

POLITECNICO DI TORINO

Collegio di Ingegneria Gestionale – Classe LM-31

Corso di Laurea Magistrale in Ingegneria Gestionale

Percorso Gestione e progettazione dei servizi digitali



Tesi di Laurea Magistrale

***Comunicazione e performance sulle piattaforme digitali:
un'analisi semantica delle descrizioni
degli annunci su Airbnb***

Relatrice

Prof.ssa Elisabetta Raguseo

Candidata

Tamara Tamburello

Anno Accademico 2025/2026

ABSTRACT

Il presente lavoro analizza il ruolo della comunicazione testuale nella determinazione delle performance economiche degli annunci pubblicati sulla piattaforma Airbnb. L'obiettivo è valutare se e in che misura le scelte linguistiche adottate dagli host influenzino indicatori chiave di performance, quali ricavi, prezzo medio, tasso di occupazione, numero di prenotazioni e reputazione online.

L'analisi si basa su un campione di annunci Airbnb localizzati in Italia e attivi nel 2024. A partire da un dataset strutturale di grandi dimensioni, è stato estratto un sotto-campione sul quale sono state raccolte le descrizioni testuali mediante tecniche di web scraping. I testi sono stati successivamente analizzati attraverso strumenti di Natural Language Processing, al fine di costruire indicatori quantitativi relativi allo stile comunicativo, allo sforzo informativo e al contenuto semantico delle descrizioni, successivamente inclusi in modelli di regressione lineare con effetti fissi.

I risultati mostrano che la dimensione formale del testo, come lunghezza e complessità linguistica, non ha un impatto significativo sulle performance economiche. Al contrario, il contenuto e l'orientamento della comunicazione risultano rilevanti, con effetti differenziati per alcune categorie semantiche, per lo stile relazionale e per il contesto territoriale dell'alloggio.

INDICE

ABSTRACT	I
INDICE.....	II
INDICE DELLE FIGURE.....	IV
INDICE DELLE TABELLE.....	V
INTRODUZIONE	1
CAPITOLO I – L’evoluzione digitale nel settore dell’ospitalita’	3
1.1 La sharing economy	3
1.2 Disruptive innovation.....	5
1.3 La costruzione della fiducia digitale	6
1.4 Comunicazione e creazione di valore nella sharing economy	8
CAPITOLO II – Airbnb.....	10
2.1 La storia	10
2.2 Il funzionamento	11
2.2.1 Airbnb come marketplace a due lati	12
2.2.2 Business Model Canvas	13
2.3 La centralità dell'annuncio nel processo di prenotazione	16
2.4 Airbnb in Italia	18
CAPITOLO III – Background teorico e obiettivo di ricerca.....	20
3.1 Analisi della letteratura	20
3.1.1 Gli attributi strutturali e la localizzazione	21
3.1.2 Profilo dell’Host.....	22
3.1.3 La qualità della presentazione e l'analisi linguistica	22
3.2 I Gap nella letteratura esistente.....	23
3.3 Obiettivi della ricerca.....	24
CAPITOLO IV – Metodologia.....	26
4.1 Origine dei dati e costruzione del campione.....	26
4.1.1 Analisi del dataset iniziale e variabili strutturali	26

4.1.2 Filtraggio dei dati e definizione del perimetro d'analisi.....	28
4.2 Integrazione dei dati testuali e analisi semantica.....	29
4.2.1 Web Scraping	29
4.2.2 Analisi testuale e creazione dei nuovi indici	29
4.3 Misure	35
4.3.1 Definizione e operazionalizzazione delle variabili	35
4.3.2 Analisi descrittiva.....	40
4.3.3 Matrice di correlazione.....	53
CAPITOLO V – Risultati empirici.....	55
5.1 Specificazione del modello econometrico	55
5.2 Analisi dei risultati delle regressioni tramite Script R	56
5.2.1 Metriche di complessità del testo	57
5.2.2 Metriche relazionali.....	58
5.2.3 Metriche semantiche	60
5.3 Approfondimento metodologico tramite software LIWC.....	62
5.3.1 Metriche di complessità del testo	63
5.3.2 Metriche relazionali.....	65
5.3.3 Metriche semantiche	67
5.4 Il ruolo del contesto urbano: effetti di moderazione della urban class	70
5.5 Il ruolo dell'host: effetti di moderazione della variabile Superhost.....	78
CAPITOLO VI – Conclusioni	84
6.1 Sintesi dei principali risultati	84
6.2 Implicazioni teoriche	85
6.3 Implicazioni manageriali	86
6.4 Limiti della ricerca e sviluppi futuri	87
BIBLIOGRAFIA	89
SITOGRAFIA	93

INDICE DELLE FIGURE

Figura 1: Capitalizzazione di mercato di Airbnb e delle principali catene alberghiere al momento dell'IPO – Fonte: elaborazione propria su dati pubblici (Yahoo Finance, 2020).....	11
Figura 2: Business Model Canvas di Airbnb.....	16
Figura 3: Script R utilizzato per l'analisi semantica e il calcolo degli indici linguistici	34
Figura 4: Capacità ricettiva degli alloggi	42
Figura 5: Distribuzione dello status di Superhost nel campione	43
Figura 6: Lunghezza delle descrizioni.....	44
Figura 7: Confronto fra i focus tematici delle descrizioni.....	45
Figura 8: Distribuzione degli annunci per intervalli di prezzo medio giornaliero	46
Figura 9: Distribuzione degli annunci per intervalli di ricavo medio per notte disponibile	47
Figura 10: Distribuzione degli annunci per intervalli di occupation rate	48
Figura 11: Distribuzione degli annunci per tipologia di alloggio.....	49
Figura 12: Distribuzione delle principali categorie di alloggio.....	49
Figura 13: Distribuzione degli alloggi per regioni	50
Figura 14: Distribuzione degli alloggi per città.....	51
Figura 15: Distribuzione delle politiche di cancellazione	52
Figura 16: Effetti delle variabili relazionali su LogADR – SCRIPT R.....	59
Figura 17: Effetti delle variabili semantiche su LogADR – SCRIPT R.....	61
Figura 18: Effetti delle variabili semantiche su Log Rating – SCRIPT R	62
Figura 19: Effetti delle variabili relazionali su LogADR – LIWC.....	66
Figura 20: Moderazione dello status di Superhost: effetto di Affective su LogADR - LIWC	80
Figura 21: Moderazione dello status di Superhost: effetto di Guest orientation su LogRevPAN - LIWC	83

INDICE DELLE TABELLE

Tabella 1: Operazionalizzazione delle variabili	39
Tabella 2: Statistiche descrittive.....	40
Tabella 3: Matrice di correlazione di Spearman.....	53
Tabella 4: Risultati delle regressioni: effetti della struttura e complessità del testo sulle performance – Script R.....	57
Tabella 5: Risultati delle regressioni: impatto delle strategie comunicative sulle performance – Script R.....	58
Tabella 6: Risultati delle regressioni: impatto delle variabili semantiche sulle performance – Script R.....	60
Tabella 7: Risultati delle regressioni: effetti della struttura e complessità del testo sulle performance – LIWC.....	64
Tabella 8: Risultati delle regressioni: impatto delle strategie comunicative sulle performance – LIWC.....	65
Tabella 9: Risultati delle regressioni: impatto delle variabili semantiche sulle performance – LIWC.....	68
Tabella 10: Risultati delle regressioni: effetto moderatori dell'area geografica sulla variabile pulizia – SCRIPT R.....	71
Tabella 11: Risultati delle regressioni: effetto moderatori dell'area geografica sulla variabile sicurezza – SCRIPT R.....	73
Tabella 12: Risultati delle regressioni: effetto moderatori dell'area geografica sulla variabile tempo libero – SCRIPT R.....	74
Tabella 13: Risultati delle regressioni: effetto moderatori dell'area geografica sulla variabile orientamento linguistico professionale – SCRIPT R.....	76
Tabella 14: Risultati delle regressioni: effetto moderatori dell'area geografica sulla percezione – LIWC.....	77
Tabella 15: Risultati delle regressioni: effetto moderatori dello status Superhost sul linguaggio affettivo – SCRIPT R	79
Tabella 16: Risultati delle regressioni: effetto moderatori dello status Superhost sull'orientamento all'host – SCRIPT R.....	81
Tabella 17: Risultati delle regressioni: effetto moderatori dello status Superhost sull'orientamento al guest – SCRIPT R	82

INTRODUZIONE

La diffusione delle piattaforme digitali ha profondamente trasformato il settore dell'ospitalità, modificando il modo in cui domanda e offerta si incontrano e ridefinendo i meccanismi attraverso cui viene creato e comunicato il valore. In particolare, le piattaforme di affitto di breve periodo, come Airbnb, hanno introdotto un modello di intermediazione peer-to-peer in cui la fiducia tra sconosciuti diventa un elemento centrale del processo di scambio.

In questo contesto, l'annuncio online assume un ruolo cruciale. In assenza di un brand alberghiero che funga da garanzia, il potenziale ospite deve basare le proprie decisioni su informazioni digitali, tra cui immagini, recensioni e descrizioni testuali. Se le recensioni sono state ampiamente studiate come meccanismo reputazionale, la descrizione dell'annuncio rappresenta una dimensione ancora parzialmente inesplorata, soprattutto per quanto riguarda il suo impatto diretto sulle performance economiche.

La letteratura esistente ha dimostrato che le caratteristiche strutturali dell'alloggio e la localizzazione geografica influenzano in modo significativo prezzo e domanda. Negli ultimi anni, l'attenzione si è spostata verso gli aspetti comunicativi, mostrando come il linguaggio possa agire da segnale informativo e incidere sulla percezione del valore. Tuttavia, gran parte degli studi si concentra sui titoli degli annunci o su contesti geografici extra-europei, lasciando aperti diversi interrogativi sul ruolo delle descrizioni complete e sul caso italiano.

A partire da queste considerazioni, il lavoro di ricerca si propone di analizzare in chiave esplorativa il legame tra comunicazione testuale e performance su Airbnb, con particolare attenzione alle descrizioni degli annunci. L'obiettivo è comprendere se specifiche strategie linguistiche, in termini di contenuto semantico, stile relazionale e orientamento comunicativo, siano associate a risultati economici differenti.

Dal punto di vista metodologico, l'analisi combina l'approccio econometrico con tecniche di elaborazione del linguaggio naturale, trasformando le descrizioni testuali in variabili quantitative utilizzabili nei modelli di regressione. L'analisi tiene inoltre conto del ruolo del contesto territoriale e della reputazione dell'host, introducendo effetti di moderazione per valutare se l'impatto del linguaggio vari in funzione di tali fattori.

La tesi è strutturata come segue. Il primo capitolo introduce il quadro teorico della sharing economy e della fiducia digitale. Il secondo capitolo è dedicato ad Airbnb, al suo

funzionamento e al ruolo centrale dell'annuncio nel processo di prenotazione. Il terzo capitolo presenta la revisione della letteratura e gli obiettivi della ricerca. Il quarto capitolo descrive la metodologia adottata e la costruzione delle variabili. Il quinto capitolo riporta i risultati empirici, mentre il capitolo conclusivo discute le principali evidenze, le implicazioni teoriche e manageriali, nonché i limiti della ricerca e le possibili direzioni future.

CAPITOLO I – L'evoluzione digitale nel settore dell'ospitalità'

La digitalizzazione ha trasformato in modo significativo i comportamenti sociali ed economici, modificando le modalità con cui individui e imprese attribuiscono valore ai beni e ai servizi. Questo processo non si traduce unicamente nella creazione di nuovi mercati, ma in un'evoluzione delle logiche di consumo. Rispetto ai modelli economici tradizionali, centrati sul possesso e sull'accumulazione di beni materiali, emerge oggi una maggiore attenzione all'accessibilità, alla flessibilità e all'uso efficiente delle risorse.

In questo scenario si afferma la Sharing Economy, come risultato di un progressivo superamento delle concezioni tradizionali sulle proprietà dei beni, dove il valore non è legato al possesso del bene, piuttosto alla possibilità di accedervi e utilizzarlo. Questo approccio consente di valorizzare risorse precedentemente inutilizzate o sottoutilizzate, trasformandole in nuove opportunità economiche e sociali per la collettività (Belk, 2014).

All'interno di questo contesto, le piattaforme online svolgono un ruolo centrale, fungendo da intermediari tra soggetti che offrono una risorsa e soggetti che ne richiedono l'utilizzo. La loro capacità di abbattere le barriere all'ingresso ha permesso a operatori non professionali di entrare nel mercato, dando origine ad un processo di "Disruptive Innovation" che ha portato a ridefinire gli equilibri del settore alberghiero (Guttentag, 2015).

Perché questo scambio tra sconosciuti possa funzionare, è necessario ridurre l'incertezza iniziale che caratterizza le interazioni online. In assenza di contatto diretto, la fiducia diventa un elemento centrale per il successo della transazione. In questo senso, le piattaforme mettono a disposizione specifici strumenti reputazionali come recensioni, descrizioni dettagliate e immagini, che aiutano gli utenti a formarsi un'idea più chiara dell'offerta, aumentando la percezione di affidabilità e qualità (Ert et al., 2016).

1.1 La sharing economy

Per comprendere la natura del cambiamento in atto nel settore dell'ospitalità, è necessario analizzare i due seguenti fenomeni distinti ma complementari: l'evoluzione dell'infrastruttura tecnologica e la ridefinizione delle priorità economiche dopo la crisi finanziaria del 2008.

Il punto di partenza di questa evoluzione va ricercato nella tecnologia. Fino all'inizio del nuovo millennio, il Web, nella sua versione 1.0, era strutturato come una vetrina digitale, ovvero uno spazio statico in cui le informazioni scorrevano in una sola direzione, dalle

aziende verso il pubblico. In quel contesto, il ruolo dell'utente era limitato alla semplice consultazione, poiché poteva visionare le pagine web senza avere alcuna possibilità di replica o di interazione.

Il Web 2.0 ha modificato in modo significativo le modalità di utilizzo della rete, riducendo la distanza tra chi produce i contenuti e chi li utilizza. La rete ha così smesso di essere un semplice archivio di informazioni, trasformandosi in un ambiente dinamico fondato sulla partecipazione attiva degli utenti e sulla produzione di contenuti generati dagli stessi utenti (User Generated Content, UGC).

In questo nuovo contesto, l'utente assume il duplice ruolo di produttore e consumatore, definito in letteratura come "Prosumer" (Buhalis & Law, 2008). Nel turismo, questo si traduce nella capacità di creare valore attraverso la condivisione di esperienze. Le recensioni, le foto e i blog di viaggio non sono più contenuti accessori, ma si trasformano nel punto di riferimento principale per chi deve scegliere, perché percepite come più autentiche e sincere rispetto ai messaggi promozionali delle aziende.

La tecnologia ha reso possibile lo sviluppo della Sharing Economy, ma non ne rappresenta l'unico fattore. Un ruolo rilevante è stato svolto anche dal contesto economico globale con specifico riferimento alla crisi economico-finanziaria del 2008 che ha spinto molte famiglie a rivedere le proprie priorità di consumo.

In quel clima di incertezza, il possesso esclusivo di beni costosi, quali una seconda casa o un'auto di proprietà, si è trasformato in un costo fisso difficile da sostenere. È proprio in questo cambio di prospettiva che trova applicazione la definizione di Botsman e Rogers (2010) secondo cui la Sharing Economy nasce come risposta intelligente per "sbloccare" il valore delle risorse che lasciamo ferme. Si tratta di valorizzare quella che gli autori definiscono "capacità inutilizzata", trasformando un bene passivo, che genera costi, in una risorsa attiva in grado di produrre reddito.

Il vero elemento innovativo, che distingue le piattaforme online dai vecchi annunci sui giornali, è la riduzione dei costi di transazione. Prima delle piattaforme digitali, affittare una stanza in casa propria a un estraneo comportava costi di ricerca, negoziazione e verifica della fiducia troppo alti per un privato. Le piattaforme hanno automatizzato questi processi, rendendo economicamente vantaggioso anche lo scambio di servizi di piccola scala tra individui.

Si passa così da un'economia basata sulla proprietà a una basata sull'accesso, dove l'utilizzo temporaneo del bene diventa preferibile al suo acquisto (Belk, 2014).

Accanto alle motivazioni economiche, la diffusione della Sharing Economy è stata favorita anche da una spinta di tipo culturale, legata alla ricerca di autenticità. Il viaggiatore contemporaneo tende, infatti, a prendere le distanze dalle esperienze standardizzate offerte dalle grandi catene alberghiere, spesso percepite come ambienti impersonali e separati dal contesto reale della destinazione.

L'home-sharing risponde a questa esigenza offrendo un'esperienza più vicina alla vita quotidiana del luogo visitato. In questo caso, l'alloggio non rappresenta solo uno spazio in cui soggiornare, ma consente di entrare più facilmente in contatto con le abitudini del quartiere, i ritmi della città e le persone che la vivono. L'esperienza di viaggio si trasforma così da semplice osservazione della destinazione a partecipazione alla vita del luogo (Dredge, Gyimóthy, 2015).

1.2 Disruptive innovation

La diffusione della Sharing Economy nel settore turistico non può essere considerata come una semplice evoluzione del mercato o l'emergere di un nuovo concorrente. Si tratta, piuttosto, di una vera e propria rivoluzione del settore. Dietro questa crescita esponenziale non c'è il caso, ma una dinamica di mercato ben definita. Per analizzarla correttamente, è fondamentale fare riferimento alla teoria della "Disruptive Innovation", utilizzata da Guttentag (2015) nel suo articolo scientifico per spiegare le dinamiche di successo di piattaforme online come Airbnb.

Secondo questa teoria, le grandi aziende leader (in questo caso, le catene alberghiere) tendono naturalmente a migliorare i propri servizi per soddisfare i clienti più esigenti e redditizi. Aggiungono servizi, lusso, reception 24 ore su 24, piscine ed altri comfort. Così facendo, però, finiscono per "superare" le esigenze del cliente medio, offrendo un prodotto eccellente ma costoso e complesso. È in questo spazio vuoto che si inserisce il "disruptor". La disruption non nasce offrendo un prodotto migliore fin da subito, ma proponendo un'alternativa che, agli occhi degli esperti, appare decisamente inferiore, ma che risulta più economica, semplice e accessibile (Clayton Christensen, 1997).

È esattamente questa la dinamica che ha ingannato gli albergatori. Nella fase iniziale, l'industria dell'ospitalità ha osservato la nascita delle piattaforme peer-to-peer con un misto di scetticismo e sufficienza. Gli hotel hanno, infatti, classificato questo fenomeno come una soluzione di fascia inferiore limitata esclusivamente a studenti e viaggiatori low-cost che non rientravano comunque nel loro target di profitto. Erano convinti che la

loro clientela business e luxury non avrebbe mai rinunciato al comfort per l'incertezza di una casa privata.

Ed è qui che è scattata quella che Christensen definisce la "trappola". Il modello della Disruption prevede che il nuovo entrante conquisti prima il mercato dal basso, per poi iniziare una rapida risalita verso l'alto. Una volta consolidata la fiducia e migliorata la piattaforma tecnologica, l'offerta della Sharing Economy si è evoluta drasticamente. Non più solo sistemazioni improvvisate, ma interi appartamenti, ville con piscina e alloggi unici. La qualità del servizio è salita fino a incrociare, e spesso superare, quella degli hotel, ma mantenendo prezzi competitivi e offrendo quella autenticità che gli alberghi non potevano replicare. Quando i grandi gruppi si sono accorti che Airbnb stava sottraendo loro non solo i turisti, ma anche la clientela business e luxury, la diffusione della Sharing Economy era ormai completa e irreversibile.

A rendere la competizione particolarmente complessa per gli operatori tradizionali contribuisce anche una rilevante differenza nella struttura dei costi. Il modello alberghiero è caratterizzato da un'elevata intensità di capitale e l'espansione in nuovi mercati richiede investimenti significativi in immobili, ristrutturazioni, autorizzazioni e personale stabile.

Viceversa, le piattaforme della Sharing Economy adottano un modello asset-light, basato su costi marginali contenuti. Airbnb, ad esempio, non è proprietaria degli alloggi presenti sulla piattaforma e non sostiene direttamente i costi operativi legati alla gestione delle strutture. Questo assetto consente una maggiore flessibilità dell'offerta, che può adattarsi rapidamente alle variazioni della domanda. In presenza di un aumento improvviso dei flussi turistici, l'offerta può crescere attraverso l'ingresso di nuovi host, mentre le strutture alberghiere restano vincolate a una capacità fisica difficilmente modificabile nel breve periodo (Zervas, Proserpio e Byers, 2017).

1.3 La costruzione della fiducia digitale

Anche se parlare di Sharing Economy suggerisce l'idea di solidarietà, la realtà delle piattaforme è più complessa e dietro l'immagine della condivisione, infatti, ci sono meccanismi di mercato e interessi economici reali.

In un mercato caratterizzato interamente da un'interfaccia digitale, dove l'acquirente e il venditore sono completi estranei e il prodotto, ovvero l'abitazione privata, non può essere ispezionato fisicamente prima dell'acquisto, la percezione del rischio risulta naturalmente elevata. L'utente non si limita ad acquistare un servizio, ma richiede di entrare negli spazi

personali di uno sconosciuto, creando un senso di insicurezza che non si prova nel turismo classico.

Nel settore alberghiero, tale incertezza viene limitata dalla reputazione del brand aziendale. La prenotazione presso una catena nota offre una garanzia implicita perchè il cliente sa esattamente cosa aspettarsi in termini di pulizia, servizi e sicurezza, indipendentemente dalla posizione geografica.

Nel contesto della peer-to-peer accommodation, dove questa garanzia istituzionale viene a mancare e ogni alloggio è una soluzione distinta, la mancanza di fiducia deve essere colmata dai contenuti digitali. In assenza di un'interazione fisica diretta, le fotografie, le descrizioni testuali e i profili utente smettono di essere semplice materiale promozionale ma diventano l'unico strumento in grado di ridurre l'asimmetria informativa. I contenuti visivi e testuali fungono da evidenza concreta dell'offerta, permettendo all'utente di valutare a distanza non solo le caratteristiche strutturali dell'immobile, ma anche l'affidabilità e la personalità dell'host (Ert, Fleischer e Magen, 2016).

Questo meccanismo ha generato un cambiamento radicale nei canoni estetici della comunicazione turistica. Mentre l'industria alberghiera ha tradizionalmente puntato su immagini professionali, studiate con grandangoli e luci artificiali per mostrare gli ambienti in una perfezione ideale, sulle piattaforme digitali emerge il valore della veridicità. Spesso, una fotografia scattata dall'host con mezzi non professionali, che mostra dettagli di vita quotidiana o piccole imperfezioni, risulta paradossalmente più efficace di uno scatto da catalogo. L'utente moderno non cerca la perfezione formale, ma la garanzia di autenticità. L'immagine amatoriale segnala l'assenza di artificio e di manipolazione di marketing, agendo come un segnale di onestà che predispone a una maggiore fiducia verso l'host.

Parallelamente alla componente visiva, è fondamentale la costruzione della reputazione, garantita dal sistema delle recensioni. Le valutazioni lasciate dagli ospiti precedenti non svolgono solo la funzione di giudicare parametri oggettivi come la pulizia o la posizione, ma costruiscono nel tempo la reputazione dell'host ed il feedback contribuisce a costruire una garanzia condivisa tra gli utenti (Filieri, 2016).

Le recensioni degli altri viaggiatori aiutano a ridurre la diffidenza iniziale, trasformando l'host da sconosciuto a soggetto percepito come affidabile. Sapendo di essere valutato, l'host è incentivato a mantenere un buon livello di servizio, mentre foto e opinioni permettono agli utenti di formarsi un giudizio prima del soggiorno. In questo modo, le

piattaforme riescono a ricreare online un clima di fiducia simile a quello che, in passato, derivava dal passaparola diretto.

1.4 Comunicazione e creazione di valore nella sharing economy

Accanto al ruolo svolto dai contenuti digitali nella costruzione della fiducia, la letteratura sulla sharing economy riconosce alla comunicazione una funzione più ampia nel processo di creazione del valore. Le informazioni scambiate all'interno delle piattaforme non si limitano a rendere possibile la transazione, ma contribuiscono a definire il modo in cui l'offerta viene percepita, valutata e confrontata dagli utenti. In questo senso, la comunicazione rappresenta una componente strutturale delle dinamiche di mercato che caratterizzano le piattaforme digitali (Sundararajan, 2016).

La comunicazione interviene nel modo in cui l'offerta viene presentata e interpretata all'interno della piattaforma. Attraverso il linguaggio e le modalità di esposizione delle informazioni, i fornitori del servizio possono costruire una rappresentazione dell'esperienza proposta, enfatizzare determinati aspetti e orientare l'attenzione dei potenziali guest. Questo processo contribuisce a differenziare offerte che, dal punto di vista strutturale, possono risultare simili, influenzando la valutazione complessiva del servizio (Pavlou & Dimoka, 2006).

La letteratura evidenzia come la comunicazione partecipi attivamente alla formazione del valore percepito, incidendo sulle aspettative degli utenti prima dello scambio. La narrazione dell'offerta e il tono comunicativo adottato possono influenzare la percezione di qualità, coerenza e adeguatezza del servizio, andando oltre la semplice trasmissione di informazioni funzionali. In questo senso, la comunicazione diventa parte integrante del posizionamento dell'offerta all'interno del mercato (Botsman & Rogers, 2010).

Un ulteriore aspetto riguarda il legame tra comunicazione e reputazione nelle piattaforme di sharing economy. Le informazioni scambiate nel tempo contribuiscono a consolidare un'immagine dell'offerta e del fornitore del servizio, influenzando le valutazioni degli utenti e la propensione allo scambio anche in presenza di alternative simili (Ba & Pavlou, 2002; Gefen et al., 2003).

La comunicazione si integra così con i meccanismi reputazionali, rafforzando il funzionamento dei mercati digitali. La narrazione dell'esperienza proposta può quindi essere interpretata come una leva attraverso cui le piattaforme di sharing economy contribuiscono alla creazione di valore economico. Attraverso testi e segnali

comunicativi, l'offerta viene posizionata, confrontata e valutata, incidendo sulle dinamiche competitive che caratterizzano questi mercati.

CAPITOLO II – Airbnb

Dopo aver delineato il quadro teorico della Sharing Economy e analizzato le dinamiche di fiducia che regolano gli scambi tra estranei, è opportuno spostare l'analisi sul soggetto che, più di ogni altro, ha saputo rappresentare questa rivoluzione trasformandola in un fenomeno globale: Airbnb.

Questa piattaforma non rappresenta solo un caso di successo imprenditoriale, ma costituisce il modello di riferimento del nuovo turismo digitale.

Analizzare Airbnb significa osservare da vicino come i principi teorici della Disruptive Innovation abbiano trovato applicazione pratica, ridefinendo gli equilibri di un settore, quello dell'ospitalità, che per decenni era rimasto immutato.

In questo capitolo verrà esaminata la piattaforma di Airbnb, partendo dalle sue origini storiche fino alla definizione del suo attuale modello di business, basato sulla logica del marketplace a due lati. Si approfondiranno inoltre le caratteristiche strutturali dell'annuncio, dalle strategie di prezzo agli elementi visivi e testuali, per comprendere quali siano i fattori che oggi determinano le scelte di prenotazione degli utenti.

2.1 La storia

Per capire davvero cosa sia Airbnb oggi, bisogna guardare alle sue origini. La storia dell'azienda è l'esempio perfetto di come un'idea nata per risolvere un piccolo problema pratico possa trasformarsi in un business mondiale. Tutto ha inizio nell'autunno del 2007 a San Francisco. I fondatori, Brian Chesky e Joe Gebbia, erano due designer neolaureati che si trovavano ad affrontare una difficoltà comune, ossia l'impossibilità di sostenere il pagamento dell'affitto.

Notando che tutti gli hotel della città erano al completo per una conferenza di design, decisero di cogliere l'occasione comprando tre materassi ad aria, sistemandoli in salotto e offrendo posto letto e colazione ai partecipanti dell'evento. Nacque così "AirBed & Breakfast". Quella che sembrava solo una soluzione d'emergenza per "arrotondare" si rivelò l'inizio di qualcosa di molto più grande.

Tuttavia, il successo non fu immediato. Nel primo anno la startup faticò a imporsi sul mercato. Gli investitori della Silicon Valley erano scettici e l'idea che le persone accettassero di dormire in casa di sconosciuti sembrava folle e pericolosa. La vera svolta arrivò nel 2009 grazie a un'intuizione fondamentale riguardante l'importanza delle

immagini. I fondatori si resero conto che gli appartamenti sul sito restavano vuoti perché le foto, scattate dai proprietari con i cellulari dell'epoca, erano buie e poco invitanti. Nessuno si fidava di prenotare ciò che non riusciva a vedere bene. Chesky e Gebbia decisero allora di volare a New York per scattare personalmente foto professionali alle case degli host. L'effetto fu immediato e attraverso l'utilizzo di immagini belle e luminose, le prenotazioni raddoppiarono in una settimana. Fu una dimostrazione concreta dell'importanza delle immagini nel processo di costruzione della fiducia online (Gallagher, 2017).

Da quel momento, la crescita è stata inarrestabile. Il nome fu accorciato in "Airbnb" e l'offerta è aumentata diversificandosi, non più solo divani o stanze condivise, ma intere case, ville e appartamenti di lusso. L'apice di questo percorso è arrivato il 10 dicembre 2020 con lo sbarco in borsa a Wall Street. In un anno difficilissimo per il turismo a causa della pandemia, Airbnb ha dato prova di una notevole resilienza, raggiungendo una valutazione di oltre 100 miliardi di dollari e superando il valore di colossi storici come Marriott e Hilton messi insieme, come mostrato in *Figura 1*.

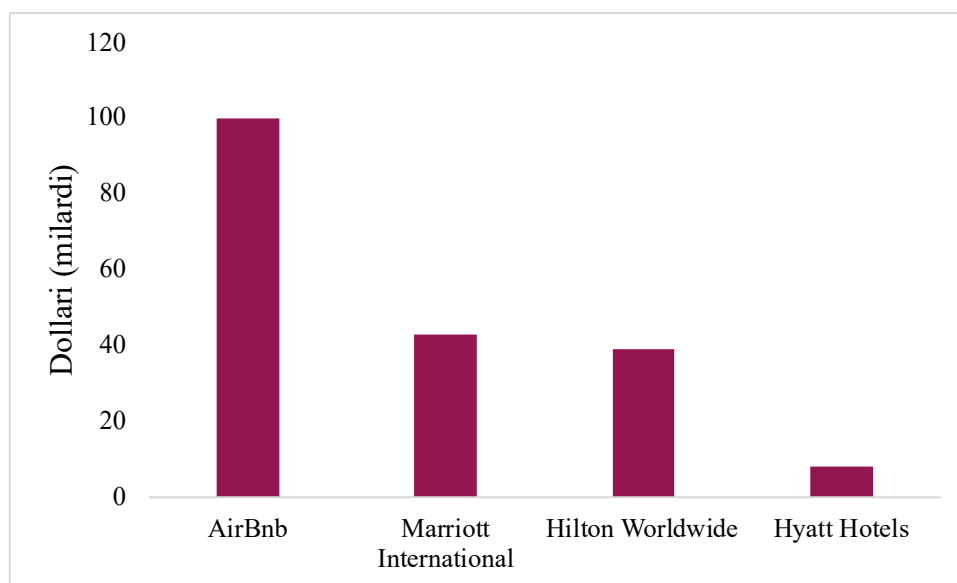


Figura 1: Capitalizzazione di mercato di Airbnb e delle principali catene alberghiere al momento dell'IPO – Fonte: elaborazione propria su dati pubblici (Yahoo Finance, 2020)

Oggi Airbnb non è più solo un sito di prenotazioni, ma ha cambiato per sempre il modo in cui le persone viaggiano e abitano (Dann et al., 2019).

2.2 Il funzionamento

A differenza delle catene alberghiere tradizionali, che seguono un modello operativo lineare, Airbnb adotta una logica basata sull'intermediazione e sull'interazione tra gli utenti. La piattaforma, infatti, non offre direttamente un prodotto di sua proprietà, ma

mette in contatto soggetti che dispongono di un alloggio con persone interessate a utilizzarlo.

2.2.1 Airbnb come marketplace a due lati

Airbnb può essere interpretata come una piattaforma digitale a due lati (two-sided platform), il cui ruolo principale consiste nell'intermediare le interazioni tra due gruppi distinti di utenti, da un lato i proprietari di immobili che desiderano affittare il proprio alloggio (host), dall'altro i viaggiatori alla ricerca di una sistemazione temporanea (guest). La letteratura economica definisce le piattaforme a due lati come intermediari che facilitano l'interazione tra gruppi differenti di utenti, creando valore facilitando la partecipazione e l'interazione tra gli utenti (Rochet & Tirole, 2003; Armstrong, 2006).

L'opportunità economica alla base della nascita di tali piattaforme deriva spesso dalla presenza di una frizione di mercato, ovvero da inefficienze nei meccanismi tradizionali di incontro tra domanda e offerta. In mercati caratterizzati da elevati costi di ricerca, informazione e coordinamento, le piattaforme digitali agiscono come matchmaker, riducendo tali frizioni attraverso processi di matching più efficienti e standardizzati (Evans & Schmalensee, 2016). Nel caso di Airbnb, la piattaforma interviene in un mercato frammentato e poco trasparente, creando un'infrastruttura digitale in grado di connettere direttamente host e guest e di abbattere i costi di ricerca associati all'affitto di alloggi di breve periodo.

Un elemento centrale che caratterizza le piattaforme a due lati è la presenza di esternalità di rete indirette e incrociate. In questi contesti, l'utilità che un utente trae dalla partecipazione alla piattaforma non dipende esclusivamente dal numero di utenti appartenenti allo stesso lato del mercato, ma anche, e soprattutto, dalla dimensione dell'altro lato (Rochet & Tirole, 2006).

Nel caso di Airbnb, un aumento del numero di host rende la piattaforma più attrattiva per i guest grazie a una maggiore varietà di offerte disponibili e allo stesso tempo, un aumento del numero di guest accresce il valore della piattaforma per gli host, aumentando la probabilità di prenotazione degli alloggi.

Questa interdipendenza tra i due lati del mercato implica che le decisioni di partecipazione siano caratterizzate da una forma di elasticità incrociata dove la domanda di un gruppo risponde alle variazioni nella dimensione dell'altro gruppo. Di conseguenza, la strategia della piattaforma non può concentrarsi su un solo lato, ma deve incentivare

simultaneamente la partecipazione di entrambi, bilanciando prezzi, regole e strumenti di interazione (Armstrong, 2006; Caillaud & Jullien, 2003).

Il successo di Airbnb può essere interpretato anche grazie alla sua capacità di raggiungere una massa critica su entrambi i lati del mercato. In mercati caratterizzati da forti esternalità di rete, tali dinamiche possono condurre a fenomeni di tipping, in cui la domanda tende a concentrarsi su una o poche piattaforme dominanti. Una volta superata una determinata soglia critica, la piattaforma diventa progressivamente più attrattiva rispetto ai concorrenti, rendendo difficile per nuove imprese entrare nel mercato o sottrarre quote significative (Evans & Schmalensee, 2016). In questo quadro, Airbnb si configura come un intermediario centrale del mercato degli affitti di breve periodo, la cui posizione è rafforzata dalla capacità di attrarre e coordinare un numero crescente di utenti su entrambi i lati della piattaforma.

Il valore della piattaforma dipende dalla qualità e dalla quantità delle informazioni disponibili per facilitare il matching tra host e guest e gli annunci pubblicati dagli host assumono un ruolo centrale nel funzionamento del mercato. In particolare, le informazioni fornite negli annunci contribuiscono a ridurre l'incertezza, a migliorare l'efficienza del matching e a sostenere le dinamiche di partecipazione che alimentano le esternalità di rete proprie della piattaforma.

2.2.2 Business Model Canvas

Per comprendere la dinamica di funzionamento di Airbnb, è necessario analizzare nel dettaglio la strategia attraverso il Business Model Canvas (Osterwalder & Pigneur, 2010). Questo framework, articolato in nove blocchi costitutivi, rappresenta una mappa logica divisa in due sezioni, unite al centro dalla Proposta di Valore. La parte destra del modello si concentra sul mercato, sulle relazioni con i clienti e sulla generazione di ricavi. In Airbnb, questo è rappresentato dall'esperienza e dall'interazione tra Host e Guest. La parte sinistra è invece dedicata all'efficienza operative e comprende le infrastrutture, le attività interne e la struttura dei costi necessarie a far funzionare il sistema.

Questa analisi consente di spiegare in che modo la piattaforma riesca a generare, distribuire e catturare valore in un mercato globale.

Segmenti di clientela

Airbnb opera in un mercato multi-versante, la cui complessità sta nel dover soddisfare simultaneamente due segmenti con bisogni opposti ma complementari: Host e Guest.

Proposte di valore

La proposta di valore di Airbnb è articolata e differenziata in base ai due principali segmenti di utenti. Per gli host, il beneficio principale consiste nella possibilità di generare reddito attraverso la messa a disposizione del proprio alloggio. A questo si affiancano ulteriori elementi di valore, quali il supporto alla sicurezza, garantito dalla verifica degli ospiti e dalla copertura assicurativa AirCover, e la semplificazione della gestione degli affitti tramite strumenti dedicati, come il calendario e i sistemi di pagamento automatici. Per i guest, il valore offerto risiede nell'accesso a un'offerta di alloggi ampia e diversificata, spesso disponibile a condizioni di prezzo più competitive rispetto alle strutture alberghiere nelle stesse aree.

Canali

Il modello è nativo digitale. La distribuzione del servizio avviene esclusivamente attraverso la piattaforma proprietaria. Il Sito web e l'App mobile sono i punti di contatto primari, progettati con un'interfaccia intuitiva che guida l'utente dalla ricerca alla prenotazione in pochi click.

Relazioni con i clienti

La relazione con gli utenti è prevalentemente automatizzata e basata su modalità di self-service, poiché la piattaforma mette a disposizione gli strumenti necessari affinché host e guest possano operare in autonomia. Allo stesso tempo Airbnb adotta sistemi di verifica degli utenti e di recensioni reciproche, finalizzati a tutelare entrambe le parti, e offre inoltre un servizio di assistenza attivo 24 ore su 24.

Attività chiave

Non dovendo gestire immobili fisici, le attività principali di Airbnb sono prevalentemente legate al funzionamento della piattaforma. Emerge con maggiore evidenza che l'azienda si occupa dello sviluppo e della manutenzione dell'infrastruttura digitale, curando l'aggiornamento dei sistemi di matching tra domanda e offerta e garantendo elevati standard di sicurezza informatica. A queste attività si affiancano le iniziative di marketing e di costruzione della community, con l'obiettivo di attrarre nuovi utenti su entrambi i lati del mercato e a sostenere gli effetti di rete. Un ulteriore ambito centrale riguarda la gestione della fiducia e della sicurezza, attraverso il controllo degli standard di utilizzo della piattaforma.

Risorse chiave

Le risorse non sono di natura fisica, come nel caso delle catene alberghiere tradizionali, ma prevalentemente digitali e relazionali. Un ruolo centrale è svolto dalla piattaforma tecnologica e dai dati generati dalle interazioni degli utenti, che consentono di migliorare il funzionamento dei sistemi di ricerca e di supportare le decisioni di prezzo. Accanto alla dimensione tecnologica, la community e il brand rappresentano una risorsa strategica fondamentale. La rete di host e guest costituisce infatti il principale patrimonio della piattaforma, poiché senza la partecipazione attiva degli host non sarebbe possibile offrire alcun servizio.

Partner chiave

I partner chiave di Airbnb sono molteplici e contribuiscono in modo determinante al funzionamento e alla scalabilità della piattaforma. Un ruolo centrale è svolto dagli host, che non sono soltanto utenti del servizio ma veri e propri partner nella creazione dell'offerta. A questi si affiancano i fornitori di servizi tecnologici e di pagamento, che garantiscono transazioni sicure, affidabili e rapide, e le compagnie assicurative, fondamentali per la gestione e la copertura dei rischi connessi alle attività di hosting. Ulteriori partner rilevanti sono i fotografi, che contribuiscono a migliorare la qualità visiva degli annunci, e le società di gestione degli affitti per conto dei privati, che supportano gli host nelle attività operative, come l'accoglienza e la pulizia. Infine, un ruolo non secondario è svolto dagli attori istituzionali e regolatori, la cui interazione con la piattaforma influenza le possibilità di diffusione ed espansione del modello di business.

Struttura dei costi

Airbnb opera con una struttura dei costi relativamente snella, tipica delle imprese tecnologiche. Le principali voci di costo sono di natura fissa e riguardano soprattutto gli investimenti in ricerca e sviluppo, necessari per sostenere il lavoro di ingegneri software e data scientist impegnati nel miglioramento continuo della piattaforma. A questi si aggiungono rilevanti spese di marketing, finalizzate a supportare la crescita del brand e l'acquisizione di nuovi utenti su scala globale, oltre ai costi legali e di lobbying, legati alla gestione delle regolamentazioni locali nei diversi mercati. Al contrario, i costi marginali associati all'inserimento di un nuovo alloggio sulla piattaforma risultano prossimi allo zero.

Flussi di ricavi

Il modello di ricavo di Airbnb è di tipo transazionale ed è basato sull'applicazione di commissioni di servizio per ogni prenotazione conclusa con successo sulla piattaforma. Airbnb trattiene una percentuale sia dal lato dell'offerta sia dal lato della domanda. Agli host viene applicata una commissione generalmente compresa tra il 3% e il 5%, mentre ai guest viene addebitata una commissione variabile, solitamente compresa tra il 5% e il 15% del subtotale della prenotazione. A queste entrate si aggiungono le commissioni applicate ai fornitori di esperienze, che possono arrivare fino al 20% del valore del servizio offerto. Questo sistema di "doppia trattenuta" consente all'azienda di allineare direttamente i propri ricavi al volume delle transazioni generate, beneficiando della crescita dell'attività sulla piattaforma.

Di seguito viene presentata una rappresentazione sintetica del Business Model Canvas di Airbnb.

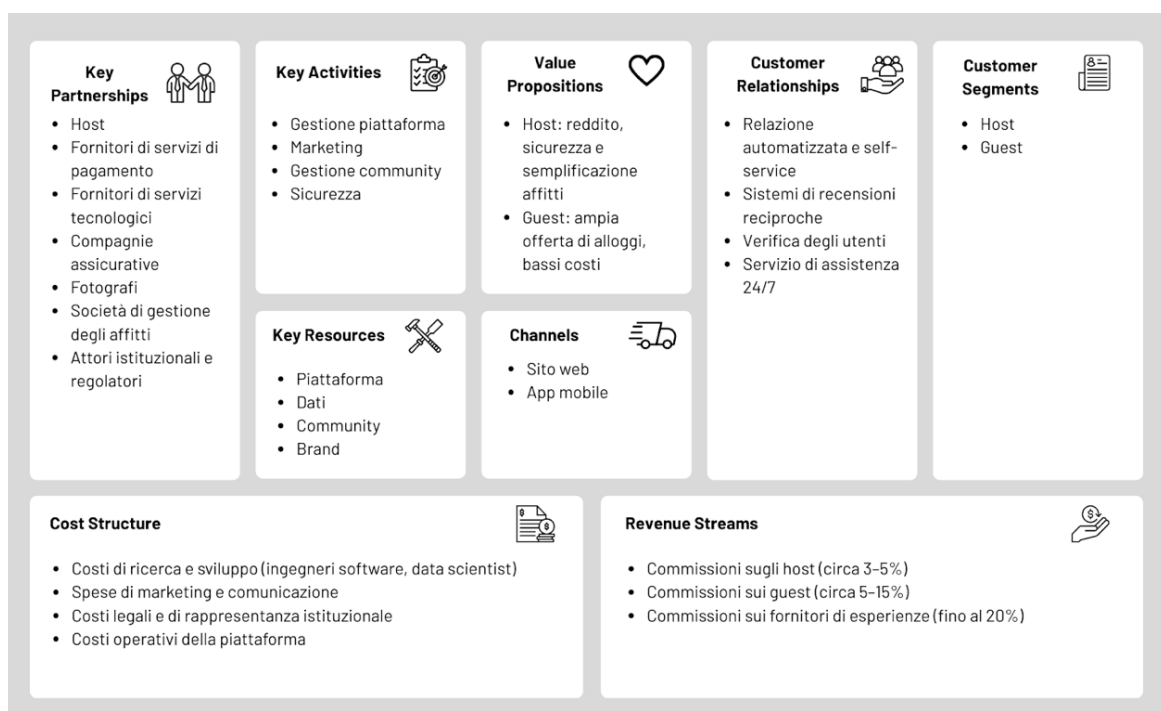


Figura 2: Business Model Canvas di Airbnb

2.3 La centralità dell'annuncio nel processo di prenotazione

Nel modello operativo di Airbnb, la transazione economica avviene in un ambiente puramente digitale, caratterizzato da asimmetria informativa. A differenza del settore alberghiero tradizionale, dove il brand funge da garanzia, nel sistema peer-to-peer ogni alloggio è unico. Il potenziale guest si trova quindi nella condizione di dover effettuare

in anticipo il pagamento per un servizio di cui potrà verificare la reale qualità solo una volta giunto a destinazione.

In questo contesto, l'annuncio assume un ruolo centrale, andando oltre la funzione di semplice strumento promozionale e rappresentando il principale punto di contatto tra host e guest. Attraverso la presentazione online, l'utente si forma un'idea preliminare dell'alloggio e di chi lo gestisce, la cura nella descrizione e nelle immagini contribuisce a trasmettere la qualità della casa e l'affidabilità dell'host (Liang et al., 2017). La probabilità di ottenere una prenotazione dipende quindi dalla capacità dell'host di utilizzare in modo efficace contenuti visivi e testuali per ridurre l'incertezza dell'utente.

Il primo punto di contatto, più immediato, è costituito dalla componente visiva. Le fotografie assumono un ruolo fondamentale nel processo decisionale. L'elaborazione delle immagini precede quella testuale con foto professionali, luminose e ordinate che non soddisfano solo un bisogno estetico, ma fungono da prova dell'esistenza e dello stato dell'immobile, riducendo la percezione del rischio (Ert et al., 2016).

Parallelamente alla dimensione visiva, la politica di prezzo rappresenta la seconda variabile determinante. Nonostante la piattaforma metta a disposizione strumenti algoritmici di Smart Pricing per adeguare le tariffe alla domanda di mercato in tempo reale, la decisione finale rimane nelle mani dell'Host. Il prezzo non è solo una cifra da pagare, ma un vero e proprio indizio sulla qualità della casa. Se da un lato una richiesta troppo alta allontana i potenziali ospiti, dall'altro lato anche una troppo bassa può essere controproducente e spesso venire letta come un campanello d'allarme, portando il guest a dubitare dell'alloggio.

Tuttavia, le foto dell'immobile e la logica del prezzo devono trovare una conferma pratica nella componente descrittiva. Se le immagini servono ad attirare l'attenzione, le parole servono a verificare se l'alloggio è funzionale alle proprie necessità, rispondendo a dubbi che una foto non può chiarire. Un titolo preciso e una descrizione sincera non sono semplici testi di riempimento, ma forniscono le istruzioni dell'alloggio. L'utente cerca nell'annuncio dettagli specifici per la propria quotidianità, dettagli concreti a permettere all'ospite di proiettarsi mentalmente nella casa prima ancora di partire (Guttentag, 2015).

Tutti questi elementi, tuttavia, restano informazioni dichiarate dall'host finché non entra in gioco la reputazione, che rappresenta il fattore decisivo nel processo di scelta. Il sistema delle recensioni svolge una funzione di verifica, poiché anche un annuncio visivamente curato e proposto a un prezzo competitivo può risultare poco convincente in presenza di

valutazioni negative legate, ad esempio, alla pulizia o alla reattività dell'host. Le valutazioni e i commenti degli altri viaggiatori forniscono quindi una conferma esterna dell'affidabilità dell'offerta, contribuendo a rendere credibili le informazioni presentate nell'annuncio (Zervas et al., 2015). Il successo dell'annuncio dipende quindi dall'equilibrio complessivo tra gli elementi che l'host presenta, ovvero le immagini, il prezzo, la descrizione testuale e le valutazioni lasciate dagli altri utenti.

2.4 Airbnb in Italia

L'Italia rappresenta uno dei contesti più rilevanti a livello globale per l'applicazione della sharing economy nel settore turistico. Le caratteristiche del territorio, contraddistinto dalla presenza diffusa di centri storici e borghi, insieme a una struttura della proprietà immobiliare fortemente frammentata tra numerosi privati, hanno favorito la diffusione di piattaforme come Airbnb (Picascia et al., 2017).

Nel contesto italiano, Airbnb si è affermata come un'alternativa ai modelli turistici tradizionali basati sulle grandi catene alberghiere, contribuendo alla valorizzazione di un ampio patrimonio di seconde case spesso inutilizzate. Inizialmente, la presenza della piattaforma si è concentrata nelle principali città d'arte, come Roma, Firenze, Venezia e Milano. Nel tempo, però, l'offerta si è estesa anche a località minori e aree interne, ampliando la distribuzione dei flussi turistici sul territorio nazionale (Celata et al., 2017). Per molti proprietari, l'affitto a breve termine ha così assunto un ruolo sempre più rilevante, evolvendo da attività occasionale a fonte di sostegno economico.

La crescita dell'offerta ha progressivamente reso il mercato più competitivo. In numerosi contesti, soprattutto urbani, l'aumento del numero di annunci ha ridotto la visibilità media delle singole inserzioni, intensificando la concorrenza tra alloggi spesso simili dal punto di vista strutturale e localizzativo. Parallelamente, il mercato italiano si caratterizza per una forte eterogeneità degli host e accanto a operatori più strutturati e professionali, operano numerosi host occasionali, con livelli molto diversi di esperienza e capacità gestionale. Questa frammentazione si riflette anche nelle modalità di presentazione degli annunci e nelle strategie comunicative adottate.

La rapida diffusione del fenomeno non è stata priva di effetti collaterali. In diversi contesti urbani, la diffusione degli affitti brevi ha contribuito alla progressiva trasformazione di alcuni quartieri in aree a forte presenza turistica, causando una riduzione della disponibilità abitativa per i residenti e di un cambiamento nella struttura commerciale delle città, con la sostituzione dei servizi utili ai residenti a favore di attività rivolte ai

visitatori. Queste dinamiche hanno reso necessario un intervento da parte delle istituzioni, che hanno introdotto misure regolamentari volte a disciplinare il settore, tra cui obblighi fiscali specifici e sistemi di identificazione degli immobili, orientando il mercato verso una gestione più strutturata e controllata.

Un ulteriore elemento distintivo del contesto italiano è rappresentato dalla forte stagionalità della domanda turistica. In molte destinazioni, l'attività di affitto breve si concentra in periodi limitati dell'anno, accentuando la competizione tra host nei mesi di minore affluenza. In questo scenario, la capacità di attirare l'attenzione degli ospiti diventa particolarmente rilevante, rendendo la descrizione dell'annuncio uno strumento centrale di differenziazione.

Il turismo in Italia, inoltre, è fortemente legato alla ricerca di autenticità e di immersione nello stile di vita locale. Gli alloggi offerti su Airbnb non vengono valutati esclusivamente per le caratteristiche strutturali, ma anche per la capacità di trasmettere un'esperienza coerente con il contesto culturale e territoriale in cui sono inseriti. Questo aspetto evidenzia come il linguaggio delle descrizioni rappresenti lo strumento attraverso cui l'host comunica non solo l'alloggio, ma anche il valore simbolico e relazionale del soggiorno.

Oggi, l'offerta italiana presente su Airbnb si caratterizza per un'elevata varietà, che riflette il patrimonio immobiliare del Paese. Accanto agli appartamenti urbani più tradizionali, la piattaforma propone soluzioni abitative uniche, come trulli, casali ristrutturati e dimore storiche. In un contesto così ampio e competitivo, la capacità di differenziarsi assume un ruolo centrale. Come evidenziato nel paragrafo precedente, la qualità dell'annuncio e l'attenzione ai dettagli rappresentano quindi non semplici elementi accessori, ma vere e proprie leve strategiche per emergere in un mercato ormai maturo e caratterizzato da un'elevata offerta (Guttentag et al., 2015).

CAPITOLO III – Background teorico e obiettivo di ricerca

Questo capitolo si propone di organizzare e analizzare i principali contributi della letteratura volti a comprendere i meccanismi di creazione del valore nella sharing economy. L'analisi si concentra sul ruolo delle diverse caratteristiche dell'offerta e sul modo in cui queste influenzano i risultati, considerando elementi quali le caratteristiche fisiche dell'alloggio, la reputazione dell'host e le modalità di presentazione dell'annuncio. Attraverso questa revisione della letteratura sarà inoltre possibile individuare gli aspetti ancora poco approfonditi e, a partire da essi, definire gli obiettivi della ricerca sviluppata nei capitoli successivi.

3.1 Analisi della letteratura

Per comprendere i meccanismi che definiscono il valore economico sulle piattaforme digitali, la letteratura accademica fa riferimento alla Teoria dei Prezzi Edonici introdotta dal celebre lavoro di Rosen (1974).

Questo modello è stato sviluppato per analizzare mercati in cui i beni non sono omogenei, ma differiscono per una serie di caratteristiche, come nel caso delle abitazioni. L'idea centrale proposta da Rosen è che un alloggio non possieda un valore unico e indivisibile. Il prezzo osservato sul mercato riflette piuttosto il contributo delle singole caratteristiche che lo compongono. In termini concreti, quando un utente prenota un alloggio su Airbnb, non acquista semplicemente un posto dove dormire, ma un insieme di attributi. Elementi come la vista, la presenza del Wi-Fi o la disponibilità dell'host contribuiscono ciascuno alla formazione del prezzo finale attraverso un valore specifico, definito shadow price. Tali valori non sono osservabili separatamente, ma risultano incorporati nella tariffa complessiva che l'ospite è disposto a pagare.

Nel contesto di Airbnb, questo approccio si rivela particolarmente efficace. A differenza delle catene alberghiere, dove le stanze sono spesso standardizzate e identiche, ogni annuncio sulla piattaforma rappresenta un pezzo unico. Di conseguenza, il prezzo finale nasce dalla combinazione di diversi elementi.

L'utilità di questa teoria risiede nella possibilità di analizzare in modo sistematico la formazione del prezzo, consentendo, attraverso strumenti statistici, di stimare il contributo delle singole caratteristiche dell'alloggio. In questo modo è possibile valutare il valore associato a specifici attributi, come la presenza dell'aria condizionata o lo status di Superhost, mantenendo costanti tutte le altre condizioni. A partire da questo approccio,

numerosi studi recenti hanno esaminato il mercato per individuare quali caratteristiche incidono maggiormente sui risultati.

3.1.1 Gli attributi strutturali e la localizzazione

L'analisi della letteratura si è focalizzata inizialmente sugli attributi fisici dell'offerta. È interessante notare come, nonostante la novità rappresentata dalla piattaforma, i driver economici tradizionali continuano a esercitare un'influenza decisiva. In questo senso, localizzazione e caratteristiche fisiche dell'immobile rappresentano ancora oggi i principali fattori alla base della valutazione iniziale dell'alloggio.

La localizzazione rappresenta un fattore rilevante nella determinazione del prezzo, sebbene non agisca in modo uniforme in tutti i contesti. La letteratura evidenzia una relazione inversa tra distanza dal centro e valore dell'alloggio e all'aumentare della distanza dai principali poli turistici, il prezzo tende a ridursi (Gibbs et al., 2018). Tuttavia, analizzando sia contesti urbani sia aree rurali, il valore della posizione dipende dal contesto. Nelle città la prossimità ai servizi e alle attrazioni urbane risulta determinante, mentre in aree alpine o rurali il prezzo è maggiormente influenzato da caratteristiche ambientali, come la vista o la tranquillità. Ne consegue che il significato di "buona posizione" varia in funzione del contesto geografico e delle motivazioni del viaggio (Falk et al., 2019).

Accanto alla localizzazione, le caratteristiche fisiche dell'immobile, come il livello di pulizia, il numero di bagni e la tipologia di alloggio, rappresentano le determinanti principali nella definizione del prezzo di base. La rilevanza di questi fattori è confermata anche da studi che utilizzano algoritmi avanzati di machine learning, come Random Forest e alberi decisionali. Applicando tali modelli a un ampio dataset di città statunitensi, si mostra che le variabili strutturali esercitano un peso statistico superiore rispetto ad altre categorie di variabili nella determinazione della fascia di prezzo iniziale (Chattopadhyay e Mitra, 2019)

Un ruolo rilevante è svolto anche dai servizi accessori offerti dall'alloggio. La letteratura distingue generalmente tra servizi considerati standard, come il Wi-Fi o una cucina di base, la cui presenza è ormai data per scontata e incide poco sulla determinazione del prezzo, e servizi differenzianti. La presenza esplicita di dotazioni meno comuni o percepite come di maggiore qualità, quali jacuzzi, sauna, parcheggio privato o aria condizionata, consente invece all'host di applicare un prezzo più elevato, distinguendosi dalla concorrenza diretta (Falk et al., 2019).

3.1.2 Profilo dell'Host

La letteratura più recente mostra come la costruzione della fiducia non dipenda esclusivamente dalle valutazioni numeriche lasciate dagli ospiti. In questa direzione, lo studio di ter Huurne et al. (2021) amplia l'analisi concentrandosi sul ruolo del profilo personale dell'host, con riferimento alla sezione descrittiva "Chi sono", spesso considerata marginale. Gli autori evidenziano come questa componente contribuisca alla formazione di una fiducia basata sulla benevolenza, fondata sulla percezione delle intenzioni e dell'affidabilità personale dell'host.

I risultati mostrano che gli utenti dedicano attenzione a tali descrizioni, soprattutto quando sono caratterizzate da un linguaggio emotivo positivo e da un certo grado di auto-rivelazione, come la condivisione di interessi personali e abitudini. Questi elementi rendono l'interazione più umana e riducono la distanza psicologica tra host e potenziale guest, aumentando la probabilità di prenotazione già nelle fasi iniziali del processo decisionale, talvolta prima ancora della consultazione delle recensioni relative all'alloggio.

3.1.3 La qualità della presentazione e l'analisi linguistica

È in questo ambito che la ricerca accademica ha registrato i progressi più rilevanti negli ultimi anni. In un contesto in cui l'interazione tra host e guest è interamente mediata dalla piattaforma digitale, il testo dell'annuncio assume un ruolo centrale e non si limita a descrivere l'alloggio, ma diventa uno strumento decisivo nel processo di scelta.

Koh et al. (2021), attraverso il loro studio, introducono nel settore turistico la Teoria dell'Aspettativa Linguistica. Analizzando oltre 110.000 annunci in sette Paesi, gli autori mostrano che l'uso di specifiche categorie lessicali, come un linguaggio maggiormente emotivo o più razionale, nel titolo dell'annuncio ha un impatto diretto sui ricavi. Un risultato particolarmente rilevante riguarda la dimensione culturale con l'efficacia delle scelte linguistiche che varia in base al contesto nazionale, suggerendo che strategie comunicative efficaci per alcuni mercati, come uno stile più diretto ed entusiastico per i turisti statunitensi, possono non esserlo in altri contesti, come quello giapponese.

Yang et al. (2022) applicano la teoria dell'Attenzione Selettiva ai titoli degli annunci Airbnb, evidenziando il loro ruolo nel catturare l'interesse degli utenti durante la fase di esplorazione. Considerato che la consultazione delle liste avviene spesso in modo rapido e superficiale, il titolo rappresenta il primo elemento in grado di attivare l'attenzione

dell'utente. Attraverso un'analisi che combina elementi linguistici e dati spaziali, gli autori mostrano come l'uso di specifiche parole chiave, in particolare quelle che richiamano l'atmosfera dell'alloggio o la vicinanza a punti di interesse, aumenti la probabilità di suscitare la curiosità e di indurre l'utente a cliccare sull'annuncio.

La letteratura recente si è interrogata sull'efficacia delle diverse strategie comunicative adottate dagli host nelle descrizioni degli annunci. Un contributo rilevante mette a confronto l'impatto di descrizioni fortemente positive, caratterizzate da un linguaggio promozionale che enfatizza i punti di forza dell'alloggio, e descrizioni più oneste, che includono anche eventuali limiti o aspetti meno favorevoli (Sun et al., 2024). I risultati evidenziano l'esistenza di un trade-off strategico perché un linguaggio positivo tende ad aumentare le prenotazioni nel breve periodo, grazie a un effetto persuasivo immediato, ma può generare aspettative eccessive. Diversamente, una comunicazione più trasparente contribuisce a costruire una fiducia più solida, favorendo una maggiore soddisfazione nel lungo periodo e valutazioni post-soggiorno più positive. Queste evidenze confermano che le scelte linguistiche non hanno una funzione meramente descrittiva, ma rappresentano una leva economica in grado di influenzare le aspettative degli utenti e i risultati di business.

3.2 I Gap nella letteratura esistente

Nonostante la crescente attenzione della letteratura verso il ruolo della comunicazione nelle piattaforme di sharing economy, emergono alcuni limiti ricorrenti che motivano la presente ricerca.

In primo luogo, si osserva un gap geografico e culturale. La maggior parte degli studi empirici sull'analisi linguistica degli annunci Airbnb si concentra su mercati asiatici o nordamericani, mentre il contesto europeo, e più nello specifico quello italiano, pur rappresentando una delle principali destinazioni turistiche a livello globale e uno dei mercati più rilevanti per la piattaforma, è stato finora scarsamente indagato sotto il profilo testuale. Questa limitazione della letteratura esistente appare significativa se si considera che il mercato italiano presenta caratteristiche distintive in termini di struttura dell'offerta, motivazioni di viaggio e aspettative degli ospiti, in grado di influenzare l'efficacia delle strategie comunicative. L'assenza di evidenze specifiche su questo contesto limita la generalizzabilità dei risultati esistenti.

Un secondo limite riguarda la dimensione testuale analizzata. Gran parte della letteratura si concentra sui titoli degli annunci o su brevi estratti testuali, privilegiando elementi di

immediata visibilità. Tuttavia, la decisione di prenotazione si costruisce attraverso un processo più articolato, in cui la descrizione completa dell'alloggio svolge un ruolo centrale nella riduzione dell'asimmetria informativa e nella costruzione della fiducia. La limitata attenzione riservata ai testi estesi comporta una comprensione parziale del ruolo effettivo del linguaggio nel determinare le performance economiche.

Un'ulteriore limitazione riguarda l'approccio metodologico all'analisi linguistica. Molti studi adottano dizionari standardizzati e categorie semantiche generaliste, spesso sviluppate in ambito psicologico, senza interrogarsi sulla loro adeguatezza rispetto a contesti economici specifici come il mercato degli affitti di breve periodo. Questo aspetto solleva interrogativi sulla capacità di tali strumenti di cogliere segnali linguistici realmente rilevanti per la formazione del valore economico.

Infine, si osserva un gap temporale legato al periodo di osservazione. Numerose analisi si basano su dati antecedenti alla pandemia, ma il contesto turistico post-Covid presenta caratteristiche nuove, con un possibile cambiamento delle preferenze dei viaggiatori verso aspetti quali la privacy, la pulizia, la disponibilità di spazi esterni o la possibilità di lavorare da remoto. Considerata la natura dinamica delle piattaforme digitali, risulta necessario verificare se le relazioni individuate in contesti precedenti mantengano la propria validità nel mercato post-pandemico.

3.3 Obiettivi della ricerca

Alla luce dell'analisi della letteratura e dei gap emersi nei paragrafi precedenti, questo studio approfondisce il ruolo della comunicazione testuale nelle performance economiche degli annunci Airbnb, concentrandosi in particolare sul contenuto delle descrizioni scritte dagli host. In un contesto digitale in cui l'interazione tra domanda e offerta è mediata quasi esclusivamente dalla piattaforma, il testo dell'annuncio rappresenta uno degli strumenti principali attraverso cui l'host può influenzare la percezione dell'alloggio e orientare le decisioni dei potenziali ospiti.

L'obiettivo della ricerca è comprendere se e in che misura le caratteristiche linguistiche delle descrizioni degli annunci siano associate a differenze nelle principali metriche di performance economica, una volta tenuto conto delle determinanti strutturali, reputazionali e di localizzazione. Il testo viene quindi interpretato non come un semplice elemento descrittivo, ma come una componente strategica dell'offerta, in grado di contribuire al posizionamento dell'annuncio all'interno di un mercato altamente competitivo.

In questa prospettiva, la comunicazione adottata dagli host viene analizzata attraverso diverse dimensioni linguistiche. Da un lato, vengono considerate le caratteristiche formali del testo, come la lunghezza e la struttura linguistica, per valutare se un maggiore sforzo comunicativo si traduca in un vantaggio misurabile in termini di risultati economici. Dall'altro lato, l'attenzione si concentra sullo stile comunicativo e sul contenuto semantico delle descrizioni, con l'obiettivo di osservare come il tono del messaggio e i temi enfatizzati possano incidere sulla percezione dell'offerta da parte degli utenti.

Il lavoro adotta un approccio empirico che integra strumenti di analisi testuale e modelli econometrici, consentendo di trasformare informazioni qualitative in variabili quantitative e di studiarne l'associazione con le performance degli annunci. Questo metodo permette di esplorare il legame tra linguaggio e risultati di mercato in modo sistematico, senza presupporre relazioni causali a priori, ma lasciando che siano i dati a fornire indicazioni sull'efficacia delle diverse strategie comunicative.

Inoltre, l'analisi si concentra sul contesto specifico del mercato italiano e del periodo post-pandemico, contribuendo a colmare un vuoto nella letteratura esistente, che si è finora concentrata prevalentemente su altri contesti geografici o su periodi precedenti. In questo senso, l'obiettivo ultimo della tesi è offrire una lettura empiricamente fondata del ruolo della comunicazione testuale su Airbnb, fornendo evidenze utili sia per l'avanzamento della ricerca accademica sia per una migliore comprensione delle dinamiche competitive che caratterizzano le piattaforme di sharing economy.

CAPITOLO IV – Metodologia

Dopo aver definito il quadro teorico e gli obiettivi di ricerca nel capitolo precedente, l'attenzione si sposta ora sulla fase empirica del lavoro. L'obiettivo di questa sezione è mostrare il procedimento seguito per verificare se e come le parole utilizzate dagli host nella descrizione del loro alloggio, influenzino le performance economiche su Airbnb.

L'analisi si sviluppa a partire da un dataset tradizionale, successivamente ampliato grazie all'integrazione di nuove informazioni testuali. L'utilizzo di tecnologie come il Web Scraping e l'elaborazione del linguaggio naturale (NLP) ha permesso di trasformare elementi qualitativi, quali le descrizioni testuali, in dati quantitativi misurabili, rendendo possibile l'analisi statistica.

4.1 Origine dei dati e costruzione del campione

4.1.1 Analisi del dataset iniziale e variabili strutturali

L'analisi si basa su un ampio dataset, composto inizialmente da 10.946.213 osservazioni. I dati coprono un arco temporale di sei anni, dal 2019 al 2024, e riguardano gli alloggi Airbnb di tre paesi europei: Italia, Spagna e Francia.

I dati sono organizzati secondo una dimensione temporale, risultando distribuiti su più anni e consentendo un'analisi dettagliata che integra le performance di mercato con le caratteristiche fisiche di ciascun alloggio.

Il dataset originario è composto da 43 variabili, relative alle caratteristiche strutturali degli alloggi, alle performance economiche, alla reputazione dell'host, alle policy di prenotazione e alla localizzazione geografica.

Più nello specifico, le variabili sono le seguenti:

1. Variabili di performance economica: rappresentano gli indicatori di successo finanziario utilizzati per stimare i modelli.

- Revenue (USD) / RevPAN (Revenue per Available Night): il fatturato totale e il ricavo per notte disponibile, indicatori sintetici della redditività.

$$\text{RevPAN} = \frac{\text{Revenue}}{\text{Reservation Days} + \text{Available Days}}$$

- ADR (Average Daily Rate): il prezzo medio giornaliero effettivo.

$$\text{Average Daily rate} = \frac{\text{Revenue}}{\text{Reservation Days}}$$

- Occupationrate: il tasso di occupazione, calcolato rapportando i giorni prenotati (Reservation Days) ai giorni disponibili (Available Days).

$$\text{Occupation Rate} = \frac{\text{Reservation Days}}{\text{Reservation Days} + \text{Available Days}}$$

- Number of Reservations: il volume totale di prenotazioni ricevute.
 - Reservationdays: numero di giorni effettivamente prenotati.
 - Availabledays: numero di giorni in cui l'alloggio è risultato disponibile.
2. Variabili strutturali dell'alloggio: definiscono le caratteristiche fisiche dell'immobile.
- Property Type / prop_type: tipologia dell'immobile (es. appartamento, casa indipendente).
 - Listing Type: tipologia di affitto (intero alloggio, stanza privata, stanza condivisa).
 - Bedrooms / Bathrooms: il numero di camere da letto e bagni.
 - Max Guests: la capacità ricettiva massima.
 - Number of Photos: il numero di fotografie caricate nell'annuncio.
 - Amenities: i servizi offerti (es. Wi-Fi, cucina, AC).
 - Listing Title / Length Title: il titolo e la lunghezza del titolo dell'annuncio.
3. Variabili di reputazione e profilo dell'host: misurano l'affidabilità e l'esperienza dell'host.
- Overall Rating: il punteggio medio delle recensioni.
 - Number of Reviews: il numero totale di recensioni ricevute.
 - Superhost: variabile dummy che indica il possesso del badge di qualità.
 - Response Rate: la reattività dell'host nelle risposte.
 - Prof Host: indicatore per distinguere gli host professionisti, che possiedono più alloggi, da quelli occasionali.
4. Variabili di policy e gestione: riflettono le scelte strategiche dell'host.
- Cancellation Policy / canc_policy: la flessibilità delle politiche di cancellazione.
 - Minimum Stay: il soggiorno minimo richiesto.
 - Instantbookenabled / instantbook: la possibilità di prenotazione immediata senza approvazione preventiva.
 - Cleaningfeusd / securitydepositusd: costi accessori come le spese di pulizia e il deposito cauzionale.
 - Publishednightlyrateusd / publishedweeklyrateusd / publishedmonthlyrateusd: prezzi pubblicati per le diverse durate di soggiorno.

5. Variabili identificative e temporali

- Year: variabile temporale per l'identificazione dell'anno di osservazione nel panel.
- Identificativi (ID): codici univoci per l'identificazione della proprietà (propertyid) e dell'host (airbnbhostid).
- Createddate: data di creazione dell'annuncio sulla piattaforma.
- Listingage: anzianità dell'annuncio, calcolata come il tempo trascorso dalla data di creazione.

6. Variabili di localizzazione geografica

- Country: Paese in cui è localizzato l'alloggio.
- State: regione o area amministrativa di riferimento.
- City: città in cui si trova l'alloggio.
- Latitude/Longitude: coordinate geografiche dell'immobile.
- Urban Class: classificazione del contesto territoriale (urbano, suburbano, rurale).

4.1.2 Filtraggio dei dati e definizione del perimetro d'analisi

Considerando le grandi dimensioni del dataset originale, è stato necessario effettuare una prima selezione per definire meglio l'ambito della ricerca.

Dall'insieme totale dei dati, si è scelto di isolare esclusivamente gli alloggi situati in Italia e attivi nell'anno 2024. Questa decisione si basa su due motivazioni principali:

- Coerenza geografica: analizzare solo il mercato italiano permette di evitare le differenze economiche e normative che ci sarebbero state confrontando paesi diversi come Francia e Spagna.
- Attualità dei dati: concentrarsi sul 2024 consente di studiare la situazione di mercato più recente, escludendo gli anni precedenti influenzati dalle anomalie della pandemia.

A seguito di questo filtraggio, si è ottenuto un dataset composto da 422.379 osservazioni. Tuttavia, poiché l'acquisizione delle descrizioni testuali richiedeva una procedura tecnica specifica, non risultava applicabile a un numero così elevato di annunci. Di conseguenza, da questo gruppo si è estratto casualmente un campione di 2.004 osservazioni, che rappresentano la base dati finale su cui è stata svolta l'analisi testuale descritta nel capitolo successivo.

4.2 Integrazione dei dati testuali e analisi semantica

Il dataset ottenuto, pur essendo ricco di informazioni numeriche, non conteneva il testo degli annunci. Per colmare questa mancanza, è stato necessario implementare una procedura tecnica specifica sui 2.004 alloggi campionati.

4.2.1 Web Scraping

La prima fase operativa ha riguardato l'acquisizione integrale delle descrizioni testuali. A partire dai codici identificativi univoci (Property ID) del campione, sono stati generati i link specifici per accedere alla pagina web di ogni singolo alloggio su Airbnb. Successivamente, è stato utilizzato un algoritmo di Web Scraping automatizzato che, seguendo questi indirizzi, ha analizzato le pagine web dedicate.

Per ogni unità, lo script ha individuato la sezione "Descrizione dell'alloggio" ed estratto tutto il testo scritto dall'host. I dati grezzi sono stati salvati e organizzati all'interno di un file Excel, strutturato in modo da associare a ogni descrizione il rispettivo codice identificativo. Questa operazione ha permesso di trasformare un contenuto testuale in un'informazione quantitativa, rendendola trattabile ai fini statistici.

4.2.2 Analisi testuale e creazione dei nuovi indici

Una volta generato il file Excel contenente le descrizioni testuali, l'attenzione si è spostata sull'elaborazione statistica dei dati in R. In questa fase è stato sviluppato uno script specifico per effettuare l'analisi del testo (Text Mining) e la Sentiment Analysis. L'obiettivo principale era convertire le informazioni qualitative, le descrizioni degli annunci, in metriche quantitative utilizzabili all'interno del modello econometrico.

Come mostrato nell'estratto riportato nella Figura 4.1, il codice ha permesso di importare i dizionari lessicali, processare le descrizioni e calcolare le frequenze delle parole chiave.

```
# SCRIPT DI ANALISI TESTUALE AIRBNB (LIWC STYLE)
# 1. CARICAMENTO LIBRERIE
library(tidyverse)
library(tidytext)
library(SnowballC)
library(stringr)
library(readxl)
library(writexl)
# Imposta la memoria per gestire testi lunghi se necessario
options(stringsAsFactors = FALSE)
```

```

# 2. DEFINIZIONE DEL DIZIONARIO MASTER (ESTESO E OTTIMIZZATO)
# Creazione delle liste di parole chiave
dict_source <- tribble(
~category,      ~term,
# --- 1. AFFECTIVE (Emozioni) ---
"affective",    "bell", "affective", "stupend", "affective", "meravigl",
"affective",    "fantast", "affective", "piacevol", "affective", "incant",
"affective",    "ador", "affective", "magnific", "affective", "favol",
"affective",    "eccellent", "affective", "perfett", "affective",
"indimentic",
"affective",    "sogn", "affective", "magi", "affective", "romantic",
"affective",    "gioi", "affective", "accoglient", "affective", "emozion",
# --- 2. SOCIAL (Relazioni) ---
"social",       "famigl", "social", "amic", "social", "bambin",
"social",       "ospit", "social", "incontr", "social", "person",
"social",       "grupp", "social", "copp", "social", "condiv",
"social",       "viv", "social", "gent", "social", "insiem",
# --- 3. COGNITIVE (Logica e Necessità) ---
"cognitive",    "ideale", "cognitive", "soluzion", "cognitive", "perch",
"cognitive",    "pens", "cognitive", "bisogn", "cognitive", "consigl",
"cognitive",    "cerc", "cognitive", "scegl", "cognitive", "garant",
"cognitive",    "adatt", "cognitive", "necess", "cognitive", "conosc",
# --- 4. PERCEPTUAL (Sensi) ---
"perceptual",  "vist", "perceptual", "lumin", "perceptual", "silenz",
"perceptual",  "panoram", "perceptual", "solegg", "perceptual", "ari",
"perceptual",  "tranquill", "perceptual", "color", "perceptual", "rumor",
"perceptual",  "profum", "perceptual", "confort", "perceptual", "caldo",
# --- 5. STATUS (Ex Drives - Lusso e Esclusività) ---
"status",      "luss", "status", "super", "status", "design",
"status",      "esclusiv", "status", "prestig", "status", "unic",
"status",      "elegan", "status", "modern", "status", "fin",
"status",      "style", "status", "vip", "status", "top",
"status",      "qualit", "status", "livell",
# --- 6. RELATIVITY (Posizione e Movimento) ---
"relativity",  "centr", "relativity", "vicin", "relativity", "metr",
"relativity",  "stazion", "relativity", "distanz", "relativity", "minut",
"relativity",  "pass", "relativity", "raggiung", "relativity", "strateg",
"relativity",  "fermat", "relativity", "colleg", "relativity", "aeroporto",
"relativity",  "posizion", "relativity", "km", "relativity", "lontan",
# --- 7. LEISURE (Svago e Cultura) ---

```

```

"leisure",      "vacanz", "leisure", "relax", "leisure", "mare",
"leisure",      "spiagg", "leisure", "ristorant", "leisure", "bar",
"leisure",      "movida", "leisure", "muse", "leisure", "art",
"leisure",      "stori",  "leisure", "cultur", "leisure", "negoz",
"leisure",      "piscina", "leisure", "shopping", "leisure", "tour",
# --- 8. HOME (Amenities e Casa) ---
"home",         "cucin",  "home",   "bagno",  "home",   "lett",
"home",         "camer",  "home",   "divan",  "home",   "wifi",
"home",         "internet", "home",   "aria",   "home",   "condizion",
"home",         "lavatric", "home",   "tv",     "home",   "terrazz",
"home",         "balcon", "home",   "ascensor", "home",   "arred",
# --- 9. MONEY (Economia) ---
"money",        "prezz",  "money",  "econom", "money",  "scont",
"money",        "cost",   "money",  "euro",   "money",  "gratuit",
"money",        "offert", "money",  "convenient", "money",  "budget",
"money",        "tass",   "money",  "risparm", "money",  "tariff",
# --- 10. CLEANLINESS (Igiene - Nuovo Indice) ---
"cleanliness", "pul",    "cleanliness", "igien", "cleanliness", "sanif",
"cleanliness", "impeccabil", "cleanliness", "ordin", "cleanliness", "cur",
"cleanliness", "fresc",  "cleanliness", "perfett",
# --- 11. SAFETY (Sicurezza - Nuovo Indice) ---
"safety",      "sicur",  "safety", "tranquill", "safety", "custod",
"safety",      "telecamer", "safety", "allarm", "safety", "privat",
"safety",      "portier", "safety", "protett", "safety", "casafort",
# --- 12. POLITENESS (Cortesia Host - Nuovo Indice) ---
"politeness",  "benvenut", "politeness", "accogli", "politeness",
"disponibil",
"politeness",  "aiut",   "politeness", "gradit", "politeness", "prego",
"politeness",  "piacer", "politeness", "felic", "politeness", "gentil"
)
# Preparazione del dizionario: Appliciamo lo stemming alle parole chiave
# In questo modo "appartamento" nel testo matcha con "appart" nel dizionario
dict_stemmed <- dict_source %>%
mutate(word_stem = wordStem(term, language = "italian")) %>%
select(category, word_stem) %>%
distinct()
# Regex per i pronomi (Indici Strutturali)
pron_io  <- "\\b(io|me|mi|mio|mia|miei|mie)\\b"
pron_noi <- "\\b(noi|ci|nostro|nostra|nostri|nostre)\\b"
pron_tu  <- "\\b(tu|te|ti|tuo|tua|tuei|tue|voi|vi|vostro)\\b"

```

```

# 3. FUNZIONE DI ANALISI COMPLETA (STRUTTURA + CONTENUTO)
analizza_annunci <- function(dataset) {
  cat("--- Inizio analisi su", nrow(dataset), "annunci ---\n")
  # A. CALCOLO METRICHE STRUTTURALI (Grammatica e Forma)
  cat("1. Calcolo metriche strutturali (WC, Pronomi, Frasi)...\n")
  structural_metrics <- dataset %>%
  rowwise() %>%
  mutate(
  # Word Count
  WC = str_count(description, "\\w+"),
  # Evita divisione per zero se WC è 0
  WC = ifelse(WC == 0, 1, WC),
  # Numero frasi
  sentence_count = str_count(description, "[.!?]+") %>% max(1),
  # Words Per Sentence
  WPS = WC / sentence_count,
  # Parole lunghe > 6 lettere (Complessità)
  long_words = length(unlist(str_extract_all(description, "\\w{7,}"))),
  Sixltr = (long_words / WC) * 100,
  # Pronomi (Percentuali)
  i_words = str_count(tolower(description), pron_io),
  we_words = str_count(tolower(description), pron_noi),
  you_words = str_count(tolower(description), pron_tu),
  i_percent = (i_words / WC) * 100,
  we_percent = (we_words / WC) * 100,
  you_percent = (you_words / WC) * 100
  ) %>%
  select(propertyid, WC, WPS, Sixltr, i_percent, we_percent, you_percent) %>%
  ungroup()
  # B. CALCOLO METRICHE SEMANTICHE (Dizionario)
  cat("2. Tokenizzazione e Stemming del testo...\n")
  tokens <- dataset %>%
  select(propertyid, description) %>%
  unnest_tokens(word, description) %>%
  mutate(word_stem = wordStem(word, language = "italian"))
  cat("3. Incrocio con il dizionario e calcolo punteggi...\n")
  matched_content <- tokens %>%

```

```

inner_join(dict_stemmed, by = "word_stem", relationship = "many-to-many")
%>%
count(propertyid, category)
# Uniamo i conteggi con il WC totale per avere le percentuali
content_metrics <- matched_content %>%
left_join(structural_metrics %>% select(propertyid, WC), by = "propertyid")
%>%
mutate(percentage = (n / WC) * 100) %>%
select(-n, -WC) %>%
# Trasforma da formato lungo a largo (colonne per ogni categoria)
pivot_wider(names_from = category, values_from = percentage, values_fill =
0)
# C. UNIONE FINALE
cat("4. Unione dati finali...\n")
final_dataset <- dataset %>%
left_join(structural_metrics, by = "propertyid") %>%
left_join(content_metrics, by = "propertyid") %>%
# Sostituiamo i NA con 0 nelle colonne delle categorie (chi non ha quelle
parole ha 0%)
mutate(across(where(is.numeric), ~replace_na(., 0))) %>%
# Arrotondiamo a 2 decimali per pulizia
mutate(across(where(is.numeric) & !propertyid, ~round(., 2)))
return(final_dataset)
}

# 4. ESECUZIONE (IMPORT -> ANALISI -> EXPORT)
# Definisci il percorso file
# Se il tuo Mac è in italiano prova "Scrivania", se è in inglese "Desktop"
file_path <- "~/Desktop/File_tesi/Descrizioni_finali_2024.xlsx"
# Oppure:
# file_path <- "~/Scrivania/File_tesi/Descrizioni_finali_2024.xlsx"
# Controllo esistenza file
if(file.exists(file_path)) {
# 1. Carica Excel
cat("Caricamento file:", file_path, "\n")
raw_data <- read_excel(file_path)
# Rinomina colonne per sicurezza (in caso di maiuscole/minuscole diverse)
# Assumiamo colonna 1 = ID, colonna 2 = Descrizione
colnames(raw_data)[1] <- "propertyid"
colnames(raw_data)[2] <- "description"
}

```

```

# Pulizia base: rimuove righe senza descrizione
raw_data <- raw_data %>% filter(!is.na(description))
# 2. Lancia l'analisi
dataset_analizzato <- analizza_annunci(raw_data)
# 3. Salva risultato
output_filename <- "~/Desktop/File_tesi/Risultati_Airbnb_12-30-05.xlsx"
write_xlsx(dataset_analizzato, output_filename)
cat("\n--- OPERAZIONE COMPLETATA CON SUCCESSO! ---\n")
cat("File salvato in:", output_filename, "\n")
cat("Totale annunci analizzati:", nrow(dataset_analizzato), "\n")
# Mostra prime righe
print(head(dataset_analizzato))
} else {
stop("ERRORE: Il file non è stato trovato. Controlla che la cartella
'File_tesi' esista e contenga 'Descrizioni_finali_2024.xlsx'.")
}

```

Figura 3: Script R utilizzato per l'analisi semantica e il calcolo degli indici linguistici

L'elaborazione eseguita dal software ha seguito due passaggi logici:

1. Pulizia dei dati (Pre-processing): i testi grezzi sono stati ripuliti da caratteri speciali e formattazioni errate per uniformare il contenuto.
2. Calcolo delle variabili: utilizzando dizionari lessicali validati (basati sulla metodologia LIWC-Linguistic Inquiry and Word Count), lo script ha conteggiato la frequenza di specifiche categorie di parole per ogni annuncio.

Da questa procedura sono stati generati 18 nuovi indici, che costituiscono le variabili indipendenti centrali della ricerca. Le principali metriche ottenute sono state classificate nei seguenti gruppi:

Metriche di complessità del testo:

- Words Count: il numero totale di parole utilizzate nella descrizione.
- Words per sentence: il numero medio di parole per frase, indicatore della complessità sintattica.
- Complexity: la percentuale di parole con più di 6 lettere. Questa variabile funge da indice di complessità del linguaggio.

Metriche relazionali:

- Guest orientation: la percentuale di pronomi di seconda persona (you, your), che indica l'orientamento verso il cliente.

- Host orientation: la percentuale di pronomi di prima persona singolare (I, me, my), segnale di un focus sull'host.
- Professional orientation: la percentuale di pronomi di prima persona plurale (we, our), che suggerisce una gestione di gruppo o aziendale.

Metriche semantiche:

- Affective: parole legate alla sfera emotiva e sentimentale.
- Cognitive: termini che indicano processi di pensiero e ragionamento.
- Perceptual: parole riferite ai sensi e alla percezione fisica.
- Social: riferimenti all'interazione umana, alla famiglia e agli amici.
- Home: termini specifici relativi all'ambiente domestico.
- Leisure: parole associate al tempo libero, allo svago e al relax.
- Safety: riferimenti alla sicurezza e alla protezione.
- Cleanliness: termini legati all'igiene e alla pulizia.
- Money: riferimenti a questioni finanziarie o economiche.
- Status: parole che denotano prestigio, esclusività o rango sociale.
- Politeness: termini di cortesia e formalità.
- Relativity: parole legate al tempo e allo spazio.

4.3 Misure

La scelta delle variabili è stata guidata dalla letteratura prevalente in ambito hospitality e revenue management, con l'obiettivo di costruire un modello econometrico robusto in grado di isolare l'effetto delle caratteristiche testuali sulla performance.

4.3.1 Definizione e operazionalizzazione delle variabili

Le variabili del modello econometrico sono state classificate in quattro categorie: dipendenti, indipendenti, di controllo e di moderazione, a cui si aggiungono gli effetti fissi.

Variabili dipendenti

Al fine di valutare le performance degli alloggi sotto diverse dimensioni, in particolare finanziaria, di domanda e reputazionale, sono state selezionate cinque metriche, quasi tutte trasformate in logaritmo naturale per correggere la distribuzione dei dati:

- Log RevPAN: questa variabile rappresenta il logaritmo naturale del fatturato per notte disponibile. È la metrica di performance standard nel settore extra-

alberghiero, analoga al RevPAR (Revenue Per Available Room) utilizzato nell'industria alberghiera. Indicatori di questo tipo sono considerati i più robusti per valutare l'efficienza economica poiché sintetizzano sia la leva del prezzo che quella dell'occupazione (Sainaghi, 2010).

- Log ADR: rappresenta il logaritmo naturale del prezzo medio giornaliero. Questa variabile è utilizzata come indicatore del valore che il mercato attribuisce all'alloggio. Basandosi sulla teoria dei prezzi edonici, lo studio di Wang & Nicolau (2017) dimostra che l'ADR non dipende solo dalle caratteristiche fisiche, ma anche dai segnali informativi presenti nell'annuncio; per cui, questa metrica consente di verificare se una descrizione efficace riesce a giustificare un prezzo superiore.
- Occupancy rate: misura l'efficienza della domanda. Il tasso di occupazione è fondamentale per comprendere se le strategie comunicative riescono a convertire le visualizzazioni in prenotazioni effettive (Varma et al., 2016).
- Log overall rating: il punteggio medio delle recensioni è utilizzato come il principale indicatore della reputazione online e della fiducia dell'host. Nei mercati peer-to-peer, caratterizzati da incertezza, il sistema di rating funge da meccanismo fondamentale per la costruzione della fiducia, permettendo agli utenti di valutare l'affidabilità dell'host prima della transazione (Ert et al., 2016).
- Log number of reservation: rappresenta il logaritmo naturale del numero totale di prenotazioni e consente di osservare il volume delle transazioni, a differenza del tasso di occupazione. Questa variabile viene impiegata per analizzare l'impatto del livello di informazione personale fornita dall'host: descrizioni più dettagliate e orientate all'auto-descrizione contribuiscono ad accrescere la fiducia degli utenti e la probabilità di scelta dell'alloggio (Ma et al., 2017).

Variabili indipendenti

Come descritto nella sezione metodologica, le variabili indipendenti sono costituite dai 18 indici derivati dall'analisi testuale, che misurano lo sforzo informativo (WordCount, WPS, Complexity), lo stile relazionale (Host/Guest/Professional orientation) e il contenuto semantico (le 12 categorie come Leisure, Social, Affective ecc.).

- Word Count: la lunghezza della descrizione è interpretata come un segnale di ricchezza informativa. Descrizioni più lunghe riducono l'asimmetria informativa e aumentano la fiducia del consumatore (Schuckert et al., 2015).

- Words per sentence, Complexity: basandosi sulla metodologia LIWC, queste misure consentono di valutare la ricchezza del linguaggio utilizzato, elemento chiave nel determinare quanto l'host venga percepito come competente e professionale agli occhi del potenziale guest (Pennebaker et al., 2015).
- Host/Guest/Professional orientation: queste variabili sono incluse per cogliere differenze nello stile comunicativo adottato dagli host e nel modo in cui il messaggio è orientato verso il potenziale guest. La letteratura mostra come il focus relazionale del linguaggio influenzi il coinvolgimento dell'utente e la percezione dell'offerta, incidendo sul processo decisionale e sulle valutazioni degli annunci nei contesti digitali (Tussyadiah et al., 2018).
- Variabili tematiche (affective, leisure, home...): permettono di distinguere le caratteristiche dell'offerta, andando oltre la semplice conta delle parole. L'inclusione di queste variabili dimostra che la scelta di specifici temi semantici e parole chiave influenza in modo determinante l'attrattività dell'annuncio e la percezione del valore in diversi contesti culturali (Brochado et al., 2020).

Variabili di controllo

Per isolare l'effetto netto delle variabili testuali, il modello include controlli relativi alle caratteristiche dell'alloggio e alle policy dell'host.

- Variabili sulle caratteristiche fisiche: il numero di camere (Log Bedrooms), il numero di bagni (Log Bathrooms) e la capacità massima di ospiti (Max Guests). L'inclusione di queste variabili segue il modello dei prezzi edonici applicato al contesto Airbnb da Gibbs et al. (2018), i quali confermano che gli attributi fisici e la capacità ricettiva rappresentano le determinanti primarie per la formazione del prezzo e la generazione della domanda.
- Log number of photos: il numero di foto è incluso poiché le immagini riducono l'incertezza visiva e fungono da complemento alla descrizione testuale (Ert et al., 2016).
- Superhost: lo status di superhost è controllato in quanto agisce come un certificato di qualità dalla piattaforma, influenzando positivamente sia il prezzo che la domanda (Liang et al., 2017).
- Response rate: il tasso di risposta misura la percentuale di nuovi messaggi a cui l'host risponde entro 24 ore. Questa variabile è stata inserita per controllare la qualità del servizio nella fase di pre-prenotazione. La tempestività nella

comunicazione riduce i costi di transazione per l'utente e funge da segnale di affidabilità e presenza attiva dell'host sulla piattaforma (Wang & Nicolau, 2017).

- Variabili su policy e restrizioni: le variabili relative alla politica di cancellazione e al soggiorno minimo sono state incluse per controllare il livello di flessibilità delle condizioni di prenotazione. La rigidità di queste condizioni agisce come un fattore di rischio trasferito sull'ospite, influenzando negativamente l'utilità percepita e, di conseguenza, la disponibilità a pagare e la domanda (Wang & Nicolau, 2017).

Variabili di moderazione

Per testare le ipotesi sugli effetti contingenti, sono state utilizzate due variabili:

- Urban Class: una variabile categorica che distingue la localizzazione dell'alloggio (urbano vs rurale/periferico), per verificare se l'importanza della descrizione cambia in base al contesto geografico.
- Superhost: oltre al suo ruolo di controllo, questa variabile viene interagita con le variabili testuali per verificare se lo status di "Superhost" amplifica o riduce l'efficacia di determinati stili comunicativi.

Effetti fissi

Per assorbire l'eterogeneità non osservata legata alla localizzazione specifica e alla tipologia immobiliare, sono stati inseriti nel modello gli effetti fissi per città e tipo di proprietà.

Tipo di variabile	Costrutto	Operazionalizzazione	Fonte	Riferimenti a studi precedenti
Variabili dipendenti	Log RevPAN	Logaritmo naturale del fatturato per notte disponibile	AirDNA	Sainaghi (2010)
	Log ADR	Logaritmo naturale del prezzo medio giornaliero	AirDNA	Wang & Nicolau (2017)
	Occupation rate	Percentuale del tasso di occupazione, misura dell'efficienza nello sfruttamento della capacità ricettiva	AirDNA	Varma et al. (2016)
	Log Number of reservation	Logaritmo naturale del numero totale di prenotazioni ricevute nel periodo	AirDNA	Ma et al. (2017)
	Log Overall rating	Logaritmo naturale del punteggio medio delle recensioni	AirDNA	Ert et al. (2016)
Variabili indipendenti	Word Count	Conteggio del numero totale di parole	Script R	Schuckert et al. (2015)
	Word per sentences, Complexity	Percentuale di parole lunghe e lunghezza media delle frasi	Script R	Pennebaker et al. (2015)
	Host/Guest/Professional orientation	Frequenza pronomi personali	Script R	Tussyadiah et al. (2018)
	Semantiche	Frequenze relative a temi edonici e funzionali	Script R	Brochado et al. (2020)
Variabili di controllo	Caratteristiche fisiche	Variabili logaritmiche per la dimensione e capacità fisica dell'alloggio.	AirDNA	Gibbs et al. (2018)
	Number of photos	Logaritmo del numero di foto	AirDNA	Ert et al. (2016)
	Superhost	Variabile dummy uguale ad 1 se l'host ha il badge di Superhost, 0 se no	AirDNA	Liang et al. (2017)
	Response rate	Percentuale del tasso di risposta	AirDNA	Wang & Nicolau (2017)
	Policy e restrizioni	Variabili dummy che definiscono la presenza di vincoli contrattuali che influenzano la flessibilità	AirDNA	Wang & Nicolau (2017)
Variabili di moderazione	Urban class	Variabile categoriale per la tipologia di area urbana	AirDNA	Falk et al. (2019)
	Superhost	Variabile dummy uguale ad 1 se l'host ha il badge di Superhost, 0 se no	AirDNA	Liang et al. (2017)

Tabella 1: Operazionalizzazione delle variabili

4.3.2 Analisi descrittiva

Variabile	Count	Mean	Stand. Dev.	Min	Max
log_revpan	1279	3.663559	1.16792	0.1769307	7.105786
log_adr	1279	4.68544	0.6909047	2.844909	7.609862
occupationrate	2004	0.2787484	0.328118	0	1
log_numberofreservations	1279	0.9521945	0.7658903	0	3.044523
log_overallrating	1892	1.555719	0.0643767	0.9162908	1.609438
words count	2004	2.543418	2.067826	5	2.817
words per sentence	2004	2.203567	1.184527	4.75	189
complexity	2004	3.000916	4.673657	0	50
host orientation	2004	0.075524	0.3630919	0	9.09
guest orientation	2004	0.3088623	0.5472483	0	6.15
professional orientation	2004	0.3480689	0.6461488	0	8.33
affective	2004	0.3958782	0.7304044	0	20
cleanliness	2003	0.1970394	0.3763747	0	3.12
cognitive	2004	0.3965968	0.5755965	0	8.33
home	2004	3.929696	306.884	0	23.08
leisure	2004	1.069731	1.316794	0	16.67
money	2004	0.402006	0.7362045	0	10.26
perceptual	2004	1.051422	1.159541	0	20
politeness	2004	0.1116916	0.3175279	0	6.67
relativity	2004	1.677011	1.581261	0	18.18
safety	2004	0.2117216	0.4009733	0	4
social	2004	1.067041	1.093917	0	11.54
status	2004	0.3618014	0.6221884	0	12.5
log_bathrooms	1994	0.2663144	0.4520642	0	2.639057
log_bedrooms	1875	0.5226835	0.5409425	0	2.564949
max guests	2004	4.656188	2.747496	1	20
minimum stay	2004	4.486527	2.023843	1	500
response rate	1970	9.236751	2.191101	0	100
superhost	2004	0.3682635	0.4824538	0	1
cancellation policy	1952	2.927766	1.142985	1	5
urban_class	1982	193.441	0.8241292	1	3

Tabella 2: Statistiche descrittive

La Tabella 2 riporta le statistiche descrittive delle principali variabili utilizzate nell'analisi e offre una panoramica complessiva delle caratteristiche del campione considerato. Il dataset finale comprende 2.004 osservazioni; tuttavia, la numerosità campionaria varia tra le diverse variabili. In particolare, le metriche di performance economica trasformate in logaritmo (*Log RevPAN*, *Log ADR* e *Log Number of Reservations*) sono disponibili per 1.279 osservazioni. Questa riduzione è dovuta alla presenza di alloggi che non hanno registrato ricavi nel periodo analizzato, per i quali il logaritmo non risulta definibile. Una dinamica simile si osserva anche per le variabili

strutturali *Log Bedrooms* e *Log Bathrooms*, il cui numero di osservazioni è inferiore al totale a causa della presenza di monocali o di alloggi privi di bagni privati, registrati con valore pari a zero nella piattaforma.

Dal punto di vista delle performance, il tasso di occupazione medio si attesta intorno al 27,8%. Ciò evidenzia un utilizzo relativamente contenuto degli alloggi nel periodo considerato. La deviazione standard elevata (0,32) segnala una forte eterogeneità all'interno del campione, con situazioni molto diverse tra alloggi scarsamente utilizzati e annunci caratterizzati da livelli di occupazione decisamente più elevati. Questo dato è coerente con la natura flessibile e spesso non continuativa dell'attività di hosting su Airbnb. Le statistiche relative alle variabili testuali mettono in luce una forte diversità nelle modalità con cui gli host costruiscono le descrizioni degli annunci. Le differenze emergono sia nella lunghezza dei testi sia nel tipo di linguaggio utilizzato.

Le metriche semantiche presentano valori medi generalmente bassi, indicando che molte categorie tematiche compaiono solo sporadicamente all'interno delle descrizioni. Nel campione emergono quindi modalità comunicative molto diverse tra loro, che si inseriscono in un contesto già eterogeneo dal punto di vista strutturale ed economico e rappresentano la base per le analisi grafiche e per le stime econometriche successive.

Di seguito vengono descritte nel dettaglio alcune delle principali caratteristiche considerate nell'analisi.

Caratteristiche strutturali: Max guests

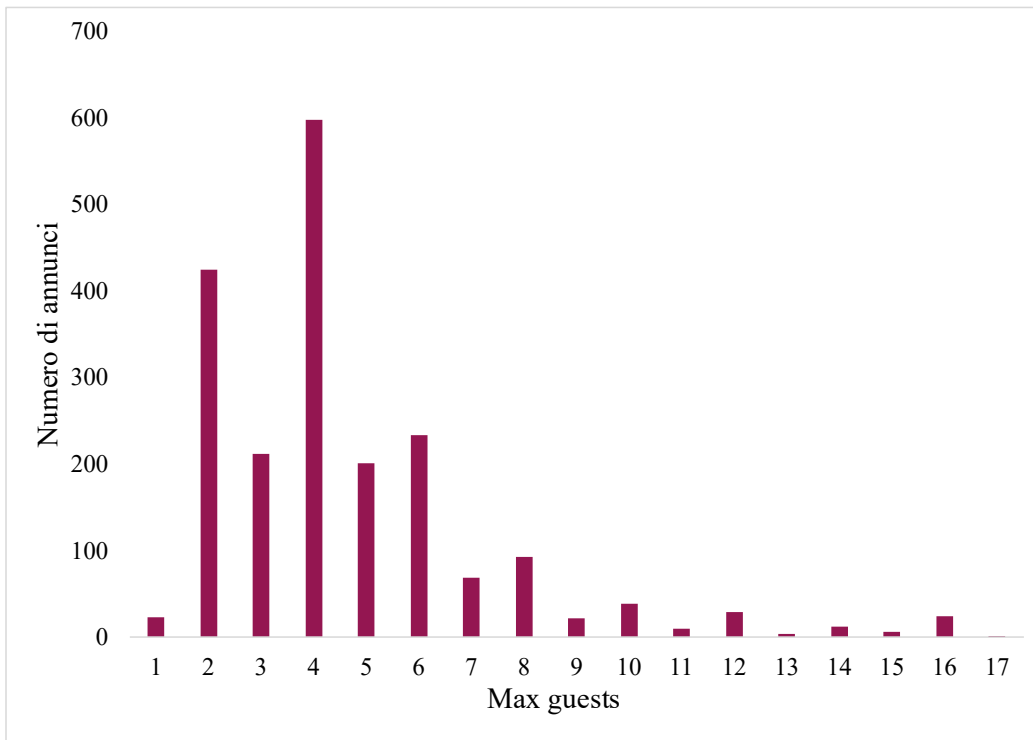


Figura 4: Capacità ricettiva degli alloggi

Il grafico in Figura 4 mostra la distribuzione della capacità ricettiva degli alloggi analizzati, misurata attraverso il numero massimo di ospiti. Il campione presenta una capacità media pari a circa 4,6 ospiti, evidenziando una prevalenza di appartamenti interi rispetto a soluzioni più piccole come stanze singole o alloggi pensati esclusivamente per coppie. La distribuzione risulta asimmetrica verso destra e se da un lato le soluzioni per due ospiti rappresentano una quota rilevante dell'offerta, dall'altro emerge una presenza consistente di alloggi in grado di ospitare famiglie e gruppi più numerosi, con capacità pari o superiore a quattro persone. Questo andamento evidenzia una buona eterogeneità dello stock immobiliare presente sulla piattaforma, coerente con la natura dell'offerta di Airbnb pensata per soddisfare esigenze di viaggio differenti.

Profilo reputazionale: Superhost

Per quanto riguarda il profilo reputazionale degli host, il campione analizzato mostra un livello di professionalità elevato, come suggerito anche dall'alto tasso medio di risposta. Tuttavia, lo status di Superhost non risulta distribuito in modo uniforme tra gli annunci.

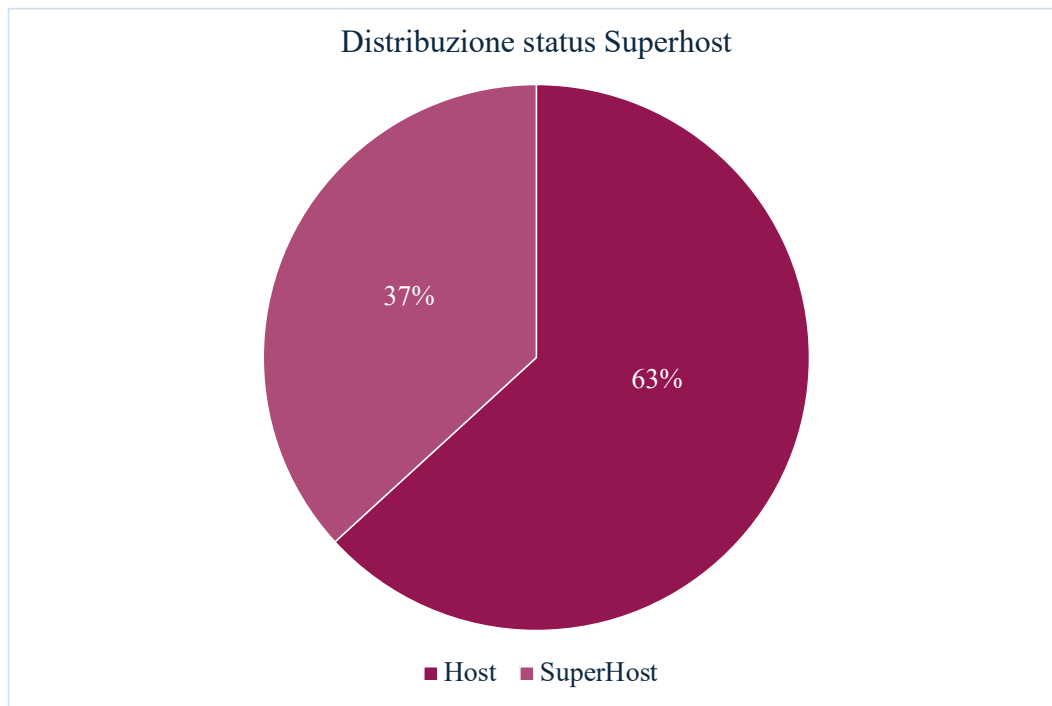


Figura 5: Distribuzione dello status di Superhost nel campione

Come mostrato in *Figura 5*, il 37% degli host possiede il badge di Superhost, mentre la maggioranza degli annunci (63%) ne è priva. Questa distribuzione indica che lo status di eccellenza rappresenta una risorsa relativamente rara all'interno del mercato analizzato. Proprio questa relativa scarsità rende il badge un segnale credibile di qualità e affidabilità, rafforzando l'idea che possa tradursi in un vantaggio competitivo e, potenzialmente, in grado di incidere sulle performance economiche.

Strategie linguistiche: Words count

Dall'analisi della lunghezza delle descrizioni emerge una forte eterogeneità nello sforzo comunicativo degli host. La descrizione media conta circa 254 parole, ma la variabilità è molto ampia: si passa da testi estremamente brevi, composti da poche parole, a descrizioni particolarmente dettagliate che superano le 2.800 parole.

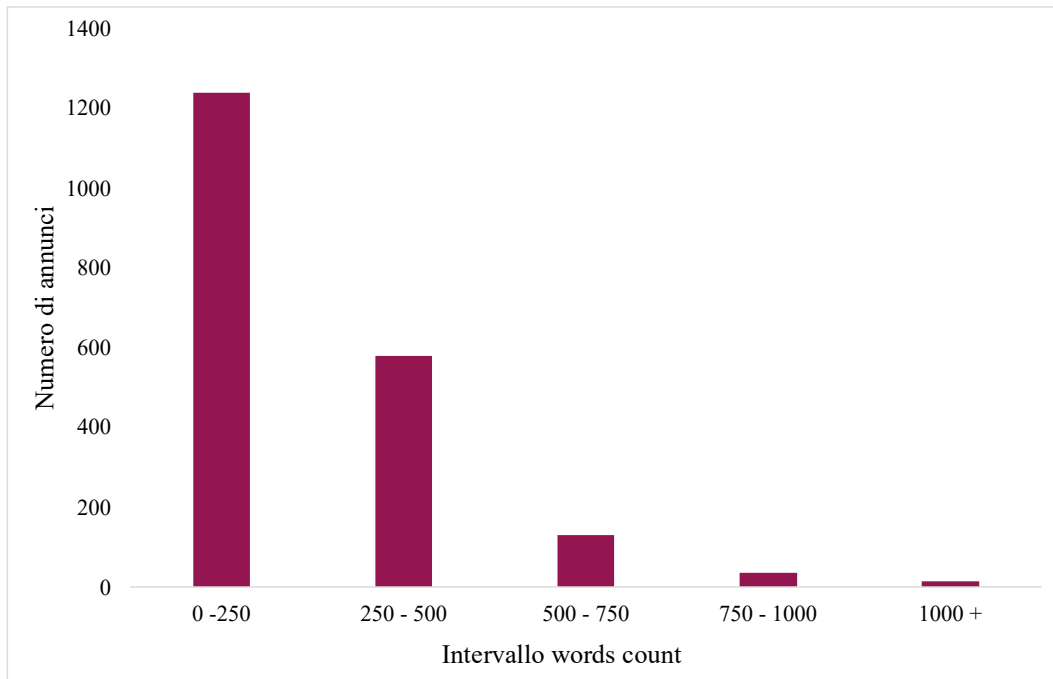


Figura 6: Lunghezza delle descrizioni

L'istogramma in *Figura 6* mostra che la distribuzione è fortemente asimmetrica positiva. La concentrazione maggiore di annunci si trova nella fascia tra le 150 e le 250 parole. Questo andamento suggerisce che la strategia comunicativa più diffusa sia orientata alla sintesi e alla trasmissione delle informazioni essenziali. Al contrario, la scelta di investire in testi lunghi e articolati appare limitata a un gruppo ristretto di host e può essere interpretata come una strategia di differenziazione rispetto allo stile comunicativo prevalente sulla piattaforma.

Analisi del contenuto: i temi dominanti

L'analisi del contenuto delle descrizioni consente di confrontare i principali temi su cui gli host concentrano la propria comunicazione. È stato, a tal proposito, effettuato un confronto tra le tematiche prevalenti nelle descrizioni.

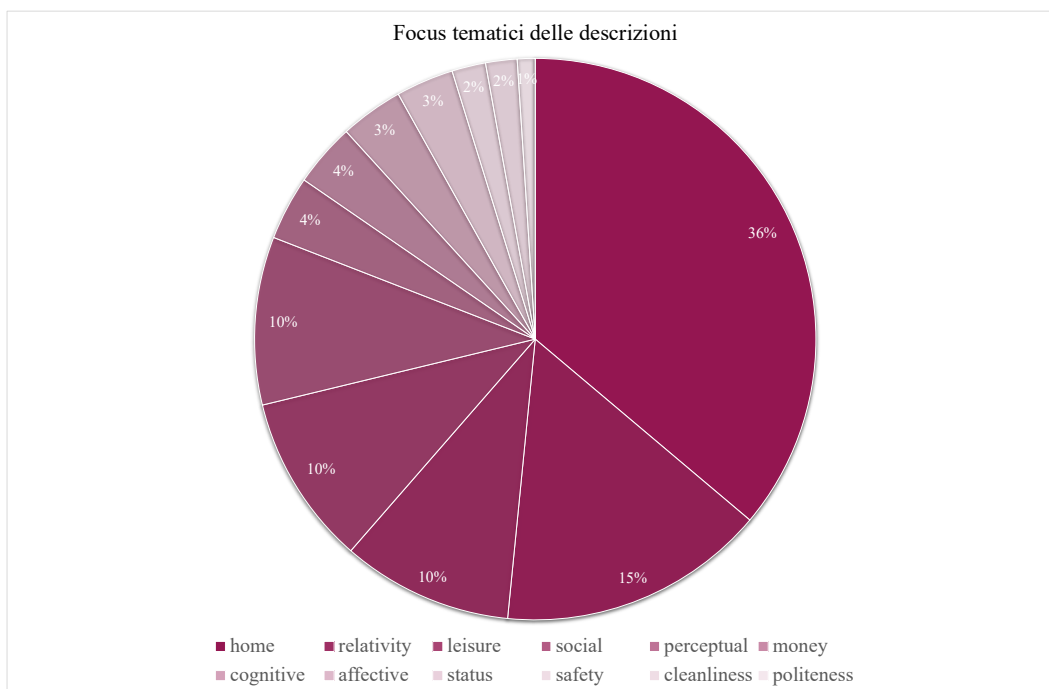


Figura 7: Confronto fra i focus tematici delle descrizioni

Il grafico in *Figura 7* mostra, infatti, la distribuzione relativa dei principali temi semantici presenti nelle descrizioni degli annunci. La componente più rilevante è rappresentata dalla dimensione *Home*, che assorbe da sola circa il 36% del contenuto testuale complessivo. Questo dato conferma che la comunicazione degli host è fortemente centrata sugli elementi fisici dell'alloggio e sulle dotazioni disponibili, come stanze, arredi e servizi domestici. Accanto a questa dimensione emergono, con un peso comunque significativo, i riferimenti alla *Relativity* (15%), legati alla posizione e alla vicinanza a punti di interesse, e un gruppo di temi con incidenza simile quali *Leisure*, *Social* e *Perceptual*, ciascuno intorno al 10%. Queste categorie richiamano aspetti legati all'esperienza di soggiorno, alle interazioni sociali e alle sensazioni associate all'alloggio, ma risultano complessivamente subordinate rispetto alla descrizione funzionale della casa. Le restanti dimensioni semantiche, come *Affective*, *Status*, *Money*, *Safety*, *Cleanliness* e *Politeness*, presentano un peso più contenuto e frammentato, suggerendo che tali registri comunicativi non vengano utilizzati in modo sistematico ma selettivo.

La distribuzione, quindi, evidenzia uno stile descrittivo prevalentemente informativo e orientato al prodotto, con un utilizzo più limitato di contenuti esperienziali ed emotivi, aprendo spazio a potenziali strategie di differenziazione basate sul contenuto della narrazione.

Caratteristiche di prezzo: Average daily rate

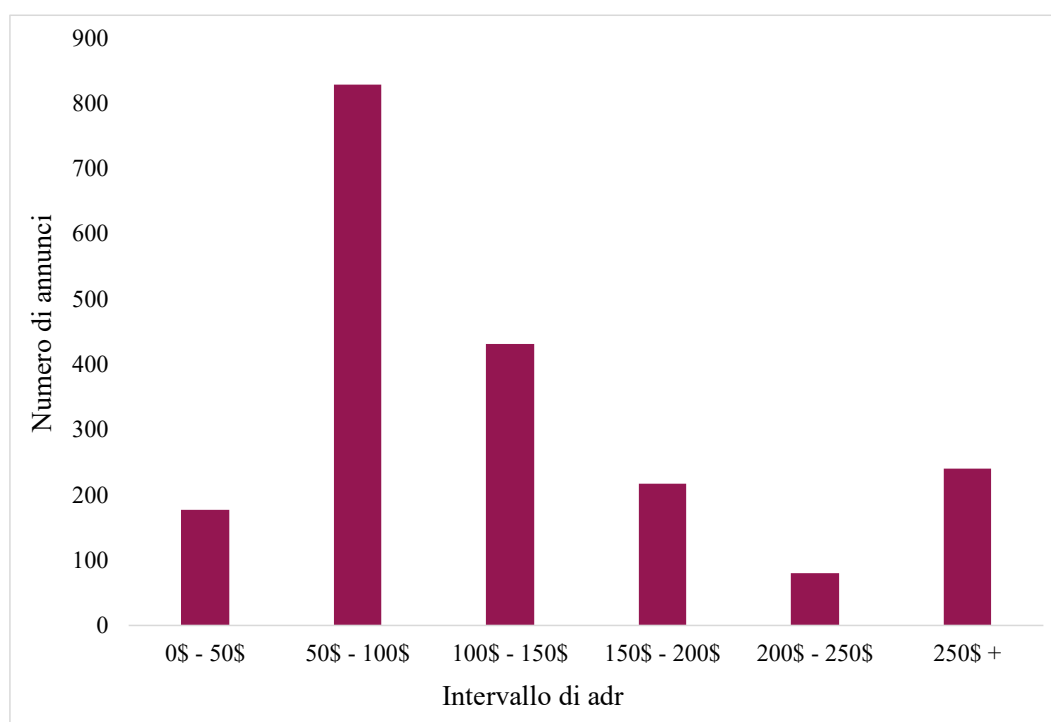


Figura 8: Distribuzione degli annunci per intervalli di prezzo medio giornaliero

Il grafico mostra la distribuzione degli annunci in base a intervalli di prezzo medio per notte. Come si osserva, la maggior parte degli alloggi si concentra nelle fasce di prezzo medio-basse, nello specifico tra i 50 e i 100 dollari a notte, che rappresentano l'intervallo più frequente nel campione. Al crescere del prezzo medio, il numero di annunci tende progressivamente a ridursi, segnalando una minore presenza di alloggi nelle fasce di prezzo più elevate. Tuttavia, la presenza non trascurabile di annunci con ADR superiore ai 250 dollari indica l'esistenza di un segmento di mercato orientato a soluzioni di fascia alta. Nel suo insieme la distribuzione evidenzia un mercato eterogeneo, caratterizzato da una forte concentrazione su livelli di prezzo accessibili ma con la coesistenza di soluzioni di fascia più alta. Questo indica che il prezzo medio per notte può essere utilizzato dagli host come strumento di posizionamento all'interno del mercato.

Caratteristiche di ricavo: RevPAN

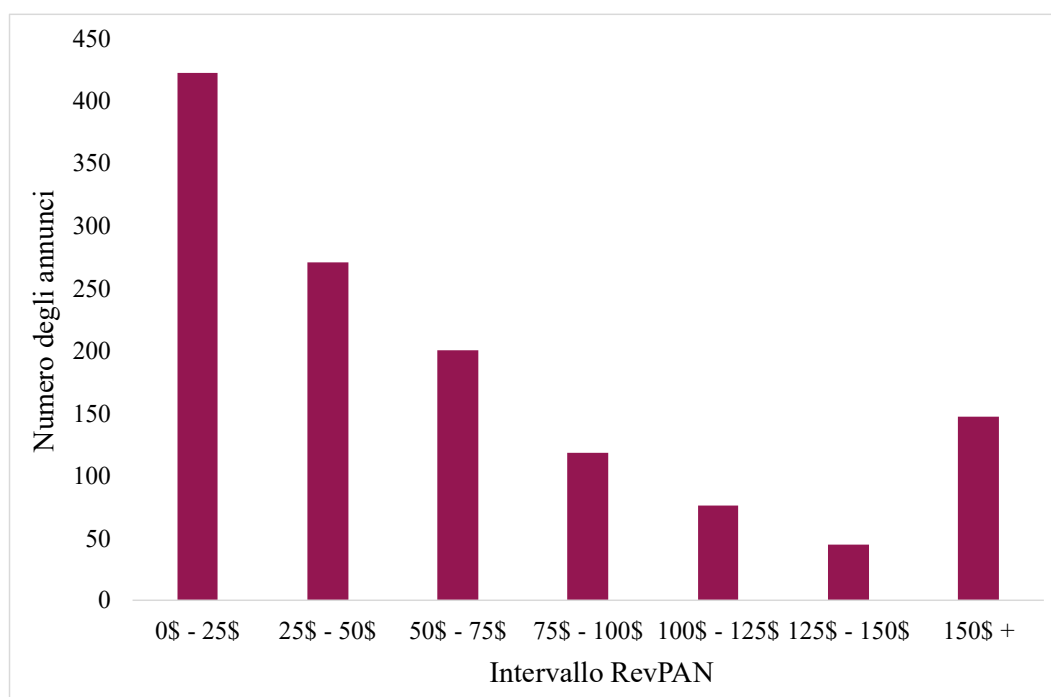


Figura 9: Distribuzione degli annunci per intervalli di ricavo medio per notte disponibile

La capacità degli alloggi di generare ricavi per notte disponibile, misurata attraverso il RevPAN, presenta una marcata dispersione all'interno del campione. La maggior parte degli annunci si concentra nelle fasce di RevPAN più basse, in particolare sotto i 50 dollari, indicando che, per molti host, la disponibilità dell'alloggio non si traduce automaticamente in ricavi elevati. Accanto a questa componente emerge un insieme più ristretto di annunci caratterizzati da valori di RevPAN sensibilmente più alti, che indicano una maggiore capacità di combinare efficacemente prezzo e utilizzo dell'alloggio. La coesistenza di questi profili suggerisce che le differenze di performance non dipendano da un unico fattore, ma da scelte gestionali e caratteristiche dell'annuncio che incidono congiuntamente sui risultati economici. In questo contesto, il RevPAN si conferma una misura particolarmente adatta a cogliere le differenze nell'efficienza di monetizzazione tra gli host.

Efficienza della domanda: Occupation rate

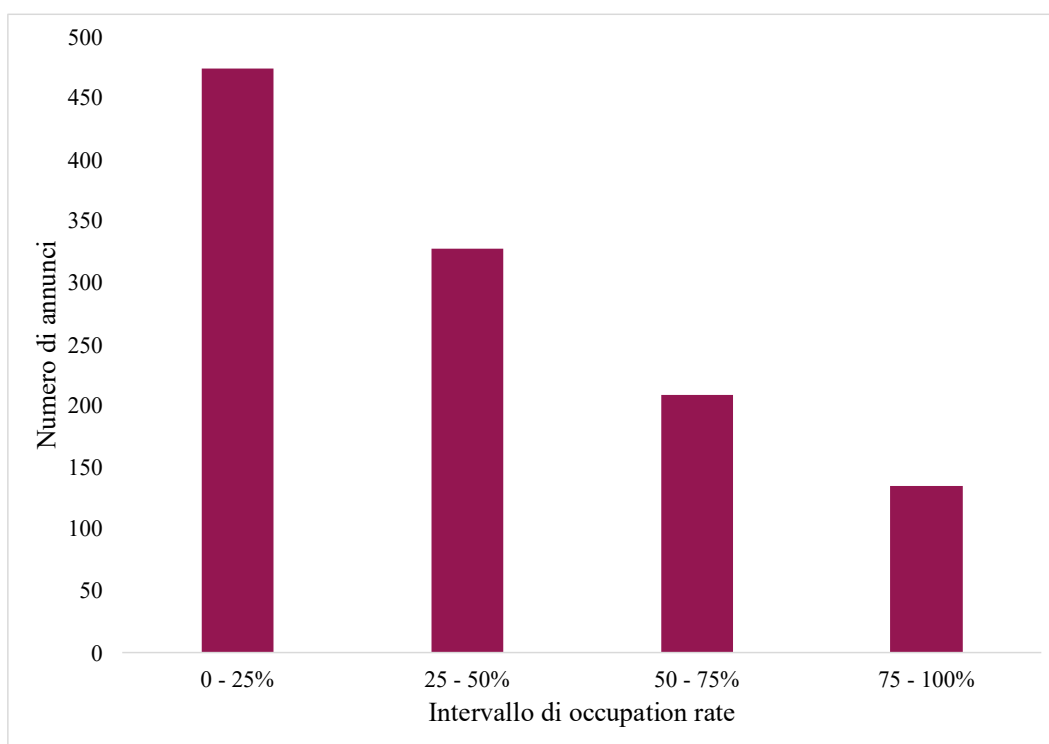


Figura 10: Distribuzione degli annunci per intervalli di occupation rate

Il grafico evidenzia una forte concentrazione di annunci nelle fasce di occupazione più basse, con una quota rilevante di alloggi che non supera il 25% di utilizzo nel periodo considerato. Questo dato suggerisce che, per molti host, l'attività di affitto su Airbnb non rappresenta una fonte di utilizzo continuativo dell'alloggio, ma piuttosto una scelta intermittente, legata a specifici momenti dell'anno o a esigenze personali. All'aumentare del tasso di occupazione, il numero di annunci si riduce progressivamente, indicando che mantenere livelli elevati di domanda nel tempo è una condizione meno frequente. Solo una parte limitata degli alloggi riesce infatti a collocarsi nelle fasce di occupazione più alte, superiori al 75%. Questa distribuzione mette in luce una forte differenziazione nelle capacità degli host di attrarre prenotazioni e di sfruttare la disponibilità dell'alloggio, rendendo il tasso di occupazione una misura particolarmente utile per cogliere le differenze di performance sul lato della domanda.

Caratteristiche dell'offerta: tipologia di alloggio

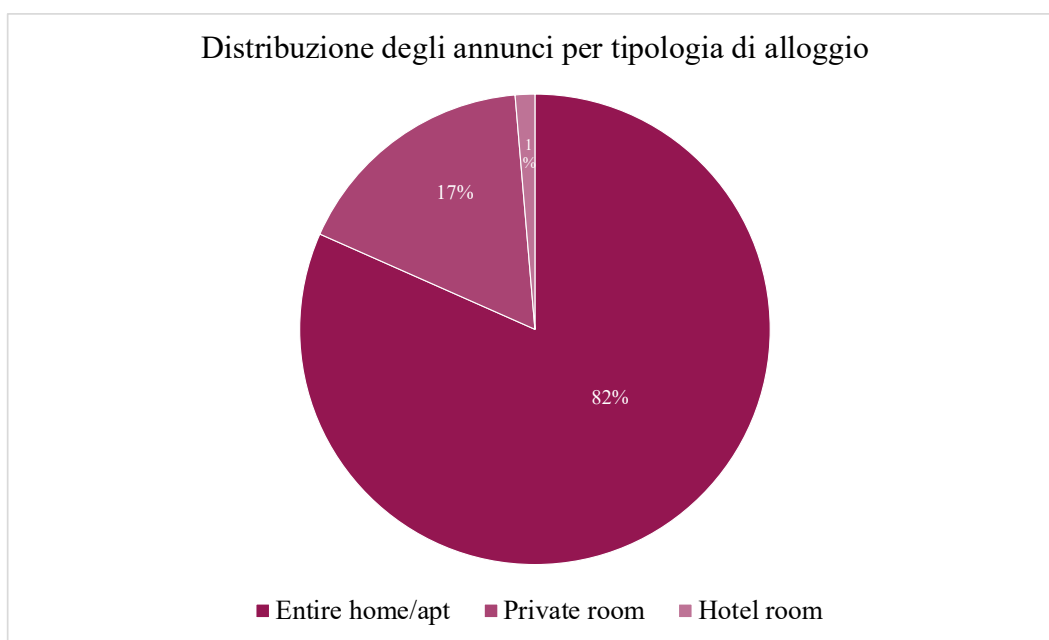


Figura 11: Distribuzione degli annunci per tipologia di alloggio

Il grafico a torta mostra la distribuzione degli annunci in base alla forma di affitto. La componente nettamente prevalente è rappresentata dagli alloggi interi, che costituiscono circa l'82% del campione. Le stanze private rappresentano una quota più contenuta, pari a circa il 17%, mentre le soluzioni di tipo alberghiero risultano marginali. Questa distribuzione evidenzia come l'offerta analizzata sia fortemente orientata verso forme di ospitalità che garantiscono maggiore privacy e autonomia agli ospiti, in linea con il posizionamento tipico di Airbnb rispetto alle strutture ricettive tradizionali.

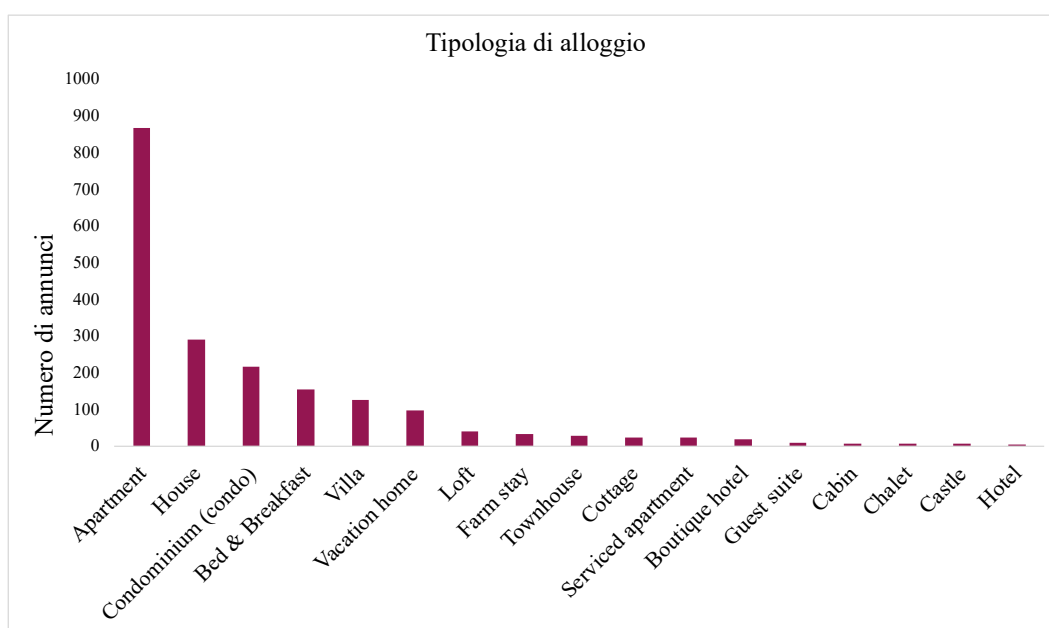


Figura 12: Distribuzione delle principali categorie di alloggio

Attraverso il grafico presente in *Figura 12*, viene approfondita la composizione dell'offerta all'interno delle principali categorie immobiliari. Tra queste, gli appartamenti risultano di gran lunga la tipologia più diffusa, seguiti dalle case indipendenti e dai condomini. Le altre categorie, come bed & breakfast, ville e case vacanza, presentano una diffusione più limitata, mentre le tipologie più atipiche o di nicchia compaiono solo in modo marginale. Questo andamento conferma che, pur in presenza di una certa varietà di soluzioni abitative, l'offerta è concentrata su poche categorie principali, riflettendo una struttura del mercato relativamente standardizzata dal punto di vista immobiliare.

Distribuzione geografica degli alloggi

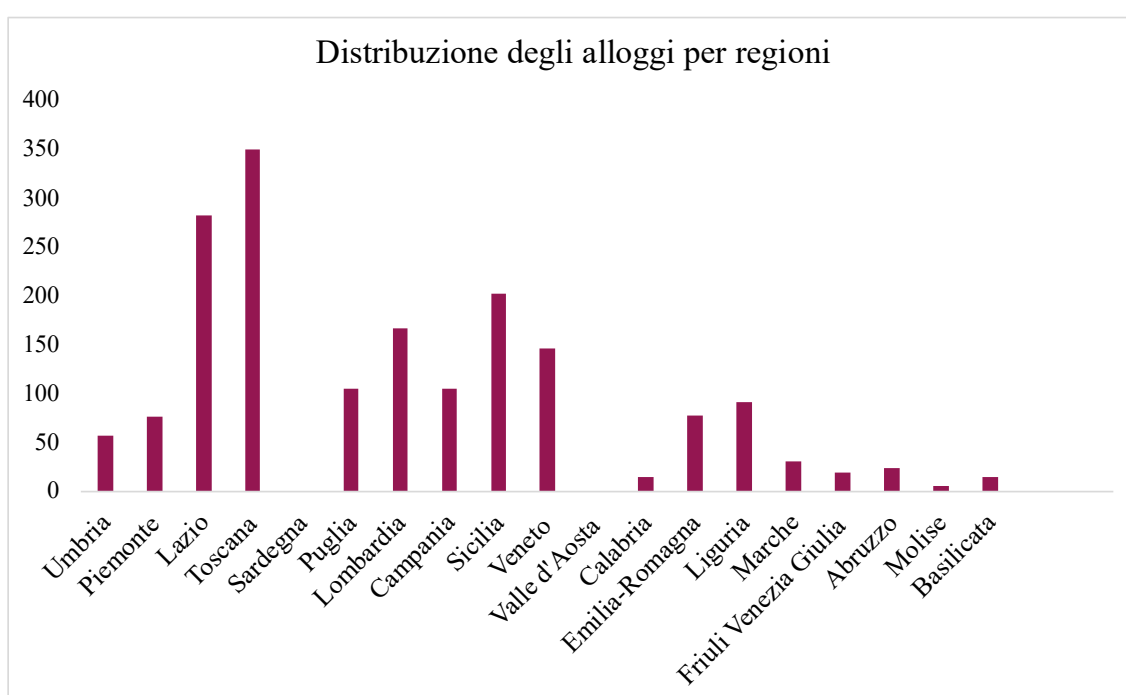


Figura 13: Distribuzione degli alloggi per regioni

Questo primo grafico mostra la distribuzione degli alloggi a livello regionale. La presenza degli annunci risulta fortemente concentrata in alcune aree, in particolare in Toscana e Lazio, che rappresentano le regioni con il maggior numero di osservazioni nel campione. Seguono Lombardia, Sicilia e Veneto, mentre molte altre regioni presentano una numerosità più contenuta. Questa distribuzione riflette il peso delle principali destinazioni turistiche e culturali all'interno del mercato Airbnb, evidenziando una concentrazione geografica dell'offerta coerente con l'attrattività dei territori.

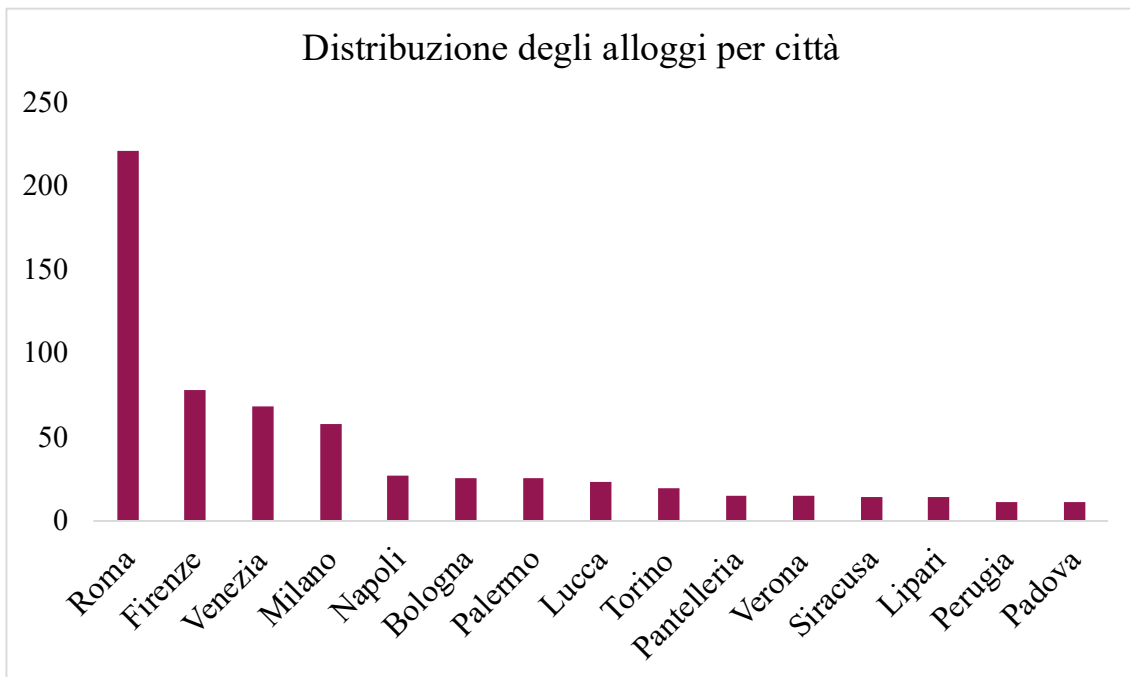


Figura 14: Distribuzione degli alloggi per città

Scendendo al livello urbano, è mostrato un secondo grafico in *Figura 14* che mette in luce una concentrazione ancora più marcata. Roma emerge come la città con il maggior numero di alloggi, seguita da Firenze, Venezia e Milano. Le restanti città presentano una diffusione più limitata degli annunci, spesso legata a specifiche vocazioni turistiche o a contesti locali particolari.

Questo andamento conferma l'esistenza di forti differenze territoriali all'interno del campione, giustificando l'inclusione di effetti fissi a livello urbano nelle analisi econometriche successive.

Politiche di cancellazione degli alloggi

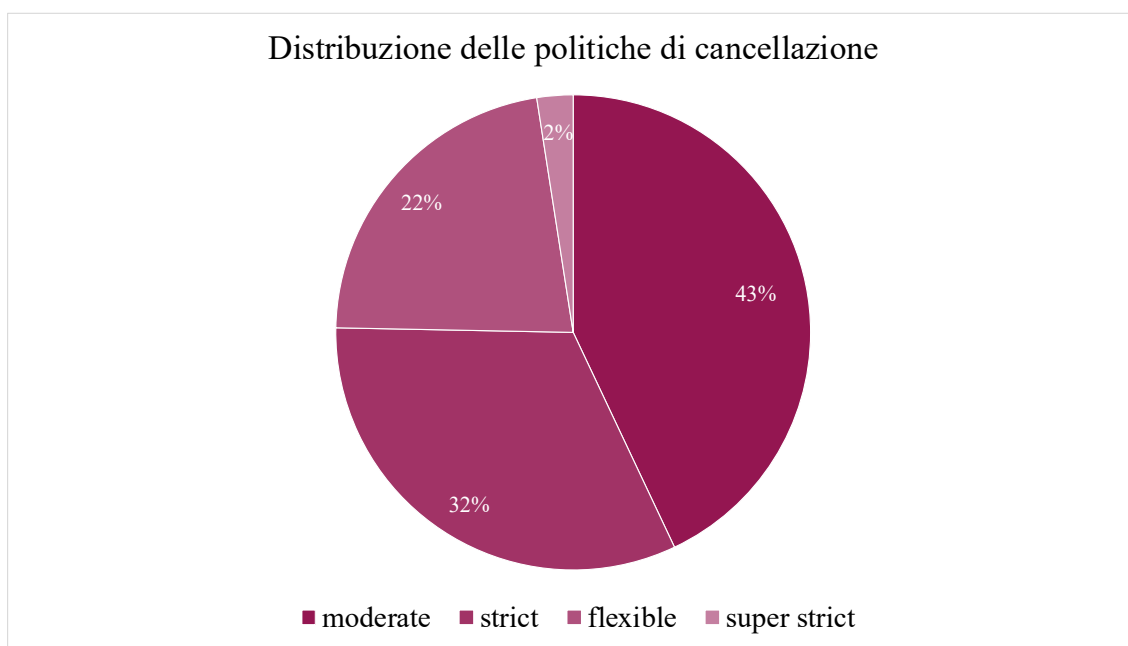


Figura 15: Distribuzione delle politiche di cancellazione

Il grafico in *Figura 15* mostra come gli annunci si distribuiscano tra le diverse tipologie di politiche di cancellazione. La politica *moderate* risulta la più diffusa, interessando circa il 43% degli alloggi, seguita dalle politiche *strict* (32%) e *flexible* (22%). Le politiche *super strict* rappresentano invece una quota residuale del campione. Questa configurazione suggerisce che molti host optino per soluzioni intermedie, che consentono una certa flessibilità agli ospiti senza esporsi eccessivamente al rischio di cancellazioni tardive. Allo stesso tempo, la presenza di una quota rilevante di politiche più rigide indica che una parte degli host privilegia la stabilità dei ricavi e una maggiore prevedibilità delle prenotazioni. La coesistenza di approcci differenti riflette quindi strategie eterogenee nella gestione del rischio e contribuisce a rendere la politica di cancellazione una dimensione rilevante nel confronto delle performance degli annunci.

4.3.3 Matrice di correlazione

Prima di procedere alla stima del modello econometrico, è fondamentale esaminare le relazioni bivariate tra le variabili oggetto di studio. Questa analisi preliminare ha un duplice obiettivo. Da un lato, permette di osservare le associazioni naturali tra le metriche di performance e le caratteristiche degli annunci, dall'altro, serve a identificare potenziali problemi di multicollinearità tra i regressori, che potrebbero ridurre la significatività statistica dei risultati.

Data l'elevata numerosità delle variabili incluse nel dataset, la *Tabella 3* riporta una selezione rappresentativa delle correlazioni di Pearson. La tabella include le principali variabili dipendenti, gli indici testuali di maggior interesse teorico e i controlli strutturali chiave.

Matrice di correlazione	log_revpan	log_adr	occupationrate	WC	social	cognitive	home	affective	log_bedrooms	superhost	responserate	urban_class
log_revpan	1.00											
log_adr	0.53***	1.00										
occupationrate	0.73***	-0.08**	1.00									
WC	0.05	0.13***	0.07**	1.00								
social	-0.04	0.03	0.00	0.03	1.00							
cognitive	-0.08**	-0.04	-0.02	0.07**	0.23***	1.00						
home	0.09**	0.07**	-0.01	-0.24***	-0.05*	-0.14***	1.00					
affective	0.02	0.01	0.00	-0.02	0.04	0.09***	-0.14***	1.00				
log_bedrooms	0.23***	0.53***	-0.08***	0.10***	0.03	-0.06*	-0.04	-0.04	1.00			
superhost	-0.01	-0.08**	0.14***	0.15***	0.02	0.02	-0.07**	0.01	-0.09***	1.00		
responserate	0.08**	0.05	0.16***	0.13***	0.02	0.04	-0.11***	-0.03	-0.01	0.24***	1.00	
urban_class	-0.10***	0.09**	-0.21***	-0.05*	0.05*	0.02	-0.04	0.02	0.25***	-0.13***	-0.04	1.00

Tabella 3: Matrice di correlazione di Spearman

L'analisi dei coefficienti in tabella conferma che i dati sono coerenti. Come prevedibile, c'è una relazione positiva ($r = 0.53$) tra la dimensione della casa (*Log Bedrooms*) e il prezzo (*Log ADR*), e un legame molto forte tra il tasso di occupazione e il ricavo generato (*Log RevPAN*, $r = 0.73$).

Per quanto riguarda le variabili testuali, le correlazioni restano basse, il che suggerisce che le diverse categorie semantiche non si sovrappongono tra loro. Ad esempio, la lunghezza della descrizione ha una bassa correlazione con lo status di Superhost ($r = 0.15$) indicando che, anche se gli host esperti tendono a scrivere un po' di più, le due variabili misurano aspetti ben distinti.

Per essere certi di non avere distorsioni nelle stime, è stato calcolato il VIF (Variance Inflation Factor) per tutte le variabili. I risultati confermano la solidità del modello poiché la media del VIF è molto bassa, pari a 1.40. Questo dato è fondamentale perché garantisce che le variabili non siano ridondanti tra loro. Ogni regressore, infatti, apporta un contributo informativo unico, permettendo di stimare il suo impatto specifico sul prezzo senza le interferenze tipiche della multicollinearità.

I valori più alti riguardano la capacità ospitativa (Max Guests) e il numero di stanze, dato che sono caratteristiche fisiche collegate tra loro; tuttavia, rimangono al di sotto delle soglie di allarme (10 o 5).

Le variabili testuali, invece, hanno valori vicinissimi a 1, confermando la loro totale indipendenza.

CAPITOLO V – Risultati empirici

L'analisi inferenziale si concentra sull'isolamento del valore economico generato dalle strategie linguistiche, stimando l'effetto delle descrizioni testuali al netto delle variabili strutturali e reputazionali. L'applicazione del modello econometrico consente di determinare se, e in che misura, l'adozione di specifici registri comunicativi si traduca in una variazione statisticamente significativa del prezzo di mercato.

5.1 Specificazione del modello econometrico

Per analizzare le determinanti delle performance su Airbnb, è stato specificato un modello di regressione lineare multipla (OLS) che include variabili di controllo, variabili indipendenti, effetti fissi e termini di moderazione.

L'equazione completa stimata è la seguente:

$$Y_i = \alpha + \underbrace{\sum \beta_c C_i}_{\text{Controlli}} + \underbrace{\sum \beta_x X_i}_{\text{Variabili Indipendenti}} + \underbrace{\sum \beta_m (X_i \times M_i)}_{\text{Moderazione}} + \underbrace{\mu_l + \lambda_t}_{\text{Effetti Fissi}} + \epsilon_i$$

In cui,

1. **Y_i** (variabili dipendenti): 5 metriche di performance analizzate.
2. **C** (variabili di controllo): variabili strutturali e reputazionali che non sono oggetto delle ipotesi, inserite nel modello per isolare l'effetto delle variabili di interesse (es. numero di bagni, di stanze, massimo numero di ospiti ecc.).
3. **X_i** (variabili indipendenti): variabili linguistiche su cui si basano le ipotesi (es. *WordCount*, *Host orientation* ecc.).
4. **X_i×M_i** (effetti di moderazione): termini di interazione volti a verificare se l'effetto delle variabili linguistiche varia in funzione di specifici moderatori.
5. **μ_l** e **λ_t** (effetti fissi): parametri inclusi per catturare l'eterogeneità non osservata legata alla localizzazione e alla tipologia di alloggio.

5.2 Analisi dei risultati delle regressioni tramite Script R

Considerata l'ampiezza dell'impianto analitico, che comprende un totale di 18 variabili indipendenti, si è scelto di strutturare l'analisi suddividendo le metriche in tre macrocategorie logiche. Questa strategia risponde da un lato a un'esigenza di evitare potenziali problemi di multicollinearità che potrebbero emergere inserendo simultaneamente tutte le variabili testuali in un'unica equazione, dall'altro consente di garantire una maggiore chiarezza interpretativa isolando i diversi livelli della comunicazione.

Di conseguenza, i risultati sono presentati attraverso tre modelli di stima distinti, ciascuno volto a indagare una specifica dimensione della strategia comunicativa dell'host:

1. *Metriche di complessità del testo*: analisi della quantità e della complessità sintattica del testo.
2. *Metriche relazionali*: analisi dello stile interpersonale attraverso l'uso dei pronomi.
3. *Metriche semantiche*: analisi dei temi dominanti e del contenuto emozionale.

5.2.1 Metriche di complessità del testo

	M1	M2	M3	M4	M5
	Log RevPAN	Log ADR	Log Rating	Log Reservation	Occup. Rate
Log Words Count	0.0146 (0.18)	0.0257 (0.57)	0.00286 (0.54)	-0.00481 (-0.07)	0.0133 (0.53)
Log Words per Sentence	0.199 (1.28)	0.0110 (0.13)	-0.00135 (-0.15)	0.225 (1.67)	0.0374 (0.83)
Log Complexity	0.0484 (0.12)	0.0363 (0.17)	0.0398 (1.05)	-0.201 (-0.57)	0.0163 (0.15)
Superhost	-0.105 (-0.92)	-0.0849 (-1.47)	0.0323*** (5.09)	-0.0628 (-0.62)	0.0113 (0.30)
Log Bedrooms	0.295 (1.54)	0.0758 (0.76)	0.00285 (0.28)	-0.0900 (-0.56)	0.0165 (0.30)
Log Bathrooms	0.428 (1.81)	0.433** (3.23)	0.0162 (1.35)	0.130 (0.65)	-0.00581 (-0.10)
Log Photos	0.138 (1.21)	0.0657 (1.13)	0.00700 (1.12)	0.122 (1.32)	0.0146 (0.47)
Max guests	0.00118 (0.02)	0.0756*** (3.33)	-0.00291 (-1.13)	-0.0326 (-0.85)	-0.00721 (-0.62)
Minimum stay	0.000931 (0.54)	-0.000709 (-1.04)	0.0000380 (0.34)	0.00168 (1.28)	0.000137 (0.14)
Response rate	0.0113 (1.88)	0.00490* (2.35)	0.000466 (1.37)	0.00683 (1.75)	0.00112 (1.15)
Cancellation policy	-0.000957 (-0.01)	-0.0776 (-1.32)	-0.00152 (-0.23)	0.0303 (0.32)	0.0278 (0.82)
Constant	1.096 (0.64)	3.484*** (4.01)	1.338*** (8.26)	0.109 (0.08)	-0.0945 (-0.23)
City FE	Yes	Yes	Yes	Yes	Yes
Property type FE	Yes	Yes	Yes	Yes	Yes
N	1153	1153	1686	1153	1781

Tabella 4: Risultati delle regressioni: effetti della struttura e complessità del testo sulle performance – Script R

Le metriche di struttura e complessità del testo non mostrano un impatto statisticamente significativo sulle principali misure di performance economica degli annunci Airbnb. Nonostante si possa notare una certa coerenza nel segno dei coefficienti relativi alla lunghezza della descrizione o alla ricchezza del vocabolario, i valori rimangono molto bassi per poter affermare che uno stile di scrittura più elevato o articolato si traduca direttamente in un aumento dei ricavi o delle prenotazioni. Questo aspetto diventa ancora più evidente e robusto quando il modello statistico isola questi fattori dalle caratteristiche fisiche dell'alloggio o dalla sua posizione, confermando che la forma del testo passa in secondo piano rispetto alla sostanza dell'offerta immobiliare.

Questo risultato trova una spiegazione logica considerando le reali abitudini di chi naviga sulle piattaforme digitali, dove l'attenzione è spesso limitata e la rapidità di consultazione diventa fondamentale per non perdere l'interesse dell'utente. Scrivere descrizioni eccessivamente lunghe o sintatticamente complesse non sembra offrire alcun vantaggio

competitivo reale perché i potenziali ospiti non sono alla ricerca di sofisticazione letteraria ma di risposte rapide e funzionali alle loro esigenze. I dati suggeriscono quindi che la strategia vincente non risiede nella ricercatezza formale ma nella capacità di essere chiari, essenziali e diretti. In un contesto dove l'utente scorre velocemente tra decine di opzioni diverse, la semplicità comunicativa permette di trasferire il valore dell'offerta in modo molto più efficace, lasciando che siano la qualità intrinseca dell'alloggio e la trasparenza delle informazioni a convincere il cliente finale.

5.2.2 Metriche relazionali

	M1 Log RevPAN	M2 Log ADR	M3 Log Rating	M4 Log Reservation	M5 Occup. Rate
Host orientation	-0.2068 (0.1257)	-0.1761*** (0.0618)	0.0018 (0.0068)	-0.1178 (0.0990)	-0.0052 (0.0313)
Professional orientation	-0.0871 (0.0637)	-0.0136 (0.0313)	0.0083** (0.0040)	-0.0022 (0.0501)	-0.0051 (0.0175)
Guest orientation	0.0876 (0.0680)	0.0333 (0.0334)	-0.0008 (0.0043)	0.1015* (0.0536)	-0.0166 (0.0173)
Superhost	-0.1191 (0.0743)	-0.0832** (0.0365)	0.0304*** (0.0045)	-0.1297** (0.0585)	0.0214 (0.0201)
Log Bedrooms	0.2634** (0.1289)	0.0812 (0.0634)	-0.0028 (0.0070)	-0.0119 (0.1015)	-0.0062 (0.0311)
Log Bathrooms	0.3218** (0.1386)	0.4042*** (0.0681)	0.0242*** (0.0077)	-0.0235 (0.1091)	-0.0287 (0.0328)
Log Photos	0.1089 (0.0683)	0.0652* (0.0336)	0.0088** (0.0037)	0.0249 (0.0538)	0.0268* (0.0162)
Max guests	0.0352 (0.0272)	0.0750*** (0.0134)	-0.0034** (0.0015)	0.0006 (0.0215)	0.0011 (0.0067)
Minimum stay	0.0009 (0.0017)	-0.0008 (0.0008)	0.0000 (0.0001)	0.0016 (0.0013)	0.0002 (0.0005)
Response rate	0.0082*** (0.0028)	0.0031** (0.0014)	0.0004*** (0.0001)	0.0053** (0.0022)	0.0017*** (0.0005)
Cancellation policy	-0.0338 (0.0828)	0.0140 (0.0407)	-0.0128*** (0.0049)	-0.0791 (0.0652)	-0.0194 (0.0222)
Constant	0.6564 -13238	4.0315*** (0.6508)	1.6253*** (0.0974)	0.2291 -10426	-0.2335 (0.4420)
City FE	Yes	Yes	Yes	Yes	Yes
Property type FE	Yes	Yes	Yes	Yes	Yes
N	1154	1154	1687	1154	1782
R ²	0.4151	0.6019	0.1070	0.1697	0.2220

Tabella 5: Risultati delle regressioni: impatto delle strategie comunicative sulle performance – Script R

I risultati esposti nella Tabella 5 mettono in luce come le diverse strategie comunicative adottate dall'host non siano neutre ma orientino le performance dell'annuncio secondo logiche ben distinte. Si osserva con chiarezza che un approccio focalizzato eccessivamente sulla figura del proprietario risulta penalizzante per il posizionamento economico dell'offerta. La variabile *host orientation* mostra infatti un impatto negativo e statisticamente significativo sul prezzo medio giornaliero (-0,176***) lasciando intendere

che una narrazione troppo autoreferenziale potrebbe distogliere le attenzioni dalle esigenze del cliente e far abbassare il valore percepito del servizio rendendo l'ospite meno propenso a pagare cifre elevate.

Al contrario il mercato sembra premiare chi riesce a trasmettere affidabilità e attenzione verso l'altro. La *professional orientation* si associa positivamente al rating complessivo (+0,0083***) dimostrando che un'immagine di gestione solida, ordinata e strutturata è fondamentale per costruire la fiducia necessaria a ottenere valutazioni eccellenti. Parallelamente si nota come la *guest orientation* svolga un ruolo cruciale nello stimolare la domanda attiva poiché incide direttamente sul volume delle prenotazioni (+0,1015*). Questo dato conferma che rivolgersi direttamente al potenziale ospite e mettere al centro la sua esperienza è la leva più efficace per favorire la conversione e trasformare l'interesse in una prenotazione concreta.

I dati suggeriscono di abbandonare l'autoreferenzialità in favore di una comunicazione che sia professionale nella forma ed empatica nella sostanza per massimizzare sia la reputazione che i risultati commerciali.

La *Figura 16* riporta una sintesi grafica dell'impatto delle variabili di orientamento relazionale sull'ADR.

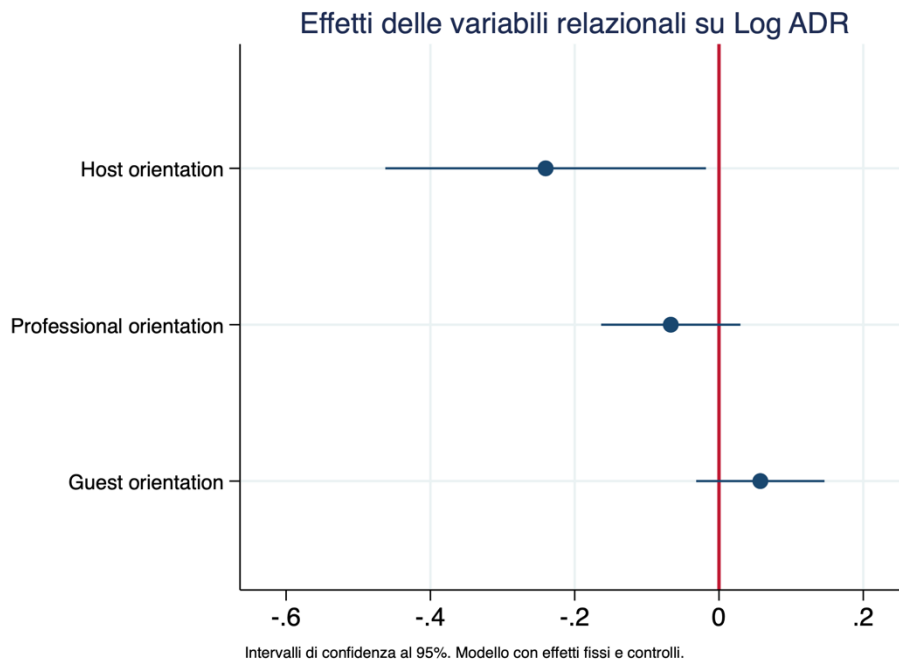


Figura 16: Effetti delle variabili relazionali su LogADR – SCRIPT R

5.2.3 Metriche semantiche

	M1 Log RevPAN	M2 Log ADR	M3 Log Rating	M4 Log Reservation	M5 Occup. Rate
Psychological process variables					
Affective	0.0526 (0.0729)	0.0316 (0.0353)	-0.0064 (0.0042)	0.1014* (0.0577)	0.0136 (0.0185)
Cleanliness	0.0225 (0.1117)	0.0333 (0.0540)	-0.0002 (0.0061)	-0.0470 (0.0883)	-0.0276 (0.0271)
Cognitive	-0.0508 (0.0771)	0.0138 (0.0373)	-0.0045 (0.0045)	-0.0277 (0.0610)	-0.0097 (0.0200)
Perceptual	0.0099 (0.0385)	0.0195 (0.0186)	0.0007 (0.0022)	-0.0251 (0.0305)	0.0012 (0.0100)
Relativity	-0.0194 (0.0265)	-0.0283** (0.0128)	0.0001 (0.0015)	0.0116 (0.0210)	0.0051 (0.0067)
Safety	0.0322 (0.1076)	-0.0206 (0.0521)	-0.0049 (0.0059)	0.0306 (0.0851)	-0.0072 (0.0261)
Social	-0.0161 (0.0365)	0.0023 (0.0177)	0.0020 (0.0021)	-0.0060 (0.0289)	0.0086 (0.0093)
Status	0.0897* (0.0488)	0.0930*** (0.0236)	0.0017 (0.0031)	0.0271 (0.0386)	-0.0039 (0.0136)
Politeness	-0.0168 (0.1289)	0.0588 (0.0624)	-0.0161** (0.0079)	-0.0804 (0.1020)	0.0018 (0.0356)
Personal concern variables					
Home	0.0134 (0.0135)	0.0082 (0.0065)	-0.0006 (0.0008)	0.0027 (0.0107)	-0.0008 (0.0033)
Leisure	0.1130** (0.0536)	0.0561** (0.0259)	0.0006 (0.0024)	0.0163 (0.0424)	0.0039 (0.0103)
Money	-0.0305 (0.0598)	-0.0371 (0.0289)	-0.0065** (0.0032)	0.0118 (0.0473)	-0.0071 (0.0140)
Superhost	-0.1131 (0.0750)	-0.0797** (0.0363)	0.0310*** (0.0045)	-0.1274** (0.0594)	0.0209 (0.0203)
Log Bedrooms	0.2584** (0.1296)	0.0795 (0.0627)	-0.0050 (0.0071)	-0.0225 (0.1025)	-0.0112 (0.0315)
Log Bathrooms	0.3261** (0.1396)	0.3987*** (0.0675)	0.0234*** (0.0077)	-0.0238 (0.1104)	-0.0270 (0.0330)
Log Photos	0.1294* (0.0679)	0.0739** (0.0328)	0.0086** (0.0037)	0.0405 (0.0537)	0.0287* (0.0163)
Max guests	0.0283 (0.0278)	0.0703*** (0.0134)	-0.0032** (0.0015)	0.0017 (0.0220)	0.0022 (0.0068)
Minimum stay	0.0011 (0.0018)	-0.0009 (0.0009)	0.0001 (0.0001)	0.0020 (0.0014)	0.0002 (0.0005)
Response rate	0.0081*** (0.0028)	0.0029** (0.0014)	0.0004*** (0.0001)	0.0056** (0.0022)	0.0017*** (0.0005)
Cancellation policy	-0.0405 (0.0837)	0.0083 (0.0405)	-0.0125** (0.0050)	-0.0808 (0.0662)	-0.0198 (0.0223)
Constant	0.7779 -13292	4.0978*** (0.6428)	1.6180*** (0.0974)	0.2037 -10513	-0.2553 (0.4439)
City FE	Yes	Yes	Yes	Yes	Yes
Property type FE	Yes	Yes	Yes	Yes	Yes
N	1154	1154	1686	1154	1781
R ²	0.4129	0.6133	0.1076	0.1594	0.2172

Tabella 6: Risultati delle regressioni: impatto delle variabili semantiche sulle performance – Script R

La distinzione tra dimensioni psicologiche e variabili legate alle preoccupazioni personali consente di delineare un quadro significativo per comprendere quali contenuti testuali

generino valore reale. All'interno della sfera psicologica si nota un comportamento contrastante in cui i riferimenti allo *status* permettono di sostenere un prezzo medio più alto (+0.0930***), a differenza di quanto si osserva per la dimensione *relativity* che influisce negativamente sullo stesso indicatore (-0.0283**), suggerendo che l'uso di confronti o termini relativi finisce per ridurre la disponibilità economica dell'utente.

L'influenza del testo diventa ancora più incisiva quando tocca aspetti tangibili dell'esperienza come dimostra la variabile *leisure* che spinge verso l'alto sia i ricavi complessivi (+0.1130**) sia il prezzo medio (+0.0561**) confermando quanto sia redditizio richiamare scenari di svago e relax. L'insistenza sui costi, invece, rappresentata dalla variabile *money* si rivela dannosa per la reputazione abbassando il rating finale (-0.0065**) ed evidenziando come l'ospite percepisca negativamente un'eccessiva attenzione agli aspetti monetari.

Appare, quindi, chiaro che un linguaggio incentrato su benefici concreti e piaceri tangibili risulta strategicamente più efficace rispetto all'uso di leve puramente emotive o astratte il cui impatto appare meno stabile.

Nelle figure successive è riportata una sintesi grafica degli effetti stimati delle variabili semantiche rispettivamente sull'ADR e sul rating, consentendo un confronto immediato tra le diverse dimensioni linguistiche considerate.

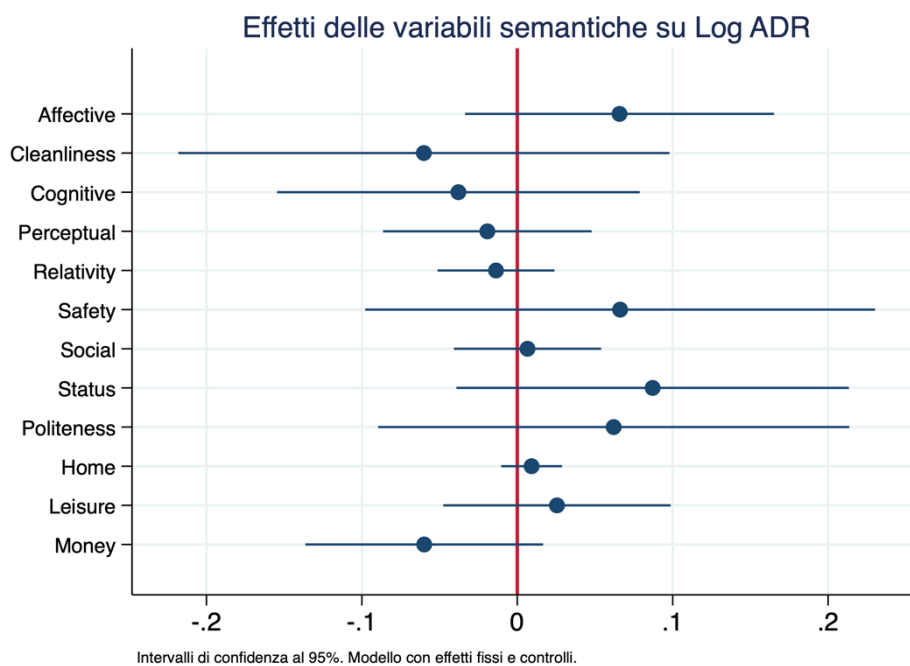


Figura 17: Effetti delle variabili semantiche su LogADR – SCRIPT R

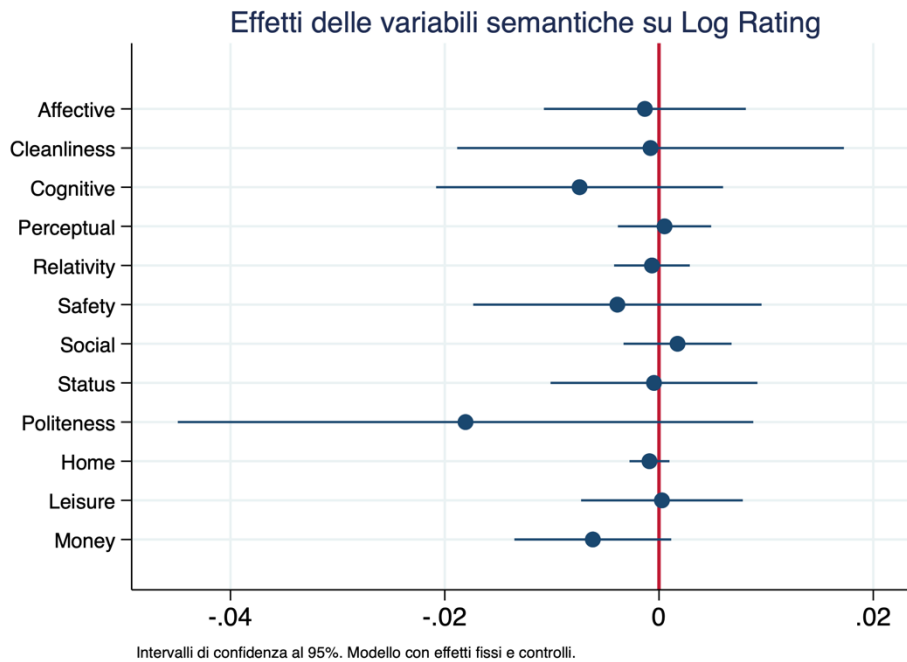


Figura 18: Effetti delle variabili semantiche su Log Rating – SCRIPT R

5.3 Approfondimento metodologico tramite software LIWC

Lo studio approfondito delle determinanti testuali richiede un approccio capace di andare oltre la semplice conta delle parole o la valutazione della complessità sintattica per esaminare come il linguaggio rifletta processi cognitivi ed emotivi. La letteratura accademica affronta questa sfida attraverso l'impiego di software di analisi automatica del contenuto progettati per identificare specifici pattern linguistici associati a diverse dimensioni psicologiche. Tra gli strumenti più accreditati, si distingue il programma LIWC il quale si basa su decenni di ricerca per classificare il lessico in categorie che spaziano dalle emozioni positive e negative fino ai processi logici e sociali.

Il software LIWC (Linguistic Inquiry and Word Count) si identifica come uno strumento di analisi testuale computerizzata in grado di convertire il contenuto grezzo delle descrizioni in dati quantitativi strutturati. La caratteristica principale del sistema risiede nell'utilizzo di un sofisticato algoritmo di *dictionary-matching* il quale esamina ogni singolo termine del testo e verifica la sua appartenenza a specifiche categorie linguistiche predefinite. Questo processo permette di superare la soggettività della lettura umana garantendo una codifica oggettiva e replicabile delle variabili testuali attraverso il confronto con un dizionario validato da decenni di ricerca.

L'applicazione di questa metodologia consente di ottenere un quadro più dettagliato delle strategie comunicative offrendo una prospettiva complementare rispetto alle metriche tradizionali.

Analogamente all'approccio metodologico adottato nei paragrafi precedenti, anche per le metriche derivanti dall'analisi LIWC è stata effettuata una ripartizione in tre gruppi distinti. Ciò risponde all'esigenza di organizzare la vasta gamma di output forniti dal software in una struttura coerente e leggibile, permettendo così di isolare e confrontare separatamente le diverse dimensioni del linguaggio

5.3.1 Metriche di complessità del testo

Per approfondire l'analisi strutturale, sono state introdotte due nuove variabili LIWC che consentono di mettere in evidenza sfumature diverse rispetto alla semplice lunghezza.

- **Dictionary Words:** calcola la percentuale di parole riconosciute dal dizionario ufficiale, indicando quanto il linguaggio usato sia standard e convenzionale.
- **Periods:** conta i punti fermi presenti nel testo e serve a misurare come viene suddiviso il discorso, offrendo un indice chiaro della sua articolazione sintattica.

	M1 Log RevPAN	M2 Log ADR	M3 Log Rating	M4 Log Reservation	M5 Occup. Rate
Log Words Count	-0.0301 (-0.36)	0.00922 (0.22)	0.00243 (0.43)	-0.0118 (-0.16)	0.00561 (0.21)
Log Word per Sentence	0.157 (0.69)	-0.0108 (-0.08)	0.00306 (0.16)	0.286 (1.48)	0.0329 (0.49)
Log Complexity	0.0805 (0.19)	0.000676 (0.00)	0.0350 (0.88)	-0.0863 (-0.23)	0.0417 (0.36)
Log Dictionary Words	-0.167 (-0.35)	-0.178 (-0.83)	-0.0161 (-0.68)	0.112 (0.31)	0.0478 (0.38)
Log Periods	-0.0827 (-0.44)	-0.0142 (-0.11)	0.00422 (0.31)	0.00757 (0.04)	-0.0254 (-0.40)
Superhost	-0.0830 (-0.73)	-0.0879 (-1.50)	0.0323*** (4.92)	-0.0439 (-0.44)	0.0152 (0.40)
Log Bedrooms	0.281 (1.45)	0.0777 (0.76)	0.00162 (0.16)	-0.0960 (-0.59)	0.0189 (0.33)
Log Bathrooms	0.454 (1.87)	0.394** (2.83)	0.0156 (1.19)	0.193 (0.94)	0.00294 (0.05)
Log Photos	0.145 (1.27)	0.0786 (1.37)	0.00628 (0.96)	0.124 (1.33)	0.0113 (0.35)
Max guests	0.00536 (0.11)	0.0776*** (3.36)	-0.00248 (-0.94)	-0.0335 (-0.85)	-0.00722 (-0.61)
Minimum stay	0.000989 (0.56)	-0.000725 (-1.03)	0.0000342 (0.29)	0.00185 (1.44)	0.000190 (0.20)
Response rate	0.0105 (1.72)	0.00416* (2.10)	0.000518 (1.46)	0.00750 (1.89)	0.00114 (1.12)
Cancellation policy	-0.00725 (-0.06)	-0.0808 (-1.36)	-0.000617 (-0.09)	0.0301 (0.31)	0.0293 (0.85)
Constant	2.128 (0.79)	4.493** (3.18)	1.393*** (7.48)	-0.998 (-0.42)	-0.269 (-0.36)
Effetti Fissi City	Yes	Yes	Yes	Yes	Yes
Effetti Fissi Property	Yes	Yes	Yes	Yes	Yes
N	1140	1140	1666	1140	1758

Tabella 7: Risultati delle regressioni: effetti della struttura e complessità del testo sulle performance – LIWC

Le misure relative alla quantità e alla struttura del testo, incluse le variabili LIWC appena introdotte (*dictionary words* e *periods*), non mostrano alcuna associazione statisticamente rilevante con ricavi, prezzi o valutazioni. Tale risultato evidenzia come la forma del testo, intesa sia come aderenza a un lessico standardizzato sia come articolazione sintattica, non costituisca un fattore discriminante per il successo dell'annuncio.

Ne deriva che la qualità percepita dall'utente non è legata alla complessità formale o alla lunghezza della descrizione, dimostrando che non è utile cercare uno stile troppo ricercato; al contrario, i dati suggeriscono di indirizzare gli sforzi comunicativi verso la

chiarezza informativa e la pertinenza dei contenuti, uniche vere leve in grado di rispondere alle aspettative degli ospiti.

5.3.2 Metriche relazionali

	M1 Log RevPAN	M2 Log ADR	M3 Log Rating	M4 Log Reservation	M5 Occup. Rate
Host orientation	-0.2736 (0.2369)	-0.1928* (0.1121)	0.0032 (0.0117)	-0.1309 (0.1565)	-0.0004 (0.0505)
Professional orientation	-0.0521 (0.1184)	-0.0397 (0.0548)	0.0037 (0.0074)	0.1026 (0.0948)	0.0150 (0.0343)
Guest orientation	11171 -13757	-0.7850 (0.8401)	0.0261 (0.0715)	12847 -14333	0.2325 (0.7116)
Superhost	-0.1006 (0.1152)	-0.0751 (0.0582)	0.0319*** (0.0063)	-0.0602 (0.1011)	0.0106 (0.0376)
Log Bedrooms	0.3211* (0.1909)	0.0929 (0.0989)	-0.0002 (0.0101)	-0.0811 (0.1622)	0.0156 (0.0553)
Log Bathrooms	0.4313* (0.2371)	0.4345*** (0.1331)	0.0172 (0.0120)	0.1524 (0.2024)	-0.0017 (0.0595)
Log Photos	0.1238 (0.1116)	0.0610 (0.0552)	0.0082 (0.0059)	0.1245 (0.0909)	0.0208 (0.0300)
Max guests	-0.0073 (0.0497)	0.0704*** (0.0226)	-0.0027 (0.0025)	-0.0381 (0.0394)	-0.0076 (0.0117)
Minimum stay	0.0007 (0.0018)	-0.0007 (0.0007)	0.0000 (0.0001)	0.0014 (0.0014)	0.0001 (0.0010)
Response rate	0.0111* (0.0058)	0.0050** (0.0022)	0.0005 (0.0003)	0.0067* (0.0038)	0.0011 (0.0010)
Cancellation policy	0.0091 (0.1168)	-0.0698 (0.0579)	-0.0009 (0.0066)	0.0393 (0.0952)	0.0281 (0.0339)
Constant	2.0575*** (0.6838)	3.8173*** (0.2483)	1.4770*** (0.0338)	0.0710 (0.4420)	0.1178 (0.1238)
Effetti Fissi City	Yes	Yes	Yes	Yes	Yes
Effetti Fissi Property	Yes	Yes	Yes	Yes	Yes
N	1154	1154	1687	1154	1782
R ²	0.4915	0.6370	0.1213	0.2048	0.2318

Tabella 8: Risultati delle regressioni: impatto delle strategie comunicative sulle performance – LIWC

I risultati ottenuti attraverso l'algoritmo LIWC offrono una visione più selettiva sulle leve relazionali, isolando la *Host Orientation* come unica determinante significativa delle performance economiche. Il coefficiente negativo sul prezzo (-0.1928*) evidenzia come una narrazione autocentrata venga percepita come un limite dell'offerta. È ragionevole ipotizzare che un testo focalizzato sull'io dell'host sposti l'attenzione dai benefici tangibili dell'alloggio, creando una distanza psicologica che frena la spesa elevata. Contrariamente, le variabili *Professional* e *Guest Orientation* non mostrano impatti rilevanti in questa

specifica analisi, lasciando intendere che il mercato non premia automaticamente la forma cortese se non è supportata da altri fattori.

Questo approfondimento consente di definire con maggiore precisione le conclusioni della fase precedente. Mentre l'analisi basata su script R aveva restituito un quadro più ottimistico, in cui l'orientamento al cliente e la professionalità sembravano trainare rispettivamente domanda e reputazione, il test con LIWC suggerisce che l'impatto di queste variabili positive sia più sfumato e dipendente dallo strumento di misura. Tuttavia, esiste un punto fermo su cui entrambe le metodologie concordano pienamente: la penalizzazione economica derivante dal protagonismo dell'host. I dati confermano che "parlare troppo di sé" riduce i margini di profitto. Trasformando questa evidenza statistica in strategia pratica diventa essenziale rimuovere ogni forma di personalismo dal testo, poiché rappresenta l'unico fattore che incide negativamente sui risultati economici. La *Figura 19* riporta una sintesi grafica dell'impatto delle variabili di orientamento relazionale sull'ADR.

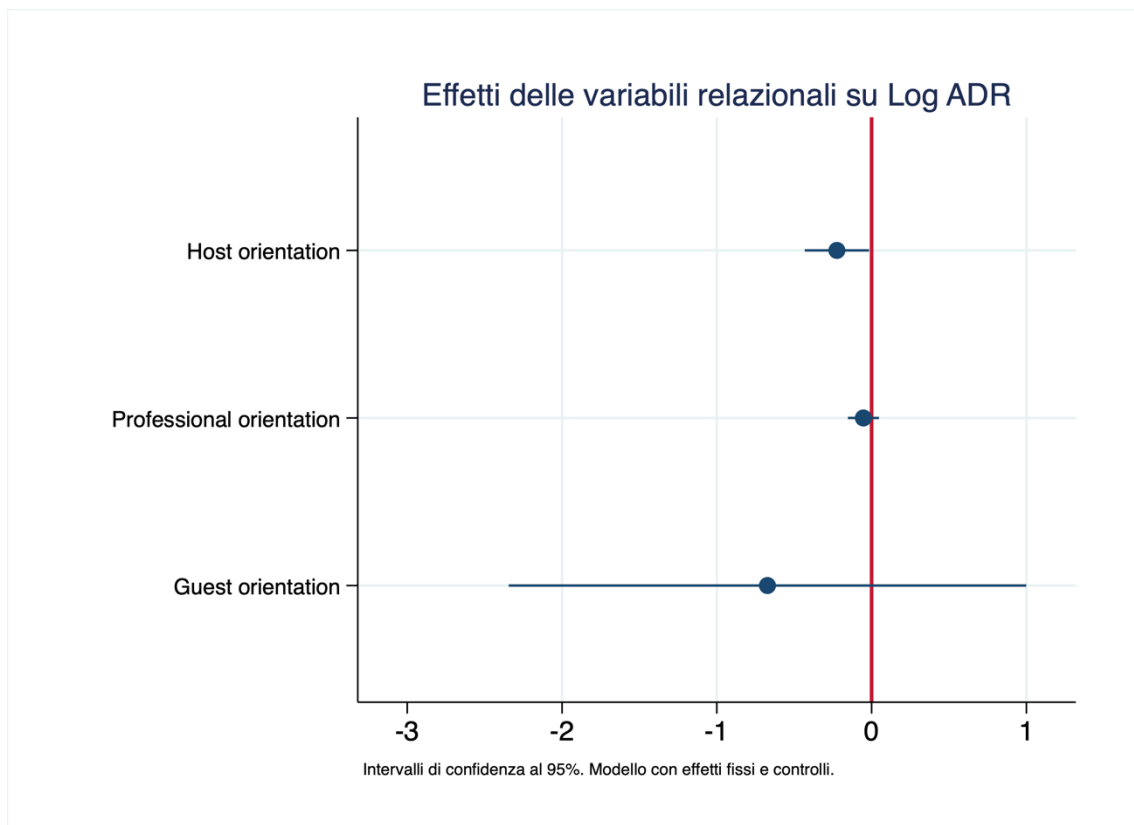


Figura 19: Effetti delle variabili relazionali su LogADR – LIWC

5.3.3 Metriche semantiche

Rispetto all'analisi semantica condotta tramite lo script in R, l'utilizzo degli indici linguistici estratti con LIWC comporta alcune differenze nella definizione delle variabili indipendenti. Alcune dimensioni precedentemente considerate non sono disponibili nella presente specificazione (*cleanliness*, *politeness*, *safety*), mentre sono state introdotte due nuove variabili:

- **Present:** misura la frequenza di verbi e termini declinati al tempo presente.
- **Possibility:** tiene conto dell'uso di espressioni che richiamano possibilità o scenari ipotetici.

L'inclusione di queste variabili consente di estendere l'analisi semantica, incorporando dimensioni temporali e modali del linguaggio non considerate in precedenza.

	M1	M2	M3	M4	M5
	Log RevPAN	Log ADR	Log Rating	Log Reservation	Occup. Rate
Psychological process variables					
Affective	0.0237 (0.55)	0.00567 (0.28)	-0.00110 (-0.41)	-0.0258 (-0.78)	-0.00636 (-0.50)
Cognitive	0.0527 (0.94)	-0.0101 (-0.42)	-0.00458 (-1.43)	0.0800 (1.85)	0.0143 (0.96)
Perceptual	-0.0708 (-0.75)	0.0130 (0.25)	-0.000177 (-0.03)	-0.106 (-1.49)	-0.0223 (-0.72)
Relativity	0.0192 (0.56)	0.00363 (0.21)	0.000402 (0.17)	0.00591 (0.22)	0.0116 (1.21)
Social	-0.0208 (-0.53)	0.00362 (0.20)	0.000796 (0.29)	-0.00366 (-0.11)	-0.00357 (-0.33)
Possibility	-0.0770 (-1.17)	-0.0262 (-0.86)	-0.00173 (-0.54)	-0.00127 (-0.02)	-0.0133 (-0.72)
Present	-0.0551 (-1.23)	-0.0376 (-1.73)	0.00315 (1.28)	-0.0252 (-0.67)	0.000787 (0.07)
Personal concern variables					
Home	0.0228 (0.51)	-0.0214 (-0.97)	-0.00298 (-0.84)	0.0259 (0.66)	0.0133 (0.98)
Leisure	-0.0219 (-0.38)	0.0102 (0.37)	0.00254 (0.63)	-0.0265 (-0.51)	-0.0190 (-1.14)
Money	-0.0550 (-0.61)	-0.0805 (-1.71)	0.00221 (0.39)	-0.00381 (-0.04)	-0.00765 (-0.23)
Superhost	-0.109 (-0.95)	-0.0740 (-1.26)	0.0323 ^{***} (5.32)	-0.0655 (-0.66)	0.0124 (0.33)
Log Bedrooms	0.334 (1.73)	0.0734 (0.76)	-0.00138 (-0.14)	-0.0509 (-0.31)	0.0320 (0.57)
Log Bathrooms	0.395 (1.65)	0.411 ^{**} (3.00)	0.0181 (1.46)	0.151 (0.75)	-0.00883 (-0.15)
Log Photos	0.148 (1.23)	0.0750 (1.32)	0.00704 (1.14)	0.118 (1.23)	0.0172 (0.57)
Max guests	-0.000869 (-0.02)	0.0805 ^{***} (3.38)	-0.00273 (-1.06)	-0.0377 (-0.98)	-0.00754 (-0.64)
Minimum stay	0.000661 (0.37)	-0.000419 (-0.55)	0.0000306 (0.27)	0.000917 (0.70)	-0.0000118 (-0.01)
Response rate	0.0101 (1.73)	0.00412 (1.94)	0.000517 (1.45)	0.00664 (1.71)	0.000953 (0.97)
Cancellation policy	0.00184 (0.02)	-0.0724 (-1.23)	-0.00150 (-0.23)	0.0177 (0.18)	0.0269 (0.80)
Constant	2.094 ^{**} (2.72)	4.008 ^{***} (13.95)	1.488 ^{***} (43.01)	0.0971 (0.18)	0.120 (0.85)
Effetti Fissi City	Yes	Yes	Yes	Yes	Yes
Effetti Fissi Propert	Yes	Yes	Yes	Yes	Yes
N	1154	1154	1687	1154	1782

Tabella 9: Risultati delle regressioni: impatto delle variabili semantiche sulle performance – LIWC

L'analisi delle dimensioni semantiche condotta attraverso il dizionario LIWC restituisce un quadro di sostanziale neutralità statistica. Osservando i coefficienti, nessuna delle variabili considerate, né quelle relative ai processi cognitivi o affettivi, né le categorie legate alle preoccupazioni personali come *Home*, *Leisure* o *Money*, mostra associazioni rilevanti con le performance economiche. Anche le nuove metriche introdotte in questa fase, *Present* e *Possibility*, non esercitano alcuna influenza significativa su ricavi, prezzi o valutazioni. Ciò indica che, se misurati attraverso le categorie generaliste di LIWC, i temi trattati nella descrizione non sembrano costituire, di per sé, un fattore discriminante per il successo dell'annuncio.

Il confronto con le evidenze emerse nelle analisi precedenti consente di interpretare in modo più articolato i risultati ottenuti tramite LIWC. Poiché entrambi gli approcci adottano una metodologia di tipo dictionary-based, le differenze osservate negli esiti non devono essere lette come incongruenze, bensì come conseguenza delle diverse modalità di definizione delle categorie linguistiche.

L'analisi effettuata con il software, LIWC restituisce risultati coerenti con l'utilizzo di categorie semantiche di carattere generale, mentre l'analisi basata sullo script in R si concentra su categorie maggiormente aderenti al contenuto descrittivo degli annunci. Di conseguenza, le due analisi colgono il linguaggio a livelli diversi di astrazione.

Nel loro insieme, questi risultati suggeriscono che l'impatto del linguaggio non dipende esclusivamente dalla categoria tematica considerata, ma anche dalle modalità con cui essa viene implementata nell'analisi.

5.4 Il ruolo del contesto urbano: effetti di moderazione della *urban class*

Per verificare se l'effetto delle caratteristiche testuali delle descrizioni vari in base al contesto territoriale, l'analisi è stata estesa introducendo la *urban class* come prima variabile di moderazione. Tale indicatore, di natura categoriale, distingue tre livelli di urbanizzazione dell'area in cui è localizzato l'alloggio.

Nelle regressioni stimate in Stata, la *urban class* è stata modellata tramite variabili dummy, selezionando una categoria di riferimento al fine di evitare problemi di multicollinearità. Più nello specifico, la categoria corrispondente alle aree più urbanizzate (*urban class* = 1) è assorbita nell'intercetta del modello, mentre le classi *urban class* = 2 (Urban areas) e *urban class* = 3 (Rural areas) rappresentano deviazioni rispetto a tale livello base.

L'inclusione dei termini di interazione tra *urban class* e le variabili testuali consente quindi di valutare se, e in che misura, l'impatto del linguaggio sulle performance economiche dell'annuncio differisca tra contesti caratterizzati da un diverso grado di urbanizzazione.

Cleanliness

	M1 Log RevPAN	M2 Log ADR	M3 Log Rating	M4 Log Reservation	M5 Occup. Rate
Cleanliness	-0.0025 (0.1169)	0.0221 (0.0578)	-0.0069 (0.0077)	0.0513 (0.0927)	-0.0151 (0.0342)
Urban areas	0.3910 (0.8960)	-0.1779 (0.4431)	0.0414 (0.1183)	0.2122 (0.7106)	-1.2186** (0.5272)
Rural areas	2.2797 -1.5517	0.5824 (0.7674)	0.0620 (0.0929)	-0.5051 -1.2307	-0.3704 (0.4156)
Urban areas x cleanliness	0.3599 (0.3395)	0.2685 (0.1679)	0.0048 (0.0149)	-0.4355 (0.2693)	0.0127 (0.0635)
Rural areas x cleanliness	0.4049 (0.3380)	0.3505** (0.1672)	0.0129 (0.0163)	0.0035 (0.2681)	-0.0424 (0.0675)
Log Bedrooms	0.2541** (0.1285)	0.0660 (0.0636)	-0.0018 (0.0073)	-0.0278 (0.1019)	-0.0043 (0.0312)
Log Bathrooms	0.3366** (0.1386)	0.4001*** (0.0685)	0.0249*** (0.0080)	-0.0182 (0.1099)	-0.0279 (0.0329)
Log Photos	0.1214* (0.0675)	0.0693** (0.0334)	0.0099*** (0.0038)	0.0380 (0.0536)	0.0276* (0.0162)
Max guests	0.0354 (0.0271)	0.0794*** (0.0134)	-0.0038** (0.0016)	0.0018 (0.0215)	0.0011 (0.0067)
Minimum stay	0.0010 (0.0017)	-0.0007 (0.0008)	-0.0000 (0.0001)	0.0017 (0.0013)	0.0002 (0.0005)
Response rate	0.0081*** (0.0028)	0.0028** (0.0014)	0.0006*** (0.0001)	0.0043* (0.0022)	0.0019*** (0.0005)
Cancellation policy	-0.0523 (0.0830)	0.0067 (0.0411)	-0.0142*** (0.0051)	-0.0781 (0.0658)	-0.0211 (0.0223)
Constant	1.0230 -1.3119	4.2231*** (0.6488)	1.4990*** (0.1006)	0.5301 -1.0405	0.1232 (0.4428)
Effetti Fissi City	Yes	Yes	Yes	Yes	Yes
Effetti Fissi Property	Yes	Yes	Yes	Yes	Yes
N	1148	1148	1671	1148	1765

Tabella 10: Risultati delle regressioni: effetto moderatori dell'area geografica sulla variabile pulizia – SCRIPT R

L'analisi dell'interazione tra la variabile *Cleanliness* e il contesto territoriale mette in evidenza una dinamica di mercato eterogenea.

Il punto di partenza dell'indagine è il comportamento della variabile nella categoria di riferimento, ovvero le aree suburbane. Qui, i coefficienti stimati non si discostano in modo significativo dallo zero, indicando che, nei contesti residenziali standard, descrivere minuziosamente l'igiene non genera variazioni misurabili nelle performance economiche. Lo stesso scenario si ripete nelle aree urbane, in cui l'interazione risulta neutrale e non significativa. Tale evidenza suggerisce che, nei contesti ad alta densità abitativa, l'igiene viene percepita dal mercato come un prerequisito essenziale per competere, la cui

presenza è data per scontata dall'ospite e non giustifica un incremento del valore percepito.

La situazione cambia completamente quando ci si sposta nelle aree rurali. L'interazione *Rural areas* × *Cleanliness* evidenzia infatti un coefficiente positivo e statisticamente robusto (+0.3505**) sulla variabile *Log ADR*. A differenza di quanto accade in città, nel contesto rurale la garanzia esplicita di igiene smette di essere un dettaglio irrilevante e si trasforma in uno strumento strategico per incrementare le tariffe. Economicamente, ciò implica che gli host di campagna possono sfruttare la rassicurazione sulla pulizia come un segnale di qualità per ridurre l'incertezza del potenziale guest e sostenere così un prezzo medio giornaliero più elevato.

È importante però notare che tale effetto è selettivo, infatti agisce sul prezzo (*ADR*) ma non sui volumi (*Occupancy, Reservations*). La pulizia in campagna, dunque, non serve ad attrarre più turisti, non aumenta la domanda, ma permette di estrarre un valore marginale superiore dalla transazione da ogni singolo ospite.

Safety

	M1 Log RevPAN	M2 Log ADR	M3 Log Rating	M4 Log Reservation	M5 Occup. Rate
Safety	0.1415 (0.1239)	0.0888 (0.0614)	0.0056 (0.0079)	-0.0939 (0.0983)	-0.0328 (0.0344)
Urban areas	0.5219 (0.8996)	-0.0829 (0.4458)	0.0118 (0.1198)	0.1125 (0.7133)	-1.1004** (0.5334)
Rural areas	2.1365 -1.5477	0.4565 (0.7669)	0.0656 (0.0926)	-0.5321 -1.2272	-0.3941 (0.4147)
Urban areas x safety	-0.3357 (0.2493)	-0.1968 (0.1235)	-0.0116 (0.0132)	0.2229 (0.1977)	0.0058 (0.0551)
Rural areas x safety	-0.1080 (0.2387)	-0.2545** (0.1183)	-0.0266* (0.0138)	0.2382 (0.1893)	0.1120* (0.0606)
Log Bedrooms	0.2425* (0.1286)	0.0591 (0.0637)	-0.0017 (0.0073)	-0.0304 (0.1019)	-0.0041 (0.0311)
Log Bathrooms	0.3590** (0.1391)	0.4069*** (0.0689)	0.0246*** (0.0080)	-0.0167 (0.1103)	-0.0265 (0.0329)
Log Photos	0.1287* (0.0671)	0.0805** (0.0333)	0.0096** (0.0038)	0.0336 (0.0532)	0.0259 (0.0161)
Max guests	0.0344 (0.0270)	0.0786*** (0.0134)	-0.0038** (0.0016)	0.0031 (0.0214)	0.0013 (0.0067)
Minimum stay	0.0010 (0.0017)	-0.0007 (0.0008)	-0.0000 (0.0001)	0.0016 (0.0013)	0.0002 (0.0005)
Response rate	0.0081*** (0.0028)	0.0027* (0.0014)	0.0006*** (0.0001)	0.0046** (0.0022)	0.0019*** (0.0005)
Cancellation policy	-0.0473 (0.0831)	0.0065 (0.0412)	-0.0142*** (0.0051)	-0.0744 (0.0659)	-0.0204 (0.0222)
Constant	0.9686 -1.3125	4.1931*** (0.6504)	1.4971*** (0.1004)	0.5217 -1.0407	0.1263 (0.4423)
Effetti Fissi City	Yes	Yes	Yes	Yes	Yes
Effetti Fissi Property	Yes	Yes	Yes	Yes	Yes
N	1148	1148	1672	1148	1766

Tabella 11: Risultati delle regressioni: effetto moderatori dell'area geografica sulla variabile sicurezza – SCRIPT R

Diversamente da quanto osservato per la pulizia, l'analisi della dimensione *Safety* non restituisce un quadro di valorizzazione dell'offerta, ma evidenzia criticità specifiche legate al contesto territoriale. Nelle aree suburbane, che fungono da categoria di riferimento, così come nei contesti urbani, l'analisi non rileva alcuna associazione significativa tra l'enfasi sulla sicurezza e le metriche di performance, suggerendo che in questi scenari il mercato consideri tali dotazioni come requisiti standard, incapaci di generare valore aggiunto.

Tuttavia, quando l'analisi si estende alle aree rurali, lo scenario cambia andando in direzione opposta. Qui l'interazione statistica evidenzia un impatto negativo sia sul prezzo

medio (-0.2545**) sia sulla valutazione complessiva (-0.0266**). Tale risultato indica che, lontano dai centri urbani, il richiamo esplicito a dotazioni di sicurezza non viene percepito come un servizio, ma tende piuttosto ad attivare un senso di allarme o di eccessivo isolamento, finendo per ridurre la disponibilità a pagare dell'ospite. Ne consegue che, per gli host rurali, la strategia comunicativa più efficace consistente nell'evitare questi riferimenti tecnici, orientando piuttosto la narrazione verso la qualità dell'esperienza e il comfort per non insinuare nell'ospite una percezione di insicurezza.

Leisure

	M1 Log RevPAN	M2 Log ADR	M3 Log Rating	M4 Log Reservation	M5 Occup. Rate
Leisure	0.0358 (0.0647)	0.0095 (0.0323)	-0.0006 (0.0038)	-0.0151 (0.0516)	0.0140 (0.0161)
Urban areas	0.5228 (0.9505)	-0.2647 (0.4750)	0.0670 (0.1211)	-0.0937 (0.7584)	-1.0842** (0.5371)
Rural areas	1.8546 -1.5460	0.4036 (0.7726)	0.0645 (0.0928)	-0.6285 -1.2335	-0.3858 (0.4148)
Urban areas x leisure	-0.0474 (0.1315)	0.0436 (0.0657)	0.0005 (0.0054)	0.0995 (0.1049)	-0.0298 (0.0226)
Rural areas x leisure	0.3172** (0.1334)	0.0545 (0.0667)	0.0055 (0.0064)	0.0874 (0.1065)	0.0095 (0.0266)
Log Bedrooms	0.2450* (0.1278)	0.0639 (0.0639)	-0.0021 (0.0073)	-0.0332 (0.1020)	-0.0036 (0.0311)
Log Bathrooms	0.3546** (0.1377)	0.4105*** (0.0688)	0.0251*** (0.0080)	-0.0139 (0.1099)	-0.0272 (0.0329)
Log Photos	0.1263* (0.0670)	0.0767** (0.0335)	0.0097** (0.0038)	0.0275 (0.0534)	0.0280* (0.0162)
Max guests	0.0361 (0.0269)	0.0781*** (0.0135)	-0.0038** (0.0016)	0.0030 (0.0215)	0.0011 (0.0067)
Minimum stay	0.0009 (0.0017)	-0.0007 (0.0008)	-0.0000 (0.0001)	0.0017 (0.0013)	0.0002 (0.0005)
Response rate	0.0080*** (0.0028)	0.0026* (0.0014)	0.0006*** (0.0001)	0.0046** (0.0022)	0.0019*** (0.0005)
Cancellation policy	-0.0492 (0.0825)	0.0108 (0.0412)	-0.0140*** (0.0051)	-0.0812 (0.0658)	-0.0218 (0.0222)
Constant	1.0861 -1.3102	4.2161*** (0.6548)	1.4982*** (0.1006)	0.5918 -1.0454	0.1354 (0.4427)
Effetti Fissi City	Yes	Yes	Yes	Yes	Yes
Effetti Fissi Property	Yes	Yes	Yes	Yes	Yes
N	1148	1148	1672	1148	1766

Tabella 12: Risultati delle regressioni: effetto moderatori dell'area geografica sulla variabile tempo libero – SCRIPT R

L'analisi della dimensione *Leisure* introduce una variazione rispetto alle metriche precedenti, spostando l'impatto economico dal prezzo unitario alla redditività

complessiva dell'annuncio. I dati evidenziano infatti un'unica evidenza statistica riferita alle aree rurali ed infatti, l'interazione *Rural areas* × *Leisure* genera un effetto positivo e significativo sulla variabile *Log RevPAN* (0.3172**).

È interessante osservare che tale vantaggio non derivi da un aumento delle tariffe giornaliere (l'ADR rimane invariato), ma da una maggiore capacità dell'annuncio di generare ricavi per notte disponibile. Ciò suggerisce che, in campagna, la narrazione focalizzata su relax e svago agisca come un driver di attrattività, non serve a vendere a un prezzo più alto, ma rende l'offerta più desiderabile, massimizzando la redditività dell'alloggio.

Nei contesti urbani questa strategia non funziona. L'host di città non deve perdere tempo a descrivere il tempo libero, perché l'ospite cerca comodità e servizi esterni. L'host di campagna, invece, deve mettere proprio le attività ricreative al centro del suo annuncio se vuole ottenere il massimo profitto.

Professional orientation

	M1 Log RevPAN	M2 Log ADR	M3 Log Rating	M4 Log Reservation	M5 Occup. Rate
Professional orientation	-0.1250 (0.0800)	-0.0474 (0.0398)	0.0013 (0.0055)	0.0964 (0.0633)	0.0185 (0.0242)
Urban areas	0.4344 (0.8950)	-0.1411 (0.4449)	0.0458 (0.1178)	0.2004 (0.7086)	-1.2306** (0.5269)
Rural areas	2.0116 -1.5557	0.3540 (0.7733)	0.0657 (0.0924)	-0.4679 -1.2317	-0.3782 (0.4153)
Urban areas x prof.orientation	0.0681 (0.1433)	0.0654 (0.0712)	0.0194** (0.0090)	-0.2508** (0.1135)	-0.0364 (0.0385)
Rural areas x prof. orientation	0.2308 (0.2009)	0.1420 (0.0999)	0.0057 (0.0117)	-0.1198 (0.1590)	-0.0326 (0.0492)
Log Bedrooms	0.2573** (0.1291)	0.0699 (0.0642)	-0.0002 (0.0073)	-0.0589 (0.1022)	-0.0053 (0.0313)
Log Bathrooms	0.3349** (0.1385)	0.4043*** (0.0689)	0.0259*** (0.0080)	-0.0128 (0.1097)	-0.0276 (0.0330)
Log Photos	0.1225* (0.0672)	0.0765** (0.0334)	0.0101*** (0.0038)	0.0359 (0.0532)	0.0270* (0.0161)
Max guests	0.0347 (0.0271)	0.0776*** (0.0135)	-0.0042*** (0.0016)	0.0066 (0.0215)	0.0013 (0.0067)
Minimum stay	0.0009 (0.0017)	-0.0007 (0.0008)	-0.0000 (0.0001)	0.0017 (0.0013)	0.0002 (0.0005)
Response rate	0.0078*** (0.0028)	0.0026* (0.0014)	0.0006*** (0.0001)	0.0047** (0.0022)	0.0019*** (0.0005)
Cancellation policy	-0.0453 (0.0829)	0.0120 (0.0412)	-0.0144*** (0.0051)	-0.0781 (0.0656)	-0.0209 (0.0223)
Constant	1.1309 -1.3265	4.3089*** (0.6594)	1.5087*** (0.1017)	0.4472 -1.0503	0.0958 (0.4467)
Effetti Fissi City	Yes	Yes	Yes	Yes	Yes
Effetti Fissi Property	Yes	Yes	Yes	Yes	Yes
N	1148	1148	1672	1148	1766

Tabella 13: Risultati delle regressioni: effetto moderatori dell'area geografica sulla variabile orientamento linguistico professionale – SCRIPT R

L'analisi dell'orientamento professionale mostra un risultato molto particolare che riguarda solo le città. Qui notiamo un effetto doppio e opposto perchè da una parte, usare un linguaggio molto professionale fa alzare il punteggio delle recensioni (+0.0194**), segno che gli ospiti apprezzano la serietà del servizio. Dall'altra parte, però, questo stesso stile riduce in modo significativo il numero di prenotazioni (-0.2508**).

Questo indica che un tono troppo formale, anche se trasmette affidabilità, rischia di rendere l'annuncio meno invitante, allontanando i potenziali guest in fase di ricerca. Tuttavia, chi supera questa barriera iniziale finisce per premiare l'organizzazione dell'host con valutazioni elevate. Nelle aree rurali e suburbane, invece, questo problema non esiste

e i risultati rimangono neutri. Solo chi opera in città deve fare attenzione a bilanciare la professionalità con un tocco più umano per non perdere clienti.

Perceptual

	M1 Log RevPAN	M2 Log ADR	M3 Log Rating	M4 Log Reservation	M5 Occup. Rate
Perceptual	-0.068 (0.096)	0.055 (0.044)	-0.002 (0.006)	-0.008 (0.072)	-0.040 (0.032)
Urban areas	0.000 (.)	0.000 (.)	-0.011 (0.011)	-1.649*** (0.116)	-0.919*** (0.082)
Urban areas x Perceptual	0.232 (0.152)	-0.181** (0.082)	0.005 (0.009)	0.050 (0.100)	0.091 (0.064)
Rural areas x Perceptual	-0.298* (0.156)	-0.200 (0.125)	-0.003 (0.011)	-0.075 (0.118)	0.001 (0.052)
Superhost	-0.101 (0.106)	-0.107* (0.055)	0.025*** (0.005)	-0.031 (0.082)	0.015 (0.036)
Log Bedrooms	0.350* (0.209)	0.171 (0.120)	-0.004 (0.010)	-0.094 (0.143)	0.003 (0.064)
Log Bathrooms	0.507 (0.366)	0.582*** (0.219)	0.017 (0.017)	-0.081 (0.210)	0.003 (0.088)
Log Photos	0.192* (0.111)	0.075 (0.057)	0.008* (0.005)	0.062 (0.072)	0.018 (0.030)
Max guests	0.017 (0.045)	0.075*** (0.023)	-0.002 (0.002)	0.022 (0.028)	-0.004 (0.011)
Minimum stay	0.001 (0.002)	-0.001 (0.001)	0.000 (0.000)	-0.000 (0.003)	0.000 (0.001)
Response rate	0.011** (0.005)	0.005*** (0.002)	0.000 (0.000)	0.006*** (0.002)	0.002** (0.001)
Cancellation policy	0.009 (0.107)	-0.071 (0.055)	-0.001 (0.005)	0.119 (0.076)	0.033 (0.032)
Constant	1.332** (0.629)	3.243*** (0.272)	1.684*** (0.026)	0.639** (0.283)	0.352*** (0.125)
Effetti Fissi City	Yes	Yes	Yes	Yes	Yes
Effetti Fissi Property	Yes	Yes	Yes	Yes	Yes
N	1225	1225	1790	1893	1893

Tabella 14: Risultati delle regressioni: effetto moderatori dell'area geografica sulla percezione – LIWC

L'analisi della variabile *Perceptual* mostra un risultato inaspettatamente negativo. Se nelle aree suburbane di riferimento l'effetto è nullo, negli altri contesti l'uso di questo linguaggio diventa controproducente.

Nelle aree urbane, l'interazione è negativa sul prezzo: chi insiste troppo su descrizioni sensoriali tende ad avere un prezzo medio più basso (-0.181**). Nelle aree rurali, il calo colpisce direttamente i ricavi complessivi (*RevPAN*), con un coefficiente negativo significativo (-0.298*).

Questo suggerisce che il mercato turistico attuale preferisce la concretezza all'astrazione. In città, dove contano posizione e funzionalità, le descrizioni troppo evocative vengono probabilmente percepite come superflue. Anche in campagna, dove ci si aspetterebbe più attenzione all'atmosfera, l'eccesso di parole "percettive" sembra confondere o non convincere l'ospite, che forse cerca informazioni più pratiche sul valore dell'alloggio. Gli host dovrebbero usare con molta cautela questo tipo di linguaggio, poiché i dati confermano che non porta vantaggi economici, ma rischia anzi di ridurre il valore dell'offerta.

5.5 Il ruolo dell'host: effetti di moderazione della variabile Superhost

L'analisi è stata estesa considerando lo status di Superhost come ulteriore variabile di moderazione, al fine di verificare se l'impatto delle caratteristiche linguistiche delle descrizioni vari in funzione della reputazione dell'host. In Stata, la variabile *Superhost* è specificata come dummy binaria, assumendo valore pari a 1 nel caso di host che possiedono lo status di Superhost e valore pari a 0 in caso contrario. Analogamente a quanto fatto per la variabile *urban_class*, la categoria con valore 0, corrispondente agli host che non possiedono il badge di qualità, è utilizzata come categoria di riferimento ed è assorbita nell'intercetta del modello, mentre i coefficienti associati alla dummy e ai termini di interazione consentono di valutare eventuali differenze negli effetti delle variabili linguistiche tra Host e Superhost.

Affective

	M1 Log RevPAN	M2 Log ADR	M3 Log Rating	M4 Log Reservation	M5 Occup. Rate
Affective	0.0658 (0.0902)	-0.0113 (0.0441)	-0.0035 (0.0052)	0.1147 (0.0709)	0.0268 (0.0222)
Superhost	-0.1313 (0.0902)	-0.1561*** (0.0441)	0.0324*** (0.0054)	-0.1033 (0.0709)	0.0395 (0.0241)
Superhost x affective	0.0187 (0.1345)	0.1768*** (0.0658)	-0.0050 (0.0079)	-0.0784 (0.1057)	-0.0476 (0.0351)
Log Bedrooms	0.2498* (0.1289)	0.0619 (0.0630)	-0.0031 (0.0070)	-0.0195 (0.1013)	-0.0081 (0.0310)
Log Bathrooms	0.3174** (0.1390)	0.3870*** (0.0680)	0.0240*** (0.0077)	-0.0250 (0.1093)	-0.0273 (0.0328)
Log Photos	0.1358** (0.0672)	0.0789** (0.0329)	0.0086** (0.0037)	0.0400 (0.0528)	0.0266* (0.0161)
Max guests	0.0373 (0.0272)	0.0820*** (0.0133)	-0.0035** (0.0015)	0.0012 (0.0214)	0.0013 (0.0067)
Minimum stay	0.0010 (0.0017)	-0.0008 (0.0008)	0.0000 (0.0001)	0.0017 (0.0013)	0.0003 (0.0005)
Response rate	0.0086*** (0.0028)	0.0032** (0.0014)	0.0004*** (0.0001)	0.0056** (0.0022)	0.0017*** (0.0005)
Cancellation policy	-0.0353 (0.0831)	0.0173 (0.0406)	-0.0127** (0.0049)	-0.0819 (0.0653)	-0.0209 (0.0222)
Constant	0.7311 -13241	3.9867*** (0.6474)	1.6137*** (0.0973)	0.2553 -10408	-0.2454 (0.4411)
Effetti Fissi City	Yes	Yes	Yes	Yes	Yes
Effetti Fissi Property	Yes	Yes	Yes	Yes	Yes
N	1154	1154	1687	1154	1782

Tabella 15: Risultati delle regressioni: effetto moderatori dello status Superhost sul linguaggio affettivo – SCRIPT R

L'efficacia del linguaggio emotivo non è uguale per tutti, ma dipende strettamente dalla reputazione dell'host. Per gli host standard, infatti, l'uso di termini affettivi non porta alcun risultato significativo perché il mercato non premia né penalizza questo stile se non c'è ancora una solida credibilità alle spalle.

La situazione cambia quando l'annuncio è gestito da un Superhost. Sebbene lo status di per sé tenda ad abbassare il prezzo base (-0.1561***) pur garantendo ottime recensioni (+0.0324***), l'interazione con il linguaggio affettivo ribalta l'effetto economico e fa salire il prezzo in modo netto (+0.1768***).

Questo andamento indica che l'emozione si traduce in valore economico solo se sostenuta dalla fiducia. L'ospite è disposto a pagare un extra per la promessa di un'esperienza "speciale" solo se si fida di chi la propone. Di conseguenza, la strategia migliore per i

Superhost è puntare sulle emozioni per valorizzare l'offerta, mentre chi non ha ancora questo distintivo dovrebbe concentrarsi su informazioni concrete, poiché le parole suggestive, senza la "garanzia" dello status, non convincono l'acquirente a spendere di più.

La *Figura 20* rappresenta graficamente l'effetto del contenuto affettivo delle descrizioni sull'ADR, moderando lo status di Superhost.

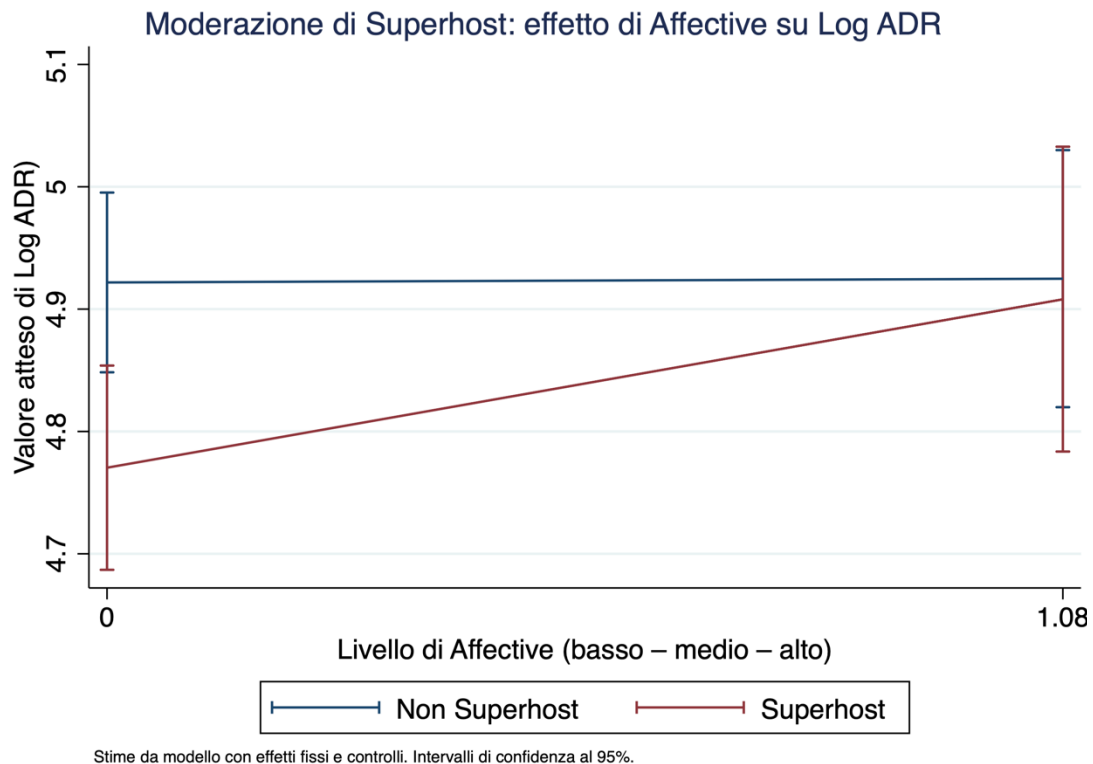


Figura 20: Moderazione dello status di Superhost: effetto di Affective su LogADR - LIWC

Esso evidenzia come l'effetto del contenuto affettivo della descrizione sull'ADR vari in funzione della reputazione dell'host. In particolare, tra i Superhost un aumento del linguaggio affettivo è associato a un incremento della disponibilità a pagare, mentre tra i non Superhost l'ADR stimato rimane sostanzialmente invariato. Questo risultato conferma quanto emerso dall'analisi di regressione, mostrando che la componente emotiva della comunicazione risulta efficace solo quando supportata da un segnale reputazionale credibile.

Host orientation

	M1	M2	M3	M4	M5
	Log RevPAN	Log ADR	Log Rating	Log Reservation	Occup. Rate
Host orientation	0.1211 (0.1973)	-0.0180 (0.0969)	0.0000 (0.0093)	-0.1675 (0.1559)	-0.0247 (0.0425)
Superhost	-0.0776 (0.0769)	-0.0638* (0.0377)	0.0301*** (0.0046)	-0.1411** (0.0607)	0.0191 (0.0207)
Superhost x host orientation	-0.4962** (0.2476)	-0.2390** (0.1215)	0.0042 (0.0134)	0.1336 (0.1956)	0.0317 (0.0610)
Log Bedrooms	0.2628** (0.1287)	0.0797 (0.0632)	-0.0032 (0.0070)	-0.0143 (0.1017)	-0.0075 (0.0311)
Log Bathrooms	0.3339** (0.1383)	0.4081*** (0.0679)	0.0235*** (0.0077)	-0.0255 (0.1093)	-0.0287 (0.0328)
Log Photos	0.1264* (0.0676)	0.0709** (0.0332)	0.0084** (0.0037)	0.0349 (0.0534)	0.0258 (0.0162)
Max guests	0.0305 (0.0271)	0.0737*** (0.0133)	-0.0033** (0.0015)	-0.0013 (0.0214)	0.0016 (0.0067)
Minimum stay	0.0007 (0.0017)	-0.0009 (0.0008)	0.0000 (0.0001)	0.0016 (0.0013)	0.0003 (0.0005)
Response rate	0.0086*** (0.0028)	0.0032** (0.0014)	0.0004*** (0.0001)	0.0052** (0.0022)	0.0017*** (0.0005)
Cancellation policy	-0.0174 (0.0832)	0.0223 (0.0408)	-0.0127** (0.0050)	-0.0845 (0.0657)	-0.0204 (0.0223)
Constant	0.8128 -13191	4.0711*** (0.6476)	1.6108*** (0.0974)	0.2895 -10423	-0.2365 (0.4413)
Effetti Fissi City	Yes	Yes	Yes	Yes	Yes
Effetti Fissi Property	Yes	Yes	Yes	Yes	Yes
N	1154	1154	1687	1154	1782

Tabella 16: Risultati delle regressioni: effetto moderatori dello status Superhost sull'orientamento all'host – SCRIPT R

L'impatto della *Host Orientation* cambia in modo netto in base alla reputazione dell'host. Tra gli host standard, questo stile narrativo risulta sostanzialmente neutro, senza mostrare interazioni significative con la domanda o i profitti. La dinamica appare diversa nel caso dei Superhost. I dati evidenziano che, quando la comunicazione enfatizza troppo la figura del proprietario, si verifica un calo significativo sia dei ricavi complessivi (-0.4962**) sia del prezzo medio (-0.2390**).

Questo risultato suggerisce che, una volta raggiunta una reputazione consolidata, l'autoreferenzialità diventa un difetto agli occhi dell'ospite. Chi prenota da un Superhost si aspetta probabilmente professionalità e attenzione all'esperienza di soggiorno. Trovare un annuncio incentrato sull'ego del proprietario entra in contrasto con gli standard previsti, abbassando il valore percepito dell'offerta. Di conseguenza, l'approccio più

efficace per coloro i quali possiedono il badge di qualità, è quello di puntare su una narrazione meno incentrata sulla persona.

Guest orientation

	M1 Log RevPAN	M2 Log ADR	M3 Log Rating	M4 Log Reservation	M5 Occup. Rate
Guest orientation	0.1660* (0.0863)	0.0299 (0.0427)	-0.0004 (0.0054)	0.1087 (0.0680)	-0.0050 (0.0202)
Superhost	-0.0315 (0.0870)	-0.0763* (0.0430)	0.0296*** (0.0053)	-0.1185* (0.0686)	0.0335 (0.0234)
Superhost x guest orientation	-0.2669** (0.1303)	-0.0370 (0.0644)	0.0027 (0.0083)	-0.0438 (0.1027)	-0.0361 (0.0354)
Log Bedrooms	0.2625** (0.1286)	0.0699 (0.0636)	-0.0031 (0.0070)	-0.0190 (0.1013)	-0.0068 (0.0311)
Log Bathrooms	0.3340** (0.1385)	0.4009*** (0.0685)	0.0235*** (0.0077)	-0.0253 (0.1092)	-0.0280 (0.0328)
Log Photos	0.1370** (0.0671)	0.0823** (0.0332)	0.0082** (0.0037)	0.0358 (0.0529)	0.0273* (0.0161)
Max guests	0.0327 (0.0271)	0.0774*** (0.0134)	-0.0032** (0.0015)	0.0022 (0.0214)	0.0009 (0.0067)
Minimum stay	0.0010 (0.0017)	-0.0008 (0.0008)	0.0000 (0.0001)	0.0016 (0.0013)	0.0003 (0.0005)
Response rate	0.0084*** (0.0028)	0.0031** (0.0014)	0.0004*** (0.0001)	0.0053** (0.0022)	0.0017*** (0.0005)
Cancellation policy	-0.0309 (0.0828)	0.0141 (0.0410)	-0.0126** (0.0049)	-0.0787 (0.0653)	-0.0188 (0.0221)
Constant	0.7029 -13210	4.0117*** (0.6532)	1.6103*** (0.0974)	0.2067 -10410	-0.2352 (0.4410)
Effetti Fissi City	Yes	Yes	Yes	Yes	Yes
Effetti Fissi Property	Yes	Yes	Yes	Yes	Yes
N	1154	1154	1687	1154	1782

Tabella 17: Risultati delle regressioni: effetto moderatori dello status Superhost sull'orientamento al guest – SCRIPT R

L'attenzione diretta verso il potenziale guest segue una direzione opposta a quella analizzata tramite il grafico precedente, rivelandosi una leva strategica soprattutto per chi deve ancora consolidare il proprio capitale reputazionale. Per gli host standard, infatti, impostare il discorso mettendo l'ospite al centro è una mossa vincente che porta ricavi concreti (+0.1660*) perché in mancanza di un distintivo che garantisca la qualità, è proprio il dialogo diretto a costruire quella fiducia necessaria per convincere a prenotare. Tuttavia, questa efficacia sembra svanire una volta raggiunto lo status di Superhost. I dati mostrano che, per loro, una narrazione incentrata sul potenziale guest non aggiunge

valore, anzi, il termine di interazione *Superhost* × *guest orientation* risulta negativo e significativo sul *RevPAN* (-0.2669^*) e va praticamente ad annullare il beneficio iniziale.

La *guest orientation* agisce come un motore di avviamento. È fondamentale nelle prime fasi per compensare la mancanza di reputazione, ma diventa superflua quando si ha già il badge di Superhost. A quel punto, la fiducia è già garantita dallo status e non serve più "costruirla" con le parole.

Il grafico sottostante rappresenta graficamente l'effetto della *guest orientation* delle descrizioni sull'*ADR*, moderando lo status di Superhost.

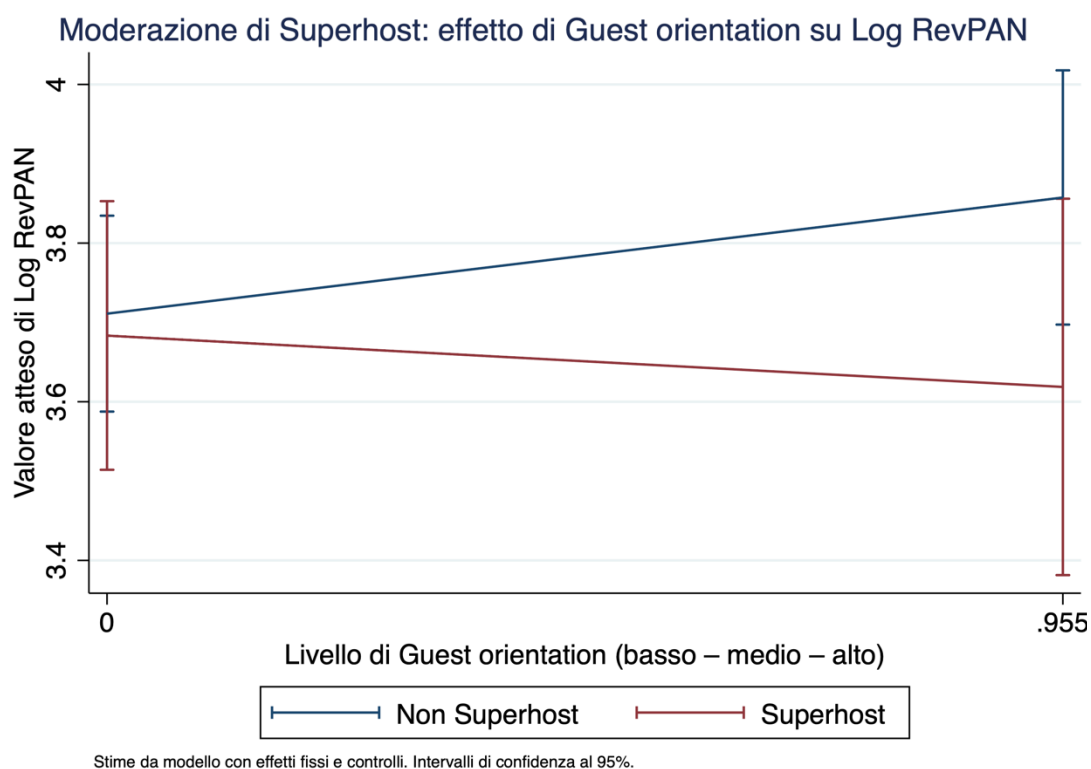


Figura 21: Moderazione dello status di Superhost: effetto di Guest orientation su LogRevPAN - LIWC

Il grafico di moderazione in *Figura 21* mostra che l'effetto della *guest orientation* sul *RevPAN* varia in funzione della reputazione dell'host. Più nello specifico, tra i non Superhost un linguaggio maggiormente orientato all'ospite è associato a un incremento della performance economica complessiva, mentre tra i Superhost l'effetto risulta di segno opposto. Questo risultato risulta in linea con le evidenze emerse dall'analisi di regressione, indicando che, in assenza di un segnale reputazionale forte, la comunicazione orientata al guest può svolgere una funzione sostitutiva, mentre per gli host con reputazione consolidata la sua efficacia appare più limitata.

CAPITOLO VI – Conclusioni

6.1 Sintesi dei principali risultati

L'obiettivo principale di questa ricerca era analizzare se e in che misura le caratteristiche linguistiche delle descrizioni degli annunci Airbnb siano associate alle performance economiche degli alloggi, una volta tenuto conto delle principali determinanti strutturali, reputazionali e di localizzazione. I risultati ottenuti suggeriscono che le scelte comunicative adottate dagli host non svolgono un ruolo puramente descrittivo, ma rappresentano una dimensione rilevante del posizionamento competitivo dell'annuncio sulla piattaforma.

In primo luogo, le analisi mostrano che le metriche legate alla lunghezza e alla complessità formale del testo non risultano sistematicamente associate a un miglioramento delle performance economiche. Descrizioni più lunghe o linguisticamente più articolate non si traducono automaticamente in prezzi più elevati, maggiore occupazione o un numero superiore di prenotazioni. Questo risultato suggerisce che un maggiore sforzo comunicativo, inteso in termini quantitativi, non costituisce di per sé un vantaggio competitivo, confermando l'idea che gli utenti privilegino testi chiari e focalizzati piuttosto che particolarmente estesi o complessi.

In senso opposto, emergono evidenze più rilevanti per quanto riguarda lo stile relazionale della comunicazione. L'uso di un linguaggio fortemente autoreferenziale, caratterizzato da un'elevata presenza della prima persona singolare, risulta associato a performance economiche più deboli. Emerge che una comunicazione centrata sull'host possa ridurre la capacità dell'annuncio di influenzare le decisioni del potenziale guest, spostando l'attenzione lontano dai bisogni e dalle aspettative. Al contrario, uno stile comunicativo maggiormente orientato al guest appare più coerente con le logiche della piattaforma e con il processo decisionale degli utenti.

L'analisi delle metriche semantiche evidenzia che il contenuto tematico delle descrizioni contribuisce a differenziare gli annunci all'interno del mercato. Sebbene molte categorie semantiche presentino valori medi contenuti, indicando un uso sporadico di specifici temi, i risultati suggeriscono che la scelta di enfatizzare determinati aspetti dell'esperienza di soggiorno possa incidere sulle performance economiche. In particolare,

emerge come la comunicazione focalizzata esclusivamente sulle caratteristiche fisiche dell'alloggio rappresenti la strategia prevalente, lasciando spazio a potenziali opportunità di differenziazione attraverso una narrazione più esperienziale.

I risultati confermano che la dimensione linguistica rappresenta una dimensione rilevante per il posizionamento degli annunci, il cui impatto si manifesta non tanto attraverso la quantità o la sofisticazione del testo, quanto attraverso il tono comunicativo e il focus del messaggio. In questo senso, la qualità della comunicazione risulta più rilevante della sua lunghezza, suggerendo che le descrizioni efficaci siano quelle in grado di allinearsi alle aspettative informative ed emotive degli utenti della piattaforma.

6.2 Implicazioni teoriche

I risultati della presente ricerca contribuiscono alla letteratura sulla sharing economy e sulla comunicazione digitale offrendo una lettura più articolata del ruolo del linguaggio nella creazione di valore economico sulle piattaforme peer-to-peer.

Lo studio mette in discussione alcune assunzioni implicite diffuse nella letteratura esistente, che tendono a interpretare la comunicazione testuale come un segnale informativo di natura prevalentemente quantitativa.

Le evidenze mostrano infatti che gli aspetti formali della comunicazione testuale non risultano sistematicamente associati a migliori performance economiche. Questo risultato suggerisce che il valore informativo del linguaggio non derivi dall'intensità dello sforzo comunicativo, ma dalla funzione che il testo svolge nel processo decisionale dell'utente. In tal senso, la ricerca contribuisce a spostare l'attenzione dalla quantità di informazione alla sua rilevanza economica, mettendo in discussione l'idea secondo cui descrizioni più estese o sofisticate costituiscano necessariamente segnali di qualità.

Un ulteriore elemento di riflessione riguarda il ruolo dello stile relazionale nella comunicazione dell'annuncio. I risultati indicano che un linguaggio fortemente autoreferenziale risulta meno coerente con le logiche di intermediazione delle piattaforme peer-to-peer, mentre uno stile orientato all'esperienza dell'ospite risulta più coerente con le logiche di funzionamento delle piattaforme peer-to-peer. Questo aspetto rafforza le interpretazioni che vedono il linguaggio non come uno strumento di auto-presentazione dell'host, ma come un meccanismo di allineamento tra le aspettative dell'utente e la rappresentazione dell'offerta in contesti caratterizzati da elevata asimmetria informativa.

L'analisi evidenzia inoltre come l'efficacia delle strategie comunicative non sia uniforme, ma vari in funzione del contesto territoriale e della reputazione dell'host.

Il linguaggio emerge quindi come un meccanismo, il cui impatto varia in funzione dell'interazione con altri segnali presenti sulla piattaforma, come la localizzazione o lo status reputazionale. Questa evidenza contribuisce a superare una visione generalista della comunicazione digitale, suggerendo un approccio più orientato allo studio dei meccanismi di creazione del valore nelle piattaforme di sharing economy.

Infine, il confronto tra diverse metodologie di analisi linguistica offre uno spunto di riflessione di natura teorico-metodologica. Le differenze osservate tra i risultati ottenuti mediante dizionari standardizzati e quelli basati su dizionari costruiti ad hoc indicano che l'impatto del linguaggio sulle performance economiche dipende in modo critico dall'adeguatezza dello strumento di analisi all'ambito economico analizzato.

Nel complesso dei risultati, la ricerca contribuisce a una comprensione più sfumata del ruolo della comunicazione nelle piattaforme di sharing economy, mostrando che il linguaggio incide sulle performance economiche non in modo diretto e uniforme, ma attraverso meccanismi selettivi, contestuali e complementari ad altri segnali informativi.

6.3 Implicazioni manageriali

Le evidenze emerse dall'analisi suggeriscono che la comunicazione testuale rappresenti una risorsa strategica per gli host, ma non nel senso di un semplice aumento dello sforzo descrittivo. Dal punto di vista manageriale, il primo risultato rilevante riguarda infatti ciò che non risulta efficace ed in particolare che investire tempo ed energie nella produzione di descrizioni più lunghe o linguisticamente più complesse non sembra tradursi in un miglioramento delle performance economiche. Questo implica che l'ottimizzazione dell'annuncio non dovrebbe concentrarsi sulla quantità di informazioni fornite, bensì sulla loro selezione e organizzazione.

In termini operativi, gli host possono trarre beneficio da una comunicazione più mirata, orientata a chiarire pochi elementi chiave dell'offerta piuttosto che a fornire descrizioni esaustive. La capacità dell'annuncio di orientare la scelta dell'ospite sembra dipendere soprattutto dal tono comunicativo e dal focus del messaggio, ovvero da come l'alloggio e l'esperienza di soggiorno vengono presentati. Ne deriva che una comunicazione efficace non richiede necessariamente competenze linguistiche avanzate, ma una maggiore attenzione alla coerenza tra messaggio, aspettative dell'utente e contesto dell'offerta.

Un'ulteriore implicazione riguarda la necessità di adattare la strategia comunicativa al contesto territoriale. In un mercato eterogeneo come quello italiano, caratterizzato dalla coesistenza di contesti urbani, turistici e rurali, l'efficacia del linguaggio non è uniforme. Di conseguenza, gli host dovrebbero evitare approcci standardizzati e considerare il contesto in cui opera l'alloggio come un elemento centrale nella costruzione della descrizione. Strategie comunicative efficaci in aree urbane ad alta densità di offerta potrebbero non produrre gli stessi effetti in contesti periferici o rurali, dove le aspettative degli ospiti e le motivazioni di scelta risultano differenti.

Inoltre, i risultati suggeriscono che le strategie comunicative dovrebbero tener conto anche del profilo dell'host. Gli host non professionali, spesso caratterizzati da risorse limitate e da una gestione meno strutturata, possono trarre vantaggio da descrizioni semplici e focalizzate sull'esperienza di soggiorno, evitando di imitare modelli comunicativi tipici di operatori più professionali. Al contrario, per host più strutturati e con una reputazione consolidata, la comunicazione può svolgere una funzione di rassicurazione e coerenza, più che di persuasione diretta. In questo senso, la comunicazione testuale non rappresenta una leva universale, ma uno strumento da calibrare in funzione delle caratteristiche dell'host e del mercato di riferimento.

6.4 Limiti della ricerca e sviluppi futuri

Nonostante i risultati ottenuti forniscano indicazioni rilevanti sul ruolo della comunicazione testuale nelle performance degli annunci Airbnb, la presente ricerca presenta alcuni limiti che devono essere considerati nell'interpretazione delle evidenze empiriche.

Un primo elemento riguarda la dimensione e la composizione del campione analizzato. Sebbene l'analisi parta da un dataset molto ampio, la necessità di acquisire manualmente le descrizioni testuali ha reso necessario lavorare su un sotto-campione di annunci. Questo approccio, pur consentendo un'analisi approfondita del contenuto testuale, limita la possibilità di generalizzare pienamente i risultati all'intero mercato. Studi futuri potrebbero superare questo vincolo ampliando il campione attraverso procedure di raccolta automatizzata su larga scala.

Un ulteriore aspetto critico è legato alla natura esplorativa dell'analisi semantica adottata. Le metriche linguistiche utilizzate, basate su dizionari e categorie lessicali predefinite, consentono di catturare pattern ricorrenti nel linguaggio, ma non permettono di cogliere

il significato contestuale o la struttura narrativa complessiva delle descrizioni. Ricerche future potrebbero integrare approcci più avanzati di Natural Language Processing, come modelli basati su tecniche di analisi del discorso, al fine di approfondire ulteriormente la relazione tra linguaggio e performance.

Il focus geografico e temporale dell'analisi rappresenta un altro elemento di delimitazione. La concentrazione sul mercato italiano e sull'anno 2024 consente di ottenere risultati coerenti e attuali, ma riduce la comparabilità con altri contesti nazionali o periodi storici. Studi successivi potrebbero adottare un approccio comparativo, analizzando le differenze nell'efficacia delle strategie linguistiche tra paesi o lungo l'arco temporale, al fine di valutare il ruolo delle variabili culturali e l'evoluzione delle preferenze degli utenti.

Pur in presenza di tali limiti, i risultati mantengono coerenza interna e definiscono con maggiore precisione il perimetro interpretativo della ricerca, indicando al contempo possibili direzioni di sviluppo per studi futuri.

Nel quadro generale, la ricerca mostra come l'analisi del linguaggio rappresenti uno strumento rilevante per comprendere i meccanismi di creazione del valore nelle piattaforme digitali, la cui efficacia dipende in modo cruciale dall'allineamento tra contenuto comunicativo, contesto di mercato e strumenti di analisi adottati.

BIBLIOGRAFIA

- Armstrong, M. (2006). Competition in two-sided markets. *The RAND Journal of Economics*, 37(3), 668–691. <https://www.jstor.org/stable/25046266>
- Ba, S., & Pavlou, P. A. (2002). Evidence of the effect of trust building technology in electronic markets: Price premiums and buyer behavior. *MIS quarterly*, 243-268. <https://doi.org/10.2307/4132332>
- Baloglu, S., Erdem, M., Brewer, P., Mayer, K., & Sainaghi, R. (2010). Hotel performance: State of the art. *International Journal of Contemporary Hospitality Management*, 22(7), 920–952. <https://doi.org/10.1108/09596111011066617>
- Belk, R. (2014). You are what you can access: Sharing and collaborative consumption online. *Journal of Business Research*, 67(8), 1595–1600. <https://doi.org/10.1016/j.jbusres.2013.10.001>
- Botsman, R., & Rogers, R. (2010). Beyond Zipcar: Collaborative consumption. *Harvard Business Review*, 88(10), 30.
- Botsman, R., & Foley, K. (2010). What's mine is yours. Tantor Media, Incorporated.
- Brochado, A., Troilo, M., & Shah, A. (2017). Airbnb customer experience: Evidence of convergence across three countries. *Annals of Tourism Research*, 63, 1–15. <https://doi.org/10.1016/j.annals.2017.01.001>
- Buhalis, D., & Law, R. (2008). Progress in information technology and tourism management. *Tourism Management*, 29(4), 609–623. <https://doi.org/10.1016/j.tourman.2008.01.005>
- Caillaud, B., & Jullien, B. (2003). Chicken & egg: Competition among intermediation service providers. *The RAND Journal of Economics*, 34(2), 309–328.
- Celata, F., & Romano, A. (2022). Overtourism and online short-term rental platforms in Italian cities. *Journal of Sustainable Tourism*, 30(5), 1020–1039. <https://doi.org/10.1080/09669582.2020.1788568>

- Chattopadhyay, M., & Mitra, S. K. (2019). Do Airbnb host listing attributes influence room pricing homogenously? *International Journal of Hospitality Management*, 81, 54–64. <https://doi.org/10.1016/j.ijhm.2019.03.008>
- Christensen, C. M. (2014). *Disruptive innovation*. Interaction Design Foundation.
- Dann, D., Teubner, T., & Weinhardt, C. (2019). Poster child and guinea pig. *International Journal of Contemporary Hospitality Management*, 31(1), 427–473. <https://doi.org/10.1108/IJCHM-03-2018-0186>
- Dredge, D., & Gyimóthy, S. (2015). The collaborative economy and tourism. *Tourism Recreation Research*, 40(3), 286–302. <https://doi.org/10.1080/02508281.2015.1086076>
- Ert, E., Fleischer, A., & Magen, N. (2016). Trust and reputation in the sharing economy. *Tourism Management*, 55, 62–73. <https://doi.org/10.1016/j.tourman.2016.01.013>
- Evans, D. S., & Schmalensee, R. (2016). *Matchmakers: The new economics of multisided platforms*. Harvard Business Review Press.
- Falk, M., Larpin, B., & Scaglione, M. (2019). The role of specific attributes in determining prices of Airbnb listings. *International Journal of Hospitality Management*, 83, 132–140. <https://doi.org/10.1016/j.ijhm.2019.04.023>
- Filieri, R. (2015). What makes online reviews helpful? *Journal of Business Research*, 68(6), 1261–1270. <https://doi.org/10.1016/j.jbusres.2014.11.006>
- Gallagher, L. (2017). *The Airbnb story*. Houghton Mifflin Harcourt.
- Gefen, D., Karahanna, E., & Straub, D. W. (2003). Trust and TAM in online shopping: An integrated model. *MIS quarterly*, 51-90. <https://doi.org/10.2307/30036519>
- Gibbs, C., Guttentag, D., Gretzel, U., Morton, J., & Tams, A. (2018). Pricing in the sharing economy. *Journal of Travel & Tourism Marketing*, 35(1), 46–56. <https://doi.org/10.1080/10548408.2017.1308292>
- Gunter, U., & Önder, I. (2018). Determinants of Airbnb demand in Vienna. *Tourism Economics*, 24(3), 270–293. <https://doi.org/10.1177/1354816617731196>

Guttentag, D. (2015). Airbnb: Disruptive innovation. *Current Issues in Tourism*, 18(12), 1192–1217. <https://doi.org/10.1080/13683500.2013.827159>

Koh, Y., Kim, J., & Vaughan, Y. (2021). How you name your Airbnb's title matters. *Journal of Travel & Tourism Marketing*, 38(1), 93–106. <https://doi.org/10.1080/10548408.2021.1875105>

Kwok, L., & Xie, K. L. (2019). Pricing strategies on Airbnb. *International Journal of Hospitality Management*, 82, 228–237. <https://doi.org/10.1016/j.ijhm.2018.09.013>

Liang, L. J., Choi, H. C., & Joppe, M. (2018). Understanding repurchase intention of Airbnb consumers. *Journal of Travel & Tourism Marketing*, 35(1), 73–89. <https://doi.org/10.1080/10548408.2016.1224750>

Liang, S., Schuckert, M., Law, R., & Chen, C. C. (2017). Be a “Superhost”. *Tourism Management*, 60, 454–465. <https://doi.org/10.1016/j.tourman.2017.01.007>

Ma, X., Hancock, J. T., Lim, K. M., & Naaman, M. (2017). Self-disclosure and perceived trustworthiness of Airbnb host profiles. *Proceedings of CSCW*, 2397–2409. <https://doi.org/10.1145/2998181.2998269>

Paul A. Pavlou, Angelika Dimoka, (2006) The Nature and Role of Feedback Text Comments in Online Marketplaces: Implications for Trust Building, Price Premiums, and Seller Differentiation. *Information Systems Research* 17(4):392-414. <https://doi.org/10.1287/isre.1060.0106>

Pennebaker, J. W., Boyd, R. L., Jordan, K., & Blackburn, K. (2015). *The development and psychometric properties of LIWC2015*.

Rochet, J. C., & Tirole, J. (2003). Platform competition in two-sided markets. *Journal of the European Economic Association*, 1(4), 990–1029. <https://doi.org/10.1162/154247603322493212>

Rochet, J. C., & Tirole, J. (2006). Two-sided markets: A progress report. *The RAND Journal of Economics*, 37(3), 645–667. <https://doi.org/10.1111/j.1756-2171.2006.tb00036.x>

Rosen, S. (1974). Hedonic prices and implicit markets. *Journal of Political Economy*, 82(1), 34–55. <https://doi.org/10.1086/260169>

Sun, X., Gui, L., & Cai, B. (2024). To be honest or positive? *Decision Support Systems*, *181*, 114200. <https://doi.org/10.1016/j.dss.2024.114200>

Sundararajan, A. (2017). The sharing economy: The end of employment and the rise of crowd-based capitalism. *MIT press*.

Ter Huurne, M., Moons, J., Ronteltap, A., Corten, R., & Buskens, V. (2021). Promoting trust through linguistic features. *International Journal of Electronic Business*, *16*(4), 417–434. <https://doi.org/10.1504/IJEB.2021.118501>

Tussyadiah, I. P., & Park, S. (2018). When guests trust hosts for their words. *Tourism Management*, *67*, 261–272. <https://doi.org/10.1016/j.tourman.2018.02.002>

Varma, A., Jukic, N., Pestek, A., Shultz, C. J., & Nestorov, S. (2016). Airbnb: Exciting innovation or passing fad? *Tourism Management Perspectives*, *20*, 228–237. <https://doi.org/10.1016/j.tmp.2016.09.002>

Wang, D., & Nicolau, J. L. (2017). Price determinants of sharing economy based accommodation rental. *International Journal of Hospitality Management*, *62*, 120–131. <https://doi.org/10.1016/j.ijhm.2016.12.007>

Yang, Y., Kim, S. I., Kim, J., & Koh, Y. (2022). How Airbnb titles influence guests' decision making. *International Journal of Hospitality & Tourism Administration*. <https://doi.org/10.1080/15256480.2022.2114973>

Zervas, G., Proserpio, D., & Byers, J. W. (2017). The rise of the sharing economy. *Journal of Marketing Research*, *54*(5), 687–705. <https://doi.org/10.1509/jmr.15.0204>

Zhang, Z., Chen, R. J., Han, L. D., & Yang, L. (2017). Key factors affecting the price of Airbnb listings. *Sustainability*, *9*(9), 1635. <https://doi.org/10.3390/su9091635>

SITOGRAFIA

Airbnb. (n.d.). *Airbnb – Official website*. <https://www.airbnb.com>

Airbnb Newsroom. (n.d.). *Airbnb Newsroom*. <https://news.airbnb.com>

ISTAT. (n.d.). *Istituto Nazionale di Statistica*. <https://www.istat.it>

Inside Airbnb. (n.d.). *Inside Airbnb*. <http://insideairbnb.com>

LIWC. (n.d.). *LIWC – Linguistic Inquiry and Word Count*. <https://www.liwc.app>

OECD. (n.d.). *Organisation for Economic Co-operation and Development*.

<https://www.oecd.org>

R Project. (n.d.). *The R Project for Statistical Computing*. <https://www.r-project.org>

Yahoo Finance. (n.d.). *Yahoo Finance*. <https://finance.yahoo.com>