region*. In *IBM Cloud Docs*. Recuperato il 29 marzo 2025, da <https://cloud.ibm.com/docs/CDN?topic=CDN-emea-region&locale=it>
- Inside Airbnb. (2024, settembre 29). *Number of Airbnb listings in selected cities in Italy as of September 2024* [Data set]. In Statista. Recuperato il 24 marzo 2025, da <https://www.statista.com/statistics/1084927/number-of-airbnb-listings-in-selected-italian-cities/>
- ISTAT. (2024). *Superfici territoriali per comune* [Dataset]. Istituto Nazionale di Statistica. https://esploradati.istat.it/databrowser/#/it/dw/categories/IT1,Z0930TER,1.0/DCV_CARGEOMOR_ST_COM/IT1,DCCV_CARGEOMOR_ST_COM,1.0
- ISTAT. (2025). *Popolazione residente al 1° gennaio – Abruzzo* [Dataset]. Istituto Nazionale di Statistica. https://esploradati.istat.it/databrowser/#/it/dw/categories/IT1,POP,1.0/POP_POPULATION/DCIS_POPRES1/DCIS_POPRES1_REG_MUN/IT1,22_289_DF_DCIS_POPRES1_16,1.0
- ISTAT. (2025). *Popolazione residente al 1° gennaio – Basilicata* [Dataset]. Istituto Nazionale di Statistica. https://esploradati.istat.it/databrowser/#/it/dw/categories/IT1,POP,1.0/POP_POPULATION/DCIS_POPRES1/DCIS_POPRES1_REG_MUN/IT1,22_289_DF_DCIS_POPRES1_20,1.0
- ISTAT. (2025). *Popolazione residente al 1° gennaio – Calabria* [Dataset]. Istituto Nazionale di Statistica. https://esploradati.istat.it/databrowser/#/it/dw/categories/IT1,POP,1.0/POP_POPULATION/DCIS_POPRES1/DCIS_POPRES1_REG_MUN/IT1,22_289_DF_DCIS_POPRES1_21,1.0

[ULATION/DCIS_POPRES1/DCIS_POPRES1_REG_MUN/IT1,22_289_DF_D_CIS_POPRES1_21,1.0](https://esploradati.istat.it/databrowser/#/it/dw/categories/IT1,POP,1.0/POP_POPULATION/DCIS_POPRES1/DCIS_POPRES1_REG_MUN/IT1,22_289_DF_D_CIS_POPRES1_21,1.0)

- ISTAT. (2025). *Popolazione residente al 1° gennaio – Campania* [Dataset]. Istituto Nazionale di Statistica. https://esploradati.istat.it/databrowser/#/it/dw/categories/IT1,POP,1.0/POP_POPULATION/DCIS_POPRES1/DCIS_POPRES1_REG_MUN/IT1,22_289_DF_D_CIS_POPRES1_18,1.0
- ISTAT. (2025). *Popolazione residente al 1° gennaio – Emilia Romagna* [Dataset]. Istituto Nazionale di Statistica. https://esploradati.istat.it/databrowser/#/it/dw/categories/IT1,POP,1.0/POP_POPULATION/DCIS_POPRES1/DCIS_POPRES1_REG_MUN/IT1,22_289_DF_D_CIS_POPRES1_11,1.0
- ISTAT. (2025). *Popolazione residente al 1° gennaio – Friuli Venezia Giulia* [Dataset]. Istituto Nazionale di Statistica. https://esploradati.istat.it/databrowser/#/it/dw/categories/IT1,POP,1.0/POP_POPULATION/DCIS_POPRES1/DCIS_POPRES1_REG_MUN/IT1,22_289_DF_D_CIS_POPRES1_10,1.0
- ISTAT. (2025). *Popolazione residente al 1° gennaio – Lazio* [Dataset]. Istituto Nazionale di Statistica. https://esploradati.istat.it/databrowser/#/it/dw/categories/IT1,POP,1.0/POP_POPULATION/DCIS_POPRES1/DCIS_POPRES1_REG_MUN/IT1,22_289_DF_D_CIS_POPRES1_15,1.0
- ISTAT. (2025). *Popolazione residente al 1° gennaio – Liguria* [Dataset]. Istituto Nazionale di Statistica. https://esploradati.istat.it/databrowser/#/it/dw/categories/IT1,POP,1.0/POP_POPULATION/DCIS_POPRES1/DCIS_POPRES1_REG_MUN/IT1,22_289_DF_D_CIS_POPRES1_5,1.0
- ISTAT. (2025). *Popolazione residente al 1° gennaio – Lombardia* [Dataset]. Istituto Nazionale di Statistica. https://esploradati.istat.it/databrowser/#/it/dw/categories/IT1,POP,1.0/POP_POPULATION/DCIS_POPRES1/DCIS_POPRES1_REG_MUN/IT1,22_289_DF_D_CIS_POPRES1_6,1.0

- ISTAT. (2025). *Popolazione residente al 1° gennaio – Marche* [Dataset]. Istituto Nazionale di Statistica. https://esploradati.istat.it/databrowser/#/it/dw/categories/IT1,POP,1.0/POP_POPULATION/DCIS_POPRES1/DCIS_POPRES1_REG_MUN/IT1,22_289_DF_DCIS_POPRES1_14,1.0
- ISTAT. (2025). *Popolazione residente al 1° gennaio – Molise* [Dataset]. Istituto Nazionale di Statistica. https://esploradati.istat.it/databrowser/#/it/dw/categories/IT1,POP,1.0/POP_POPULATION/DCIS_POPRES1/DCIS_POPRES1_REG_MUN/IT1,22_289_DF_DCIS_POPRES1_17,1.0
- ISTAT. (2025). *Popolazione residente al 1° gennaio – Provincia Autonoma Bolzano* [Dataset]. Istituto Nazionale di Statistica. https://esploradati.istat.it/databrowser/#/it/dw/categories/IT1,POP,1.0/POP_POPULATION/DCIS_POPRES1/DCIS_POPRES1_REG_MUN/IT1,22_289_DF_DCIS_POPRES1_7,1.0
- ISTAT. (2025). *Popolazione residente al 1° gennaio – Provincia Autonoma Trento* [Dataset]. Istituto Nazionale di Statistica. https://esploradati.istat.it/databrowser/#/it/dw/categories/IT1,POP,1.0/POP_POPULATION/DCIS_POPRES1/DCIS_POPRES1_REG_MUN/IT1,22_289_DF_DCIS_POPRES1_8,1.0
- ISTAT. (2025). *Popolazione residente al 1° gennaio – Puglia* [Dataset]. Istituto Nazionale di Statistica. https://esploradati.istat.it/databrowser/#/it/dw/categories/IT1,POP,1.0/POP_POPULATION/DCIS_POPRES1/DCIS_POPRES1_REG_MUN/IT1,22_289_DF_DCIS_POPRES1_19,1.0
- ISTAT. (2025). *Popolazione residente al 1° gennaio – Regione Piemonte* [Dataset]. Istituto Nazionale di Statistica. https://esploradati.istat.it/databrowser/#/it/dw/categories/IT1,POP,1.0/POP_POPULATION/DCIS_POPRES1/DCIS_POPRES1_REG_MUN/IT1,22_289_DF_DCIS_POPRES1_3B,1.0
- ISTAT. (2025). *Popolazione residente al 1° gennaio – Regione Valle d'Aosta* [Dataset]. Istituto Nazionale di Statistica. https://esploradati.istat.it/databrowser/#/it/dw/categories/IT1,POP,1.0/POP_POPULATION/DCIS_POPRES1/DCIS_POPRES1_REG_MUN/IT1,22_289_DF_DCIS_POPRES1_3B,1.0

[ULATION/DCIS_POPRES1/DCIS_POPRES1_REG_MUN/IT1,22_289_DF_DCIS_POPRES1_4,1.0](https://esploradati.istat.it/databrowser/#/it/dw/categories/IT1,POP,1.0/POP_POPULATION/DCIS_POPRES1/DCIS_POPRES1_REG_MUN/IT1,22_289_DF_DCIS_POPRES1_4,1.0)

- ISTAT. (2025). *Popolazione residente al 1° gennaio – Sardegna* [Dataset]. Istituto Nazionale di Statistica. https://esploradati.istat.it/databrowser/#/it/dw/categories/IT1,POP,1.0/POP_POPULATION/DCIS_POPRES1/DCIS_POPRES1_REG_MUN/IT1,22_289_DF_DCIS_POPRES1_23,1.0
- ISTAT. (2025). *Popolazione residente al 1° gennaio – Sicilia* [Dataset]. Istituto Nazionale di Statistica. https://esploradati.istat.it/databrowser/#/it/dw/categories/IT1,POP,1.0/POP_POPULATION/DCIS_POPRES1/DCIS_POPRES1_REG_MUN/IT1,22_289_DF_DCIS_POPRES1_22,1.0
- ISTAT. (2025). *Popolazione residente al 1° gennaio – Toscana* [Dataset]. Istituto Nazionale di Statistica. https://esploradati.istat.it/databrowser/#/it/dw/categories/IT1,POP,1.0/POP_POPULATION/DCIS_POPRES1/DCIS_POPRES1_REG_MUN/IT1,22_289_DF_DCIS_POPRES1_12,1.0
- ISTAT. (2025). *Popolazione residente al 1° gennaio – Umbria* [Dataset]. Istituto Nazionale di Statistica. https://esploradati.istat.it/databrowser/#/it/dw/categories/IT1,POP,1.0/POP_POPULATION/DCIS_POPRES1/DCIS_POPRES1_REG_MUN/IT1,22_289_DF_DCIS_POPRES1_13,1.0
- ISTAT. (2025). *Popolazione residente al 1° gennaio – Veneto* [Dataset]. Istituto Nazionale di Statistica. https://esploradati.istat.it/databrowser/#/it/dw/categories/IT1,POP,1.0/POP_POPULATION/DCIS_POPRES1/DCIS_POPRES1_REG_MUN/IT1,22_289_DF_DCIS_POPRES1_9,1.0
- Istituto Nazionale di Statistica (ISTAT). (2025). *Italia, regioni, province* [Data set]. ISTAT EsploraDati. Recuperato il 29 marzo 2025, da https://esploradati.istat.it/databrowser/#/it/dw/categories/IT1,POP,1.0/POP_POPULATION/DCIS_POPRES1/IT1,22_289_DF_DCIS_POPRES1_1,1.0
- Istituto Nazionale di Statistica (ISTAT). (2025). *Movimento dei clienti negli esercizi ricettivi per tipo di esercizio - annuali* [Data set]. ISTAT EsploraDati.

- Recuperato il 24 marzo 2025, da https://esploradati.istat.it/databrowser/#/it/dw/categories/IT1,Z0700SER,1.0/SE_R_TOURISM/DCSC_TUR_OCCYEAR/IT1,122_54_DF_DCSC_TUR_7,1.0
- Istituto Nazionale di Statistica (ISTAT). (2025). *Vacanze* [Data set]. ISTAT EsploraDati. Recuperato il 24 marzo 2025, da https://esploradati.istat.it/databrowser/#/it/dw/categories/IT1,Z0830COM,1.0/COM_TRIPS/DCCV_VIAGGI_CAPI_2/IT1,68_1221_DF_DCCV_VIAGGI_CAPI_2_1,1.0
 - Kavlakoglu, E., & Winland, V. (2024, Giugno 26). *Cos'è il k-means clustering?* IBM. Recuperato il 10 giugno, 2025, da <https://www.ibm.com/it-it/think/topics/k-means-clustering>
 - Kirkos, E. (2022). Airbnb listings' performance: determinants and predictive models. *European Journal of Tourism Research*, 30, Article 3012. <https://doi.org/10.54055/ejtr.v30i.2142>
 - Klegr, T. (2024). Comparing Airbnb and traditional accommodation experiences using text-mining methods – the hedonic and utilitarian values framework. *European Journal of Tourism Research*, 38, 3804. <https://doi.org/10.54055/ejtr.v38i.3299>
 - Li, Z., Chen, H., & Huang, X. (2020). Airbnb or hotel? A comparative study on the sentiment of Airbnb guests in Sydney – Text analysis based on big data. *International Journal of Tourism and Hospitality Management in the Digital Age*, 4(2), 1–10. <https://doi.org/10.4018/IJTHMDA.2020070101>
 - Liang, S., Zhang, X., & Li, H. (2020). How to make guest reviews more informative? A case study of Airbnb. *e-Review of Tourism Research*, 17(4). <https://ertr-ojs-tamu.tdl.org/ertr/article/view/545>
 - OpenAI. (2025). *ChatGPT* (Versione GPT-4.5) [Modello linguistico]. <https://chat.openai.com>
 - Oxford University Press. (n.d.). *Market capitalization*. In Oxford Learner's Dictionaries. Recuperato il 29 marzo 2025, da <https://www.oxfordlearnersdictionaries.com/definition/english/market-capitalization?q=market+capitalization>
 - Oxford University Press. (n.d.). *Moral hazard*. In *Oxford Learner's Dictionaries*. <https://www.oxfordlearnersdictionaries.com/definition/english/moral-hazard>

- Oxford University Press. (n.d.). *Sharing economy*. In *Oxford Learner's Dictionaries*. Recuperato il 25 marzo 2025, da <https://www.oxfordlearnersdictionaries.com/definition/english/sharing-economy?q=sharing+economy>
- Qiu, W.-W., Yu, H.-T., Tsai, C.-H. (K.), Zhu, D. (J.), Chen, M.-H., & Kim, H. J. (2023). Understanding the value of host-guest intimacy behind online reviews of Airbnb. *International Journal of Hospitality Management*, 115, 103599. <https://doi.org/10.1016/j.ijhm.2023.103599>
- Raguseo, E. (2025). Dataset "IT_19-24.dta" fornito per la seguente tesi di laurea "Studio empirico sulle performance di Airbnb: analisi semantica applicata al caso Italiano" [Dataset non pubblicato].
- Reuters. (2025, marzo 12). *Market cap of selected leading online travel companies worldwide as of March 2025 (in billion U.S. dollars)* [Data set]. In Statista. Recuperato il 24 marzo 2025, da <https://www.statista.com/statistics/1039616/leading-online-travel-companies-by-market-cap/>
- Santos, A. I. G. P., Costa Perinotto, A. R., Rodrigues Soares, J. R., & Savi Mondo, T. (2022). Feeling at home while traveling: An analysis of the experiences of Airbnb users. *Tourism and Hospitality Management*, 28(1), 167–192. <https://doi.org/10.20867/thm.28.1.9>
- So, K. K. F., Kim, H., & Min, S. (2022). Creating customer value in the sharing economy: An investigation of Airbnb users and their tripographic characteristics. *International Journal of Contemporary Hospitality Management*, 34(1), 23–45. <https://doi.org/10.1108/IJCHM-12-2020-1500>
- Spence, M. (1981). Signaling, screening, and information. In S. Rosen (Ed.), *Studies in labor markets* (pp. 319–358). University of Chicago Press. <http://www.nber.org/chapters/c8915>
- Spence, M. (2002). *Signaling in retrospect and the informational structure of markets*. *The American Economic Review*, 92(3), 434–459. https://www.jstor.org/stable/pdf/3083350.pdf?casa_token=GOUnWLO0mqEAAAAA:jX6PIOlGnEEumVHbv-yrqUwmLj16aXw5h0nnA4BMYMDL7Fn9FRafjNfGAfulRDoybv1e57a7RR9Gni0vo9MGbJuUi_8uy9zl5adI2ZLfcG_mQAQbOL8

- Statista. (2024, luglio 1). *Online travel market size worldwide from 2017 to 2023, with a forecast until 2029 (in billion U.S. dollars)* [Data set]. Recuperato il 24 marzo 2025, da <https://www.statista.com/statistics/1179020/online-travel-agent-market-size-worldwide/>
- Statista. (2024, settembre 16). *Revenue share of sales channels of the travel and tourism market worldwide from 2019 to 2029* [Data set]. Recuperato il 24 marzo 2025, da <https://www.statista.com/forecasts/1239068/sales-channels-travel-tourism-worldwide>
- Sthapit, E., Björk, P., Coudounaris, D. N., & Stone, M. J. (2022). A new conceptual framework for memorable Airbnb experiences: guests' perspectives. *International Journal of Culture, Tourism and Hospitality Research*, 16(1), 75–86. <https://doi.org/10.1108/IJCTHR-01-2021-0002>
- Tajeddini, K., Gamage, T. C., Hameed, W. U., Qumsieh-Mussalam, G., Chaijani, M. H., Rasoolimanesh, S. M., & Kallmuenzer, A. (2022). How self-gratification and social values shape revisit intention and customer loyalty of Airbnb customers. *International Journal of Hospitality Management*, 100, 103093. <https://doi.org/10.1016/j.ijhm.2021.103093>
- U.S. Securities and Exchange Commission. (2025, febbraio 13). *Airbnb's gross booking value from 2019 to 2024, by region (in billion U.S. dollars)* [Data set]. In Statista. Recuperato il 24 marzo 2025, da <https://www.statista.com/statistics/1193554/airbnb-gross-booking-value-by-region-worldwide/>
- U.S. Securities and Exchange Commission. (2025, febbraio 13). *Average number of nights per Airbnb booking from 2019 to 2024, by region* [Data set]. In Statista. Recuperato il 24 marzo 2025, da <https://www.statista.com/statistics/1193597/airbnb-average-nights-per-booking-by-region-worldwide/>
- U.S. Securities and Exchange Commission. (2025, febbraio 13). *Costs and expenses of Airbnb from 2017 to 2024 (in billion U.S. dollars)* [Data set]. In Statista. Recuperato il 24 marzo 2025, da <https://www.statista.com/statistics/1193151/airbnb-expenses-worldwide/>
- U.S. Securities and Exchange Commission. (2025, febbraio 13). *Net income of Airbnb from 2017 to 2024 (in million U.S. dollars)* [Data set]. In Statista.

- Recuperato il 24 marzo 2025, da <https://www.statista.com/statistics/1193179/airbnb-net-loss-worldwide/>
- U.S. Securities and Exchange Commission. (2025, febbraio 13). *Number of nights and experiences booked on Airbnb from 2019 to 2024, by region (in millions)* [Data set]. In Statista. Recuperato il 24 marzo 2025, da <https://www.statista.com/statistics/1193543/airbnb-nights-experiences-by-region-worldwide/>
 - U.S. Securities and Exchange Commission. (2025, febbraio 13). *Revenue of Airbnb from 2019 to 2024, by region (in billion U.S. dollars)* [Data set]. In Statista. Recuperato il 24 marzo 2025, da <https://www.statista.com/statistics/1193565/airbnb-revenue-by-region-worldwide/>
 - Vinogradov, E., Leick, B., & Kivedal, B. K. (2020). An agent-based modelling approach to housing market regulations and Airbnb-induced tourism. *Tourism Management*, 77, 104004. <https://doi.org/10.1016/j.tourman.2019.104004>
 - World Tourism Organization (UN Tourism). (20 gennaio, 2025). *Number of international tourist arrivals worldwide from 1950 to 2024 (in millions)* [Data set]. In Statista. Recuperato il 24 marzo 2025, da <https://www.statista.com/statistics/209334/total-number-of-international-tourist-arrivals/>
 - Xie, K., & Mao, Z. (2017). The impacts of quality and quantity attributes of Airbnb hosts on listing performance. *International Journal of Contemporary Hospitality Management*, 29(9), 2240–2260. <https://doi.org/10.1108/IJCHM-07-2016-0345>
 - Yilmaz, O. (2022, Febbraio 20). *Introduction to Sentence-BERT (SBERT)*. Search Candy. Recuperato il 10 giugno, 2025, da <https://www.searchcandy.uk/nlp/sentence-bert/>
 - Zheng, C., & Zhang, J. (2023). Inspiring guests’ imagination of “home away from home” to choose Airbnb through brand storytelling. *International Journal of Contemporary Hospitality Management*, 35(6), 2136–2156. <https://doi.org/10.1108/IJCHM-04-2022-0444>