



Politecnico di Torino

Corso di Laurea Magistrale in Ingegneria Gestionale

Percorso ICT e Data Analytics per il Management

A.a. 2024/2025

Sessione di Laurea Novembre 2025

Analisi empirica dell'influenza di nazionalità e immagine sulle performance degli host di Airbnb

Relatori:

Elisabetta Raguseo

Francesco Luigi Milone

Candidato:

Riccardo Battaglia

Abstract

La discriminazione nelle piattaforme digitali rappresenta una problematica delicata che sta emergendo negli ultimi anni nel contesto della trasformazione digitale. Anche se la sharing economy viene spesso vista come un modello capace di creare maggiore uguaglianza e opportunità per tutti, la realtà è più complessa. Studi recenti dimostrano che, anche in questi spazi apparentemente neutrali, sopravvivono bias impliciti che possono influenzare i risultati economici di chi vi partecipa.

Partendo da questa premessa, il lavoro si concentra su Airbnb, una delle piattaforme simbolo della condivisione. L'obiettivo è capire, attraverso un'analisi empirica, se la nazionalità e le caratteristiche visive del profilo, in particolare la foto dell'host, possano avere un impatto reale sulle performance economiche. I dati osservati riguardano vari indicatori: il prezzo medio, la tariffa giornaliera media, il tasso di occupazione, i ricavi per notte, il punteggio delle recensioni e il numero di prenotazioni.

Questo studio si inserisce nel più ampio dibattito sui pregiudizi culturali che emergono nelle piattaforme peer-to-peer, dove le interazioni tra utenti sembrano libere, ma non sempre lo sono davvero. Il valore di questa ricerca sta nel suo sguardo sul mercato italiano e in particolare sulla provincia di Torino: un territorio poco analizzato, ma utile per comprendere meglio come certi meccanismi di discriminazione si riproducano anche nel digitale, spesso in modo sottile e inconsapevole.

La prima parte dell'elaborato è dedicato all'introduzione e alla revisione della letteratura con un focus particolare sulle piattaforme digitali, sulla Sharing Economy, su Airbnb e sulla discriminazione. Dopo una fase di definizione delle ipotesi e delle domande di ricerca, segue la sezione di spiegazione delle metodologie utilizzate per l'analisi dei dati. Segue il capitolo di presentazione dei risultati ottenuti. Infine vengono presentate le conclusioni e i possibili sviluppi futuri.

Indice

Abstract	ii
1 Introduzione	1
2 Revisione della letteratura	3
2.1 Racial Discrimination	5
2.1.1 Premesse teoriche	5
2.1.2 Racial Pay Gap	9
2.2 Piattaforme digitali	12
2.2.1 Introduzione alle piattaforme	12
2.2.2 Caratteristiche e classificazione	13
2.2.3 Modello economico e dinamiche di mercato	14
2.2.4 Strategie competitive e gestione delle piattaforme	15
2.2.5 Sharing economy	16
2.3 Airbnb	17
2.3.1 La piattaforma Airbnb	17
2.3.2 Evoluzione e impatto della piattaforma	18
2.3.3 Discriminazione in Airbnb	19
3 Ipotesi di ricerca	21
4 Metodologie	22
4.1 Analisi del database AirDNA	22
4.2 Scraping dei nomi degli host e classificazione	24
4.2.1 Estrazione dei nomi host tramite web scraping	24
4.2.2 Pulizia e normalizzazione dei nomi	27
4.2.3 Analisi del genere e della nazionalità tramite API NamSor .	28
4.2.4 Analisi della foto profilo degli host	30
4.2.5 Considerazioni metodologiche e limiti	31
4.3 Misure di performance	32
4.4 Strumenti per l'analisi	33
4.4.1 Excel	33

4.4.2	Stata	34
4.5	Analisi delle descrittive	35
4.5.1	Distribuzione degli annunci per nazionalità e genere	35
4.5.2	Distribuzione degli host per nazionalità e genere	36
4.5.3	Distribuzione dei ricavi medi	37
4.5.4	Distribuzione del Rating medio	38
4.5.5	Distribuzione del numero medio di prenotazioni	39
4.5.6	Distribuzione della tipologia di alloggio	40
4.5.7	Distribuzione dello status Superhost	41
4.5.8	Distribuzione del numero medio di foto	42
4.5.9	Distribuzione della politica di cancellazione	43
4.5.10	Distribuzione delle caratteristiche visive	44
4.5.11	Distribuzione dell'ADR medio	45
4.5.12	Distribuzione dell'OCC medio	46
4.5.13	Distribuzione del RevPAN medio	47
4.5.14	Andamento delle metriche chiave	47
5	Risultati	50
5.1	Analisi dati panel	51
5.1.1	Regressione multivariata logaritmica	51
5.1.2	Regressione multivariata logaritmica con interazione	55
5.1.3	Interaction plot	62
5.2	Analisi foto profilo degli host	68
5.2.1	Regressione multivariata logaritmica	68
5.2.2	Regressione multivariata logaritmica con interazione	72
6	Conclusioni	76
6.1	Sintesi dei risultati	76
6.2	Implicazioni manageriali	77
6.3	Limiti della ricerca	78
6.4	Considerazioni finali	78
7	Sviluppi futuri	79
	Bibliografia	80
	Elenco delle figure	85
	Elenco delle tavelle	87

Capitolo 1

Introduzione

In questo primo capitolo viene introdotto il tema trattato nella tesi e il contesto in cui si sviluppa. Il contesto viene analizzato da due punti di vista, uno tecnologico, introducendo le innovazioni che lo rendono possibile, e uno più teorico, con una panoramica della letteratura scientifica in cui si inserisce questa ricerca.

Il contesto tecnologico in cui la tesi si inserisce è quello della trasformazione digitale, un fenomeno economico e sociale che permea ogni aspetto vita quotidiana. L'impatto di questo fenomeno non è solo sulla routine dei privati, ma sta rivoluzionando anche il mondo del business, introducendo nuovi modelli di produzione e consumo, basati sulla condivisione. Dal concetto di condivisione nasce il modello della Sharing Economy, ossia un sistema economico che si basa sulla condivisione di beni o risorse attraverso le piattaforme digitali. Questo fenomeno ha dalla sua svariati punti di forza come l'ampia accessibilità, la riduzione dei costi di transazione e l'utilizzo di risorse sottoutilizzate, ma non è esente da criticità come le diseguaglianze algoritmiche e i bias cognitivi che dal mondo reale si riversano in quello online.

Un esempio emblematico di piattaforma digitale che sfrutta tutte le potenzialità della Sharing Economy è Airbnb, su cui questa tesi pone il suo focus.

Questo ambiente digitale, fondato nel 2007, rappresenta una rivoluzione per il settore dell'ospitalità introducendo quello che viene chiamato modello peer to peer basato sulla fiducia, sulla reputazione e sulla visibilità digitale degli host, per compensare la totale assenza di un interazione fisica tradizionale. Diversi studi mostrano che delle differenze legate alla nazionalità e alla presentazione visiva delle foto profilo possono influenzare la fiducia percepita e di conseguenza la propensione degli utenti a prenotare, a lasciare delle recensioni positive e a pagare un prezzo più elevato. Essendo la scelta del guest basata fortemente sulla fiducia nei confronti degli host, spesso i pregiudizi e delle sensazioni inconsce possono fare la

differenza. Ad esempio gli studi di Marchenko e quelli di Edelman [1, 2] mettono in evidenza una minore fiducia dei guest nei confronti degli stranieri. Nel caso studio di Banerjee [3], invece, emerge come gli utenti ripongano maggiormente la loro fiducia negli host ritratti sorridenti nelle loro foto profilo e penalizzano gli host che si presentano con gli occhiali da sole in foto. Considerando che maggiore fiducia spesso si trasforma in valutazioni migliori e ricavi maggiori, l’obiettivo dello studio è quello di analizzare eventuali impatti della nazionalità e di alcuni segnali visivi particolarmente significativi sulle performance degli host di Airbnb nella provincia di Torino. La base per lo sviluppo di questa analisi è il dataset fornito da AirDNA che contiene dati reali degli annunci della provincia di Torino tra il 2017 e il 2024. Il dataset di partenza non era però sufficiente ai fini della ricerca. Infatti la nazionalità è stata dedotta dai nomi propri degli host. Il dataset di partenza non presentava i nomi di questi, per questo è stata necessaria una procedura di web scraping dal sito di Airbnb tramite uno script python sviluppato in autonomia. Inoltre per ogni annuncio sono stati calcolati alcuni indicatori di performance molto utilizzati nel mondo dell’ospitalità, come il prezzo medio per notte (ADR), la percentuale di occupazione (OCC) e i ricavi per notte (RevPAN).

Capitolo 2

Revisione della letteratura

La struttura della tesi prevede un iniziale trattazione della problematica della discriminazione, spiegando la distinzione tra discriminazione "di preferenza" e discriminazione "statistica". [4, 5, 6]. Segue una panoramica di studi che attestano la presenza di discriminazione razziale in molteplici contesti come le assunzioni [7], gli affitti [8] e la scuola [9].

Segue una disamina sulle piattaforme digitali e successivamente sulla Sharing Economy. Si evidenzia come queste innovazioni potrebbero abbattere le barriere e i pregiudizi, e invece spesso hanno l'effetto opposto, ossia quello di rispecchiare le divisioni che si riscontrano nel mondo, favorendo il verificarsi di quella che viene chiamata la discriminazione algoritmica [10]. Nel capitolo successivo segue una analisi della piattaforma su cui si basa lo studio: Airbnb. Dopo una breve spiegazione della storia e del suo funzionamento ci si concentra sul focus della ricerca, ossia cercare evidenze di discriminazione emersa nel contesto di Airbnb. Nel paper "The Impact of Host Race and Gender on Prices" [1] di Marchenko del 2019, l'autore analizza oltre 70.000 host negli Stati Uniti e le evidenze lo portano a concludere che esiste una componente discriminatoria in Airbnb, nello specifico osserva che gli host asiatici e di colore guadagnano dal 3% al 4% in meno rispetto agli host bianchi. Marchenko mette in risalto come la piattaforma non abbatte le disuguaglianze come ci si aspetterebbe e propone anche delle possibili azioni correttive come rendere anonimi i profili.

Park et al. nel paper "Fighting Bias with Bias" [11] del 2023 mostrano come l'effetto discriminatorio diminuisce nel caso in cui l'host abbia ricevuto recensioni positive da guest della stessa etnia del potenziale guest. Questo dimostra ulteriormente come la percezione di fiducia sia legata ad affinità culturali ed etniche. Parker et al. propongono di diminuire la discriminazione in modo più efficace con lo sfruttamento dell'endorsement delle recensioni e delle valutazioni, invece che con l'anonimato dei profili.

Anche i guest non sono esenti da discriminazione, come evidenziato da Edelman et

al. nello studio "Racial Discrimination in the Sharing Economy" [12] del 2017. Gli autori osservano che i guest afro-americani hanno una probabilità del 16% in meno di vedersi accettare la prenotazione rispetto ai guest bianchi. Risulta evidente la presenza di pregiudizio in quanto questo effetto scompare una volta che l'host ha già ospitato guest afro-americani in passato. Anche in questo caso il bias è agevolato dal design della piattaforma, in quanto rende visibili nome e foto del potenziale guest.

Edelman et al. nel paper "The impact of host race and gender on prices on Airbnb" [2] del 2014 analizzano la presenza di discriminazione su Airbnb per le prenotazioni nella città di New York. Anche in questo caso hanno osservato che gli host afro-americani ottenevano performance peggiori, in particolare registrano un prezzo inferiore del 12% rispetto agli host bianchi.

Per quanto riguarda le donne, invece, si osserva una controtendenza: le donne ispirano maggiore fiducia degli uomini. Questo è quanto è stato osservato da Ert et al. nel paper "Gender earning gap on digital platforms: The Airbnb case" [13] del 2024, nel quale si osservano maggiori ricavi per stanza per gli host femminili rispetto a quelli maschili.

Un ulteriore veicolo di fiducia è la foto profilo degli host, su cui hanno concentrato il loro focus Banerjee, Lens e Pal. Nel paper "Put on your sunglasses and smile: The secret of Airbnb hosts' profile photos" [3] del 2022, hanno osservato che i segnali visivi (visual cues) che si presentano nelle foto profilo influenzano in modo significativo la percezione di fiducia e di conseguenza le scelte dei guest. Nello specifico i volti sorridenti aumentano la percezione di fiducia e migliorano le prestazioni, al contrario, la presenza di occhiali da sole trasmette minore fiducia nei guest.

Complessivamente, la letteratura suggerisce che le piattaforme digitali non siano spazi neutrali, ma contesti in cui interagiscono elementi economici, culturali e algoritmici. Le differenze di razza possono tradursi in disparità nelle performance economiche e nella costruzione della reputazione. Gli studi più recenti hanno inoltre richiamato l'attenzione sui segnali visivi dei profili, come il sorriso, che influenzano la fiducia percepita e la propensione alla prenotazione.

Alla luce di tali contributi, la presente ricerca si propone di estendere le evidenze esistenti al contesto italiano, ancora poco esplorato, analizzando la relazione tra nazionalità, rappresentazione visiva e performance economiche degli host. L'obiettivo è duplice: da un lato, verificare se le dinamiche osservate a livello internazionale si manifestino anche nel mercato di Airbnb in Italia; dall'altro, valutare il ruolo che le fotografie del profilo e i segnali visivi hanno sulla percezione di fiducia che trasmettono gli host. In tal modo, la tesi contribuisce alla letteratura sulla discriminazione nelle piattaforme digitali fornendo un'analisi integrata che combina

dimensioni economiche, reputazionali e visive, ponendosi come ponte tra gli studi sulla discriminazione tradizionale e quelli sui bias algoritmici.

Studio	Descrizione sintetica
Banerjee, S., Lens, M., & Pal, A. (2022) <i>Put on your sunglasses and smile: The secret of Airbnb hosts' profile photos. International Journal of Hospitality Management.</i>	<p>Obiettivo: Analizzare come genere, espressione facciale e uso degli occhiali nelle foto profilo degli host influenzino fiducia e intenzione di prenotazione.</p> <p>Variabili indipendenti (VI): Genere (maschio/femmina); espressione facciale (positiva/neutra); occhiali da sole (presente/assente).</p> <p>Variabili dipendenti (VD): Intenzione di fidarsi; intenzione di prenotare.</p> <p>Variabili di controllo: Età e genere del partecipante; percezione di realismo dello scenario sperimentale.</p> <p>Fonti teoriche: Teoria della percezione affettiva; teoria della self-presentation online; studi su trust e facial cues in Airbnb (Ert et al., 2016; Su & Mattila, 2020).</p> <p>Metodologia: Esperimento online $2 \times 2 \times 2$ con 524 partecipanti; immagini dal database Cohn-Kanade; analisi tramite regressioni multiple e modelli di mediazione (Hayes, 2013).</p> <p>Risultati: Le host donne e i volti sorridenti generano maggiore fiducia e intenzione di prenotazione; gli occhiali riducono la fiducia ma l'effetto è attenuato dal sorriso.</p> <p>Conclusioni: Le caratteristiche visive influenzano significativamente le decisioni dei guest; preferibili foto con volto visibile e sorriso.</p>
Ert, E., Fleischer, A., & Kopolovich, D. (2024) <i>Gender earning gap on digital platforms: The Airbnb case. Tourism Economics, 31(3), 453–476.</i>	<p>Obiettivo: Analizzare se esista un gender earnings gap tra host Airbnb e determinarne la direzione e le cause.</p> <p>Variabili indipendenti (VI): Genere dell'host (maschio/femmina).</p> <p>Variabili dipendenti (VD): RevPAR (revenue per available room); ADR (average daily rate); occupancy rate.</p> <p>Variabili di controllo: Caratteristiche dell'appartamento (dimensione, distanza dal mare/centro, rating, foto, Airbnb Plus, pets, ecc.); caratteristiche dell'host (esperienza, Superhost, listings, verifica, ecc.).</p> <p>Fonti teoriche: Gender wage gap (Blau & Kahn, 2017); ICT empowerment (Goldin, 2014); fiducia e reputazione (Ert & Fleischer, 2020).</p> <p>Metodologia: Analisi econometrica su 11.679 listing Airbnb a Tel Aviv (2018–2019) con modello di mercato differenziato (Berry, 1994); stima con GMM e test di robustezza 3SLS; esperimenti aggiuntivi sulla percezione di fiducia.</p> <p>Risultati: Le host donne ottengono RevPAR più alto a parità di condizioni; nessuna differenza significativa in ADR; l'effetto è legato a maggiore fiducia e tasso di occupazione.</p> <p>Conclusioni: Si osserva un'inversione del gender earnings gap, con un vantaggio economico per le donne dovuto alla percezione di fiducia.</p>
Edelman, B. G., & Luca, M. (2014) <i>The impact of host race and gender on prices on Airbnb. Harvard Business School Working Paper.</i>	<p>Obiettivo: Esaminare se razza e genere dell'host influenzino i prezzi degli alloggi, indicando la presenza di discriminazioni implicite.</p> <p>Variabili indipendenti (VI): Razza percepita dell'host (caucasico / afroamericano / asiatico / altro); genere (maschio / femmina).</p> <p>Variabili dipendenti (VD): Prezzo medio dell'annuncio.</p> <p>Variabili di controllo: Caratteristiche dell'appartamento (posizione, tipologia, recensioni, valutazione, dimensione, foto, ecc.); caratteristiche del quartiere (reddito medio, diversità etnica, ecc.).</p> <p>Fonti teoriche: Discrimination economics (Becker, 1971); digital discrimination e trust nelle piattaforme (Edelman & Luca, 2014; Ert et al., 2016).</p> <p>Metodologia: Analisi quantitativa OLS su dataset Airbnb di New York City; classificazione razziale tramite valutatori indipendenti; controlli di robustezza geografica.</p> <p>Risultati: Gli host afroamericani applicano prezzi inferiori (12%) rispetto a host caucasici per proprietà comparabili; differenze non spiegate da qualità o posizione.</p> <p>Conclusioni: Evidenza di discriminazione razziale nei prezzi Airbnb; la trasparenza visiva (foto profilo) amplifica bias impliciti.</p>

Figura 2.1: Tabella sinottica dei principali studi analizzati

2.1 Racial Discrimination

2.1.1 Premesse teoriche

La discriminazione razziale comprende ogni comportamento che, direttamente o indirettamente, comporta distinzione o preferenza basata sulla razza, il colore, l'ascendenza, l'origine o la convinzione religiosa. In particolare, oltre a essere oggettivamente discriminatorio, il comportamento deve avere lo scopo o l'effetto

di distruggere, o quantomeno di compromettere, il riconoscimento, il godimento o l'esercizio dei diritti umani e delle libertà fondamentali in campo politico, economico, sociale e culturale e in ogni altro settore della vita pubblica. La definizione è comprensiva sia dei casi di discriminazione diretta, sia di quelli di discriminazione indiretta. Inoltre, per considerare illecito il comportamento è sufficiente che il fatto abbia l'effetto di produrre la discriminazione, non è quindi necessario che lo scopo sia quello di perseguire il risultato discriminatorio [14].

Questo fenomeno persiste ancora oggi e affonda le sue radici in fattori culturali, pregiudizi e stereotipi trasmessi a livello sociale. Proprio per la persistenza di questo comportamento la letteratura scientifica si è dedicata per decenni alla sua analisi, mettendo in evidenza come la razza e l'origine abbiano effetti sull'accesso al lavoro, all'abitazione, alla scuola e a molte altre attività, con effetti tangibili e misurabili a livello economico.

La discriminazione si può suddividere tra discriminazione "di preferenza" e "statistica".[4, 5, 6] La prima si verifica quando un individuo prova una vera e propria avversione verso un gruppo etnico, tramite comportamenti di razzismo esplicito. In questi casi la persona è disposta anche ad avere un profitto minore pur di evitare il contatto diretto con persone diverse da lui. Quella statistica, invece, deriva da una asimmetria informativa per la quale un individuo, non avendo informazioni sufficienti informazioni sull'altro, gli attribuisce un comportamento statisticamente medio del gruppo, senza verificare che questa media sia veritiera o causata da pregiudizi diffusi [15].

Zschirnt e Ruedin (2016) [7] concentrano il loro studio sul mondo delle selezioni dei candidati per svariati lavori in un arco temporale di 25 anni nei paesi OCSE (Organizzazione per la cooperazione e lo sviluppo economico), hanno osservato che i candidati appartenenti a minoranze etniche devono inviare circa il 50% di candidature in più per ottenere lo stesso numero di convocazioni ai colloqui rispetto ai candidati autoctoni. Anche a parità di qualifiche, nomi che richiamano ad origine arabe o africane hanno probabilità molto più basse di essere contattati rispetto a candidati con nomi europei o nordamericani.

A metà degli anni Novanta, Ayres e Siegelman (1995) [16] hanno osservato che la discriminazione è presente anche nei confronti dei consumatori. Nello studio venivano mandati degli attori di etnia e sesso diverso ad acquistare un'auto in alcuni concessionari americani. A parità di comportamento e di offerta iniziale, gli autori hanno osservato delle differenze significative. Infatti gli uomini bianchi ottenevano il prezzo finale inferiore e poi in ordine crescente di prezzo troviamo le donne bianche, gli uomini afroamericani e infine, con il prezzo finale maggiore, le donne afroamericane.

Neumark et al. (1996) [17], invece osservano fenomeni di discriminazione inversa. Lo studio si concentra sulle assunzioni nei ristoranti, ed osservano differenze tra

candidati maschili e femminili. Le donne vengono preferite per posizioni a contatto con i clienti, mentre gli uomini hanno maggiore probabilità di lavorare nelle cucine. Queste differenze si ribaltano poi sui salari, in quanto gli impiegati di sala ottengono paghe superiori di quelli in cucina. Anche in questo caso queste differenze vengono osservate a parità di condizioni, infatti i candidati erano attori addestrati a comportarsi allo stesso modo. In questo caso emerge uno stereotipo di genere che porta però vantaggi ai candidati femminili.

Anche Pager (2003) [18] studia gli impatti della razza sulla probabilità di essere richiamati da parte di un datore di lavoro, ma aggiunge una variabile che segnala la presenza di precedenti penali. Lo studio si concentra nell'area di Milwaukee, nel Wisconsin (USA). L'autore osserva che a parità di istruzione, esperienza e comportamento la presenza di un precedente penale dimezza la possibilità di essere richiamati. Emerge inoltre che i candidati bianchi con precedente penale hanno una percentuale più alta di essere richiamati rispetto ai neri con il medesimo precedente. Un bianco con precedente penale riscontrava un recall di circa 17%, un nero senza precedente circa il 14% e un nero con precedente penale si attestava solo intorno al 5%, evidenziando una forte penalizzazione discriminatoria, che causa differenze osservabili anche nella distribuzione del tasso di povertà, come osservato negli Stati Uniti nel 2023 (Figura 2.2)

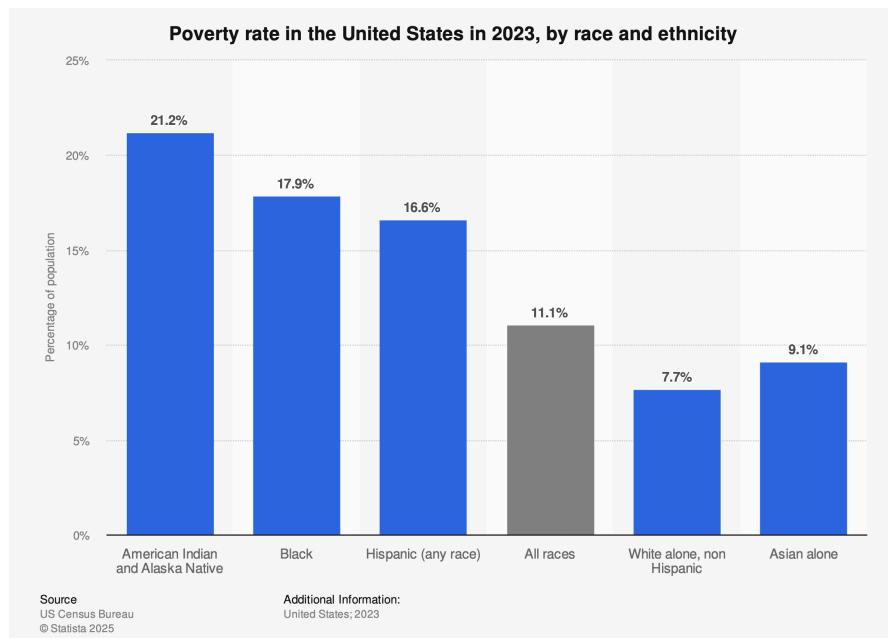


Figura 2.2: Tasso di povertà per etnia negli Stati Uniti (2023)

Fonte: US Census Bureau [19]

Nelle zone di Boston e Chicago invece, Bertrand et al. (2004) [20] hanno osservato che candidati con nomi bianchi hanno ottenuto il 50% di probabilità in più di essere richiamati rispetto a quelli con nomi neri. Risulta inoltre che la discriminazione è presente sia in piccole e medie imprese sia in multinazionali, e che CV migliori non mitigano l'effetto discriminatorio, dimostrando che le aziende ignorano il merito quando il nome è afroamericano.

In America Latina, Galarza e Yamada (2014) [21] studiano la situazione discriminatoria in Perù, nello specifico nella capitale Lima. Di nuovo tramite l'invio di candidature fintizie hanno osservato che i nomi indigeni registravano tassi di callback inferiori rispetto a nomi solitamente associati a persone bianche. Un candidato doveva inviare l'80% di candidature in più per avere lo stesso numero di colloqui di un candidato bianco, inoltre questo effetto si intensificava per posizione lavorative più qualificate.

Anche in Europa sono stati sviluppate analisi di questo genere, nello specifico nei Paesi Bassi Blommaert et al. (2014) [22] hanno inviato CV fintizi con medesime caratteristiche, ma con nomi olandesi e arabi. I CV con nomi arabi nella prima fase, ossia quella di apertura del CV, avevano il 50% di probabilità in meno di essere aperti. Una volta aperti l'effetto discriminatorio non è più significativo. Complessivamente si è osservato che i nomi olandesi avevano il 60% di probabilità in più di ottenere riscontro positivo rispetto ai nomi arabi.

In Svezia Carlsson ed Eriksson (2014) [23] studiano discriminazioni da parte di locatori nei confronti dei potenziali inquilini. Tramite l'invio di candidature fintizie tra il 2010 e il 2011, gli autori hanno osservato che gli stranieri e i disoccupati ricevevano un numero significativamente minore di visite rispetto agli svedesi e agli occupati. Inoltre si è osservato che le disparità aumentavano nelle città più piccole e nei casi in cui i locatori erano privati e non agenzie. In compenso però non sono state osservate discriminazioni dovute a genere ed età.

Anche Flage (2018) [8] si focalizza sullo studio degli affitti nei paesi OCSE, ma nel periodo che va dal 2006 al 2017. L'autore osserva che i candidati con nomi arabi o africani avevano minori probabilità di essere contattati e la disparità aumentava ulteriormente nei casi in cui il locatore era privato. Gli stati presi in esame erano 14 e in tutti si sono osservate prestazioni analoghe, confermando la diffusione capillare del fenomeno.

Nel contesto italiano sono emersi effetti discriminatori simili. Busetta et al (2018) [24], tramite l'invio di CV fintizi, osservano discriminazioni tra candidati italiani, immigrati di prima e di seconda generazione (nati e istruiti in Italia). Gli italiani vengono contattati con probabilità doppia rispetto agli immigrati, e anche gli immigrati nati e istruiti in Italia subiscono la medesima penalizzazione, evidenziando

la presenza di un pregiudizio puro e non dovuto a differenze linguistiche. Inoltre, le donne italiane ricevono meno risposte degli uomini italiani, e per le immigrate donne l'effetto discriminatorio è doppio, in quanto si sommano l'effetto genere con quello etnico.

Alesina et al. (2018) [9] analizzano la possibile presenza di discriminazione nell'assegnazione dei voti da parte degli insegnanti nelle scuole medie italiane. I risultati hanno evidenziato una leggera discriminazione non intenzionale nei voti di matematica. Una volta resi gli insegnanti consapevoli della presenza di questo bias l'effetto discriminatorio scompariva.

In conclusione, risulta evidente dagli studi analizzati che la discriminazione razziale è un fenomeno ancora oggi presente e osservabile anche con i dati. Il contesto in cui questo fenomeno emerge in modo più chiaro è quello delle assunzioni, nel quale a parità di condizioni il semplice nome può influenzare le possibilità di essere contattati, assunti o valutati. Questo dimostra che non si tratta solo di differenze reali nelle competenze, ma di pregiudizi e stereotipi che agiscono spesso in modo automatico. Inoltre, la discriminazione non si manifesta sempre in maniera evidente o intenzionale: in molti casi, opera nelle prime fasi dei processi decisionali, come l'apertura di un curriculum o l'impressione iniziale data da un nome straniero. Proprio per questo, gli effetti diventano ancora più gravi, perché colpiscono prima ancora che venga valutato il merito. Tutti i risultati analizzati mostrano quindi una dinamica sistematica, presente nel lavoro, negli affitti e perfino nella scuola, capace di generare svantaggi che si accumulano nel tempo. Per questo motivo, non basta intervenire sulle regole formali, ma è necessario riconoscere l'esistenza di bias impliciti e lavorare per creare contesti in cui le persone vengano giudicate per ciò che sanno fare, e non per l'origine che il loro nome lascia intuire.

2.1.2 Racial Pay Gap

L'esistenza del racial pay gap (Figura 2.3), ossia di una differenza salariale tra persone di origini ed etnie diverse, si allinea con le evidenze discriminatorie messe in evidenza nella sezione precedente. La discriminazione inizia nella fase di selezione dei candidati, come visto in precedenza, e prosegue anche nella progressione della carriera e di conseguenza anche nella assegnazione dei salari.

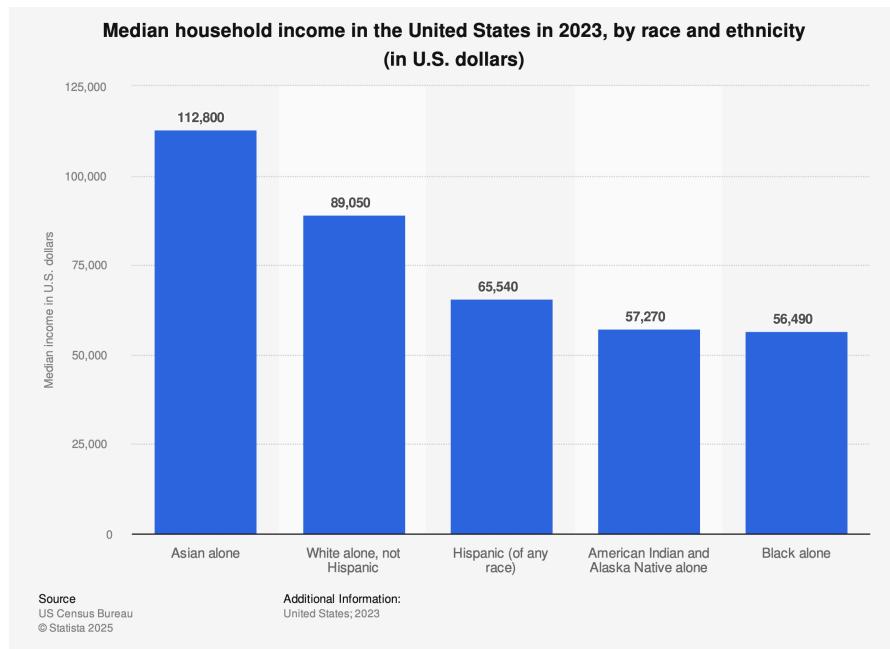


Figura 2.3: Reddito mediano delle famiglie negli Stati Uniti (2023)

Fonte: *US Census Bureau [19]*

Nel paper di Budig, Lim e Hodges (2021) [25] emerge come l'accesso a un titolo accademico non produce lo stesso ritorno economico per tutti i gruppi etnici. Analizzando dati per oltre tre decenni, gli autori rilevano che gli afroamericani, in particolare le donne di colore, percepiscono salari inferiori rispetto ai lavoratori bianchi con pari livello di istruzione e medesimo campo di studio. Ciò indica che non sarebbe sufficiente colmare il divario educativo per ridurre le disuguaglianze, ma che entra in gioco una minore valorizzazione dei titoli conseguiti dalle minoranze nei contesti professionali. Tale meccanismo produce una doppia penalizzazione, sia nella fase di ingresso nel mercato del lavoro sia nelle opportunità di avanzamento. Heywood e Parent (2012) [26] evidenziano che nei contesti in cui il salario dipende da obiettivi o bonus individuali, il gap tra lavoratori bianchi e neri tende ad aumentare, soprattutto nelle fasce più elevate di salario. Al contrario, nei settori caratterizzati da salari fissi, come nel pubblico impiego, il divario si attenua o scompare, suggerendo che i sistemi retributivi variabili possono rafforzare le disuguaglianze esistenti.

Gli studi di Blau e Kahn (1992) [27] contribuiscono a inquadrare il racial pay gap come fenomeno radicato che, nonostante alcuni progressi nella riduzione del gender pay gap, rimane stabile. Come si può osservare in Figura 2.4 nonostante i salari di tutti i gruppi aumentino, a causa dell'inflazione, nessun gruppo che subisce discriminazione riesce a recuperare il gap con i salari dei bianchi.

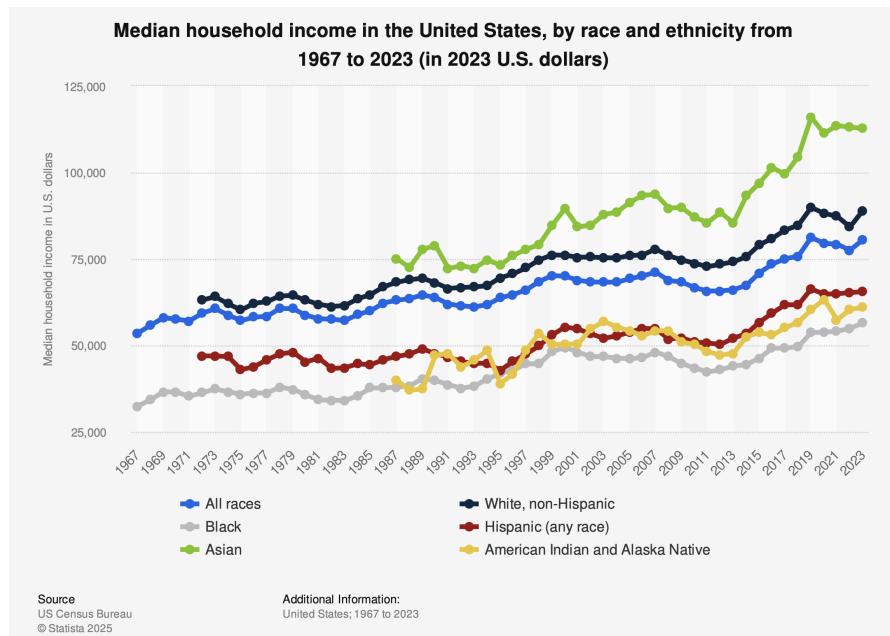


Figura 2.4: Evoluzione storica del reddito negli Stati Uniti (1967–2023)
Fonte: *US Census Bureau* [19]

Tale persistenza è riconducibile principalmente alla differenza di istruzione e al declino della sindacalizzazione. Le differenze retributive non possono dunque essere spiegate unicamente da caratteristiche individuali, ma devono essere interpretate alla luce di vincoli sistematici che incidono sulle possibilità occupazionali.

Il ruolo dell'istruzione si intreccia con la segmentazione del mercato del lavoro. Le minoranze sono spesso sotto-rappresentate nei settori STEM e nelle posizioni manageriali, dove i ritorni salariali sono più elevati, e tendono a essere concentrate in ruoli a basso margine di crescita. Anche quando accedono agli stessi livelli di istruzione, incontrano percorsi professionali caratterizzati da minore mobilità, ridotta stabilità contrattuale e prospettive di carriera più limitate.

Il racial pay gap si configura dunque come il risultato di un insieme di fenomeni come segregazione occupazionale, sottovalutazione delle competenze e assunzioni soggette a bias. Tali meccanismi producono effetti che si amplificano nel tempo, generando svantaggi persistenti non solo sui redditi, ma anche in termini di accumulazione di ricchezza, sicurezza economica e trasmissione delle opportunità alle generazioni successive.

Intervenire sul divario retributivo razziale significa intervenire non solo sull'accesso all'istruzione, ma anche sui criteri di valutazione e remunerazione nel mercato del lavoro. Ciò richiede sistemi trasparenti di assegnazione dei bonus, programmi di tutoraggio e inclusione professionale, introduzione di controlli antidiscriminatori

e promozione di ambienti equi.

In definitiva, il racial pay gap non può essere considerato come il risultato di differenze individuali, ma come una questione di disparità culturale radicata nel tessuto sociale.

2.2 Piattaforme digitali

Le piattaforme digitali, come definito dall'Organisation for Economic Co-operation and Development (OECD), sono dei "servizi online che facilitano le interazioni tra due o più insiemi di utenti distinti, siano essi imprese o individui, che interagiscono tramite il servizio via Internet". [28]

Una trasformazione dei modelli economici e dei vari tipi di interazione sociale è stata determinata grazie alle nuove tecnologie digitali. Le piattaforme digitali rappresentano uno dei fenomeni più interessanti da questo punto di vista, definite come vere e proprie infrastrutture tecnologiche che uniscono tra loro diversi gruppi di utenti e favoriscono lo scambio di beni o di servizi. Con esse viene completamente modificato il confine o il legame tra il produttore e il consumatore e viene anche creato un nuovo tipo di mercato indicato con il termine multi side dove la creazione del valore è basata sull'interazione di più attori contemporaneamente.

Con la loro introduzione l'efficacia dei processi di scambio e la produzione di nuove tipologie di partecipazione economica hanno iniziato a crescere profondamente. Purtroppo però la natura di questi sistemi digitali comporta delle questioni in termini di equità, trasparenza e inclusione e ovvero che i meccanismi algoritmici che vanno a regolarne il funzionamento producono delle disuguaglianze.

2.2.1 Introduzione alle piattaforme

Quando si parla di cambiamento portato dalle piattaforme digitali si riferisce al modo in cui le persone si incontrano nel mercato. Non si tratta solo di semplici siti o di app, sono delle vere e proprie infrastrutture tecnologiche che permettono a diversi gruppi di utenti di scambiarsi valore. Da un lato ci sono i produttori e dall'altro i consumatori. La cosa interessante è che le piattaforme non possiedono direttamente ciò che viene scambiato ma creano uno spazio digitale dove tutto diventa più facile, veloce e meno costoso. Con esse si ha un allontanamento e un conseguente miglioramento rispetto al tradizionale poiché riducono i costi delle transazioni, aumentano l'efficienza e riescono a far incontrare più persone, ampliando così il mercato.

Il funzionamento di questo ecosistema digitale è connesso alla teoria dei mercati multi-side ovvero quei mercati definiti dalla presenza di più lati interdipendenti il cui valore deriva dalle interazioni degli utenti. Il concetto di mercato bilaterale viene introdotto per la prima volta dallo studio condotto da Rochet e Tirole, i quali volevano parlare di un tipo di mercato dove la partecipazione di ciascun lato dipendeva dalla presenza e dal comportamento dell'altra parte. Questa ricerca ha valorizzato le basi teoriche che aiutano a comprendere al meglio il funzionamento delle piattaforme.

Ci sono da ricordare anche altri articoli che legati al precedente sottolineano la grande importanza delle caratteristiche legate al processo di innovazione nell'economia digitale grazie a nuovi modelli di business.[29, 30, 31]

2.2.2 Caratteristiche e classificazione

Adesso viene approfondito il punto di vista più strutturale delle piattaforme. Questi ecosistemi digitali si fondano su dei meccanismi di interazione e di scambio che arrivano a creare delle esternalità di rete e quindi degli effetti positivi che sono derivati dall'aumento del numero di utenti all'interno della piattaforma. Le esternalità di rete citate si manifestano in due modi diversi: o quelle che vengono definite same side network effects ovvero quelle che si manifestano all'interno dello stesso lato o quelle che si esprimono con il termine di cross-side network effects ovvero tra lati opposti. In particolare, same-side inteso come la partecipazione di nuovi utenti che va ad incrementare il valore percepito dagli altri appartenenti allo stesso gruppo; cross-side invece come crescita di un gruppo di utenti che va ad aumentare l'attrattività della piattaforma per l'altro lato. Per fare un esempio un numero maggiore di host su Airbnb tende a rendere la piattaforma ancora più interessante per i guests, ciò va ad incrementarne il valore complessivo.

Si possono utilizzare diversi criteri per andare a classificare le piattaforme, tra cui la natura delle transazioni, il tipo di valore scambiato e le relazioni economiche tra i diversi lati. Inoltre, alcune piattaforme (es. Amazon, Airbnb) operano in contesti di scambi che possono essere osservati o monetizzati, mentre altre piattaforme come i social network si basano invece su transazioni non direttamente economiche ovvero lo scambio di informazioni o sull'attenzione degli utenti. Inoltre, si possono distinguere le piattaforme da quelle che sono orientate ai servizi a quelle basate su delle logiche di sussidio economico che serve ad aiutare a comprendere al meglio le diverse modalità con cui costruire del valore e andare a fidelizzare gli utenti.

Tutte queste caratteristiche vanno ad incidere sulla capacità della piattaforma di raggiungere ciò che viene definito come massa critica e cioè la soglia minima di partecipazione necessaria a garantire il funzionamento più efficiente del mercato e la sostenibilità economica del modello. [32]

TIPOLOGIA	DESCRIZIONE	ESEMPI
Piattaforme di transazione	Tecnologie, prodotti o servizi che operano come collegamenti (o intermediari) al fine di facilitare gli scambi o le transazioni tra utilizzatori, clienti o fornitori.	Uber, AirBnB, Ebay.
Piattaforme di innovazione	Tecnologie, prodotti o servizi che rendono possibile a un gruppo di imprese – potenzialmente identificabili come partner di un ecosistema innovativo ^a – di sviluppare tecnologie, prodotti o servizi complementari tra loro.	Microsoft, Linux, Oracle.
Piattaforme di integrazione	Tecnologie, prodotti o servizi che integrano una funzione di transazione e una di innovazione.	Apple, Alphabet, Amazon, Facebook.
Piattaforme di investimento	Imprese con una strategia di portafoglio basata sull'investimento in piattaforme digitali (operando come <i>holding companies</i> , investitori in capitale di rischio o in entrambe le funzioni).	Softbank, Naspers, Priceline.

Figura 2.5: Tipologie di piattaforme digitali

Fonte: Evans e Gawer 2016 [33]

Come si vede dall’immagine 2.5 descritta dallo studio di Evans e Gawer le piattaforme vengono ulteriormente suddivise in 4 tipologie diverse quelle di transazione, di innovazione, di integrazione e infine di investimento, ognuna caratterizzata da delle funzioni specifiche. Le piattaforme di transazione facilitano l’incontro tra la domanda e l’offerta e portano scambi o interazioni tra gli utenti, i clienti e i fornitori (es. Uber, Airbnb, eBay). Le piattaforme di innovazione forniscono invece un’infrastruttura tecnologica condivisa che consente a diversi attori di collaborare per lo sviluppo di soluzioni complementari (es. Microsoft, Linux, Oracle). Le piattaforme di integrazione combinano entrambe le funzioni precedenti, integrando transazioni e innovazione in un unico sistema (es. Apple, Alphabet, Amazon, Facebook). Per finire le piattaforme di investimento sono imprese che vanno ad adottare una strategia di portafoglio basata sull’investimento in piattaforme digitali, operando come società che detengono partecipazioni in altre aziende o come investitori che finanzianno nuove imprese ad alto potenziale (es. Softbank, Naspers, Priceline).[33]

2.2.3 Modello economico e dinamiche di mercato

Andando adesso ad analizzare le piattaforme digitali più dal punto di vista economico possiamo notare come esse si caratterizzano per una struttura di mercato più bilaterale dove il valore complessivo dipende dalla partecipazione in contemporanea di due o più gruppi di utenti. Quindi così le decisioni di prezzo e di accesso devono essere determinate in funzione dell’interazione tra i due lati del mercato e della sensibilità della domanda di ognuno di essi.

Un elemento da attenzionare è quello di individuare il lato da sussidiare e quello da monetizzare. La piattaforma può decidere di incentivare la partecipazione di un gruppo di utenti (per esempio offrendo dei servizi gratuiti o delle condizioni vantaggiose) per andare ad aumentare il valore percepito dall’altro lato e a stimolare la crescita. In questo modo la massa critica viene raggiunta progressivamente e ciò consente alla piattaforma di diventare autosostenibile.

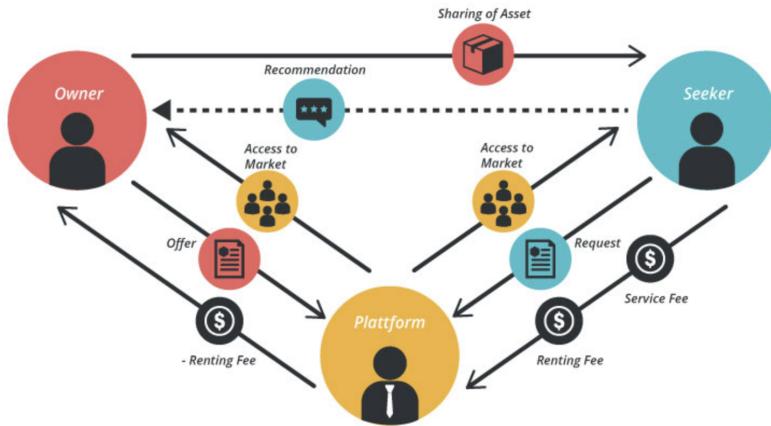


Figura 2.6: Visualizzazione delle relazioni in una piattaforma a due parti

Fonte: *A European Agenda for the Collaborative Economy 2016* [34]

Uno dei problemi con cui le piattaforme devono confrontarsi nelle fasi iniziali del loro sviluppo è quello del “chicken-and-egg problem”, esso è riferito al riuscire ad attrarre allo stesso tempo produttori e consumatori senza possedere una base di utenti già consolidata. Tra le strategie più efficaci per andare a superare questo problema troviamo quella del tipo “divide and conquer” caratterizzata dal fatto che gli sforzi iniziali vengono concentrati su un lato specifico del mercato per generare un primo nucleo di utenti e andare ad innescare le dinamiche di rete che andranno ad attirare l’altro lato, che favoriranno l’espansione della piattaforma. [35]

2.2.4 Strategie competitive e gestione delle piattaforme

Quando si vuole gestire una piattaforma digitale in modo corretto bisogna fare attenzione alla costruzione di meccanismi efficaci di fiducia e di reputazione soprattutto per quanto riguarda i contesti peer-to-peer che sono caratterizzati da forte asimmetria informativa tra gli utenti. Elementi come le recensioni, le valutazioni, i badge o gli storici delle transazioni hanno un ruolo molto importante nel ridurre l’incertezza per lo scambio e nell’aiutare a favorire dei comportamenti di maggiore collaborazione tra gli utenti della piattaforma che non si conoscono direttamente. Quindi la credibilità e l’affidabilità del sistema di reputazione diventa una componente importante nell’idea di crescita per una piattaforma digitale. [36] Le piattaforme digitali per quanto riguarda il punto di vista competitivo operano in mercati dove si manifestano spesso forti economie di scala o anche forti effetti di rete che potrebbero condurre a dinamiche di “winner takes all”. Quando siamo di fronte a una situazione del genere pochi operatori dominano il mercato grazie alla loro capacità di attrarre la maggior parte degli utenti e di consolidare quindi il loro

vantaggio competitivo. Però, per raggiungere una coesistenza di più piattaforme bisogna porsi come obiettivo quello di riuscire a differenziarsi in termini di segmentazione, qualità del servizio, caratteristiche tecniche o target di riferimento.[37] Per quanto riguarda le strategie di prezzo adottate dalle piattaforme digitali esse differiscono di molto da quelle dei mercati tradizionali. L'impatto di ciascun lato del mercato è diverso e quindi la piattaforma può adottare politiche di prezzo diverse per incoraggiare più utenti a partecipare. Ad esempio per quanto riguarda il mondo dei videogiochi, le console vengono spesso vendute a un prezzo economico per favorirne la diffusione e incentivare quindi la produzione di nuovi elementi da parte degli sviluppatori.

2.2.5 Sharing economy

La Sharing Economy, o economia collaborativa, è un fenomeno che si basa sulla condivisione e lo scambio di risorse tra persone tramite l'utilizzo delle piattaforme digitali, di cui abbiamo trattato in questo capitolo.

La nascita di questo nuovo modello di mercato ha permesso non solo di ridurre i costi di transazione e quindi di favorire l'interazione tra le parti, ma anche di utilizzare in modo più efficiente beni sottoutilizzati [38].

Data l'eterogeneità del fenomeno non esiste una definizione univoca del concetto di sharing economy, ma si possono individuare alcuni elementi principali:

- **Condivisione di risorse sottoutilizzate:** materiali o immateriali, più persone accedono alla risorsa più questa verrà sfruttata appieno.[39]
- **Centralità dell'accesso rispetto alla proprietà:** il valore economi l'acquisto viene sostituito dall'utilizzo temporaneo del bene.[40]
- **Uso di piattaforme digitali:** permettono come detto in precedenza di favorire gli scambi.
- **Sistemi di reputazione e fiducia:** possibilità di ridurre l'asimmetria informativa tramite l'introduzione di strumenti come recensioni e rating.[41]
- **Effetti di rete:** il valore generato dalla piattaforma aumenta proporzionalmente con l'aumentare degli utenti che la utilizzano.[39]
- **Flessibilità e scalabilità:** l'assenza di asset consente una elevata capacità di adattamento ai vari contesti del mercato.[42]

Una possibile definizione viene fornita dal *Oxford English Dictionary* [43] ossia: un sistema economico in cui beni o servizi sono condivisi tra individui attraverso Internet, spesso con remunerazione.

Dati i vantaggi e la flessibilità di questo modello economico, viene utilizzato in una grande varietà di settori. I servizi di car sharing e bike sharing (es. BlaBlaCar, Uber, Zipcar, Lime) stanno trasformando la mobilità urbana la mobilità urbana [38]. Il settore del turismo e alberghiero viene rivoluzionato da piattaforme come Airbnb, HomeExchange e Couchsurfing [40, 39]. Applicazioni come TaskRabbit, Care.com o Upwork forniscono nuove opportunità di lavoro e favoriscono la diffusione della gig economy, ossia un modello che si prestazioni di lavoro temporanee [44]. Vinted, eBay o Swap promuovono lo scambio e la vendita di beni usati diminuendo gli sprechi e i rifiuti [42]. Modelli di peer-to-peer lending, crowdfunding e criptovalute stanno trasformando il mercato finanziario e in questo contesto stanno emergendo le fintech. Anche la formazione e la conoscenza passano attraverso dei nuovi canali e diventano accessibili ad una platea sempre maggiore di utenti, ad esempio tramite Udemy e Coursera.

In sintesi, la sharing economy è un fenomeno eterogeneo che trova terreno fertile nell'innovazione delle piattaforme digitali per diffondersi e permeare molteplici aspetti economici e sociali. Nei capitoli successivi focalizzeremo l'attenzione sul settore turistico e nello specifico sulla piattaforma Airbnb.

2.3 Airbnb

Airbnb [45] è la piattaforma scelta per lo svolgimento delle analisi illustrate nella tesi. In seguito vengono indagate le modalità di funzionamento del sito e le possibili dinamiche che possono causare disparità culturali o di genere.

2.3.1 La piattaforma Airbnb

Airbnb viene fondata a San Francisco nell'ottobre del 2007 da Brian Chesky, Joe Gebbia e Nathan Blecharczyk [46]. Nasce come progetto sperimentale di “ospitalità condivisa” e assume rapidamente una dimensione globale, diventando un attore centrale nei settori del turismo e dell’immobiliare. Come sottolinea Hall et al. (2022) [47], il successo della piattaforma si deve alla sua capacità di rispondere alla crescente domanda di esperienze di viaggio personalizzate, posizionandosi come un’alternativa più flessibile all’ospitalità tradizionale. Grazie all’intermediazione online, Airbnb ha permesso una connessione diretta tra host e ospiti, facilitando lo scambio di beni e servizi al di fuori dei circuiti istituzionali.

Alla base del suo funzionamento vi è il modello delle piattaforme a due lati, in cui domanda e offerta vengono messe in relazione tramite un’infrastruttura digitale. Gli host possono offrire spazi abitativi come stanze condivise, appartamenti o case intere e i guest possono ricercare questi spazi utilizzando strumenti di

ricerca e prenotarli anche per lunghi periodi. Come mostrano Zervas, Proserpio e Byers (2017) [48], questo modello ha introdotto un nuovo tipo di concorrenza per il settore alberghiero, contribuendo a modificare profondamente il mercato turistico.

Un aspetto chiave è rappresentato dai sistemi di recensioni tra host e guest che, insieme a punteggi numerici, badge (come il riconoscimento di Superhost) e feedback testuali, fungono da segnali reputazionali fondamentali. Ikkala e Lampinen (2015) [49] osservano che questi strumenti non sono semplici supporti tecnici, ma veri e propri dispositivi sociali, che influenzano la percezione di affidabilità e possono determinare l'esito delle transazioni. La fiducia, infatti, è una caratteristica centrale nel funzionamento di Airbnb, ma può diventare anche un veicolo per la riproduzione di pregiudizi impliciti.

Oltre ai sistemi reputazionali, la piattaforma adotta algoritmi di pricing dinamico e criteri di visibilità personalizzata. Gli host possono scegliere autonomamente il prezzo delle proprie strutture, ma ricevono suggerimenti automatici basati su variabili come la stagionalità, la domanda nella zona e le caratteristiche dell'alloggio. Inoltre, la posizione di un annuncio nei risultati di ricerca dipende da indicatori come la frequenza di risposta, il numero di recensioni, la qualità del profilo e l'utilizzo di funzionalità come la prenotazione immediata. Secondo Li, Moreno e Zhang (2016) [50], questi strumenti premiano chi adotta strategie più sofisticate di gestione, creando asimmetrie tra host occasionali e host professionali.

2.3.2 Evoluzione e impatto della piattaforma

Inizialmente la piattaforma contava soprattutto utenti che mettevano a disposizione stanze libere nelle proprie abitazioni in modo occasionale, ma nel tempo si è verificata un'importante professionalizzazione dell'offerta e oggi una quota rilevante degli host gestisce più annunci in modo sistematico, spesso con finalità d'investimento. Horn e Merante (2017) [51] hanno evidenziato come questa evoluzione abbia avuto impatti sul mercato immobiliare, contribuendo alla riduzione dell'offerta abitativa tradizionale e all'aumento dei canoni di locazione in alcune città, ad esempio Boston.

Wachsmuth e Weisler (2018) [52] evidenziano la presenza di un effetto di gentrificazione guidato dalla sharing economy.

La gentrificazione è un concetto sociologico che indica un processo di trasformazione, fisica e socioculturale, dei quartieri di una grande città che porta a un graduale cambio della cittadinanza che può permettersi di viverci (tipicamente da proletaria a borghese)[53].

Infatti come osservato da Quattrone et al. (2016) [54] la distribuzione degli annunci tende a concentrarsi in aree centrali e ad alta attrattività turistica rendendo il centro delle città sempre più dedito al turismo invece che alla residenza degli abitanti, costretti a spostarsi verso la periferia.

Airbnb è quindi una piattaforma complessa in cui intervengono dinamiche economiche e sociali. Non solo connette host e guest, ma modifica interi quartieri e il tessuto sociale che lo abita. Crea nuove opportunità, ma espone anche a rischi di diseguaglianze culturali e sociali.

2.3.3 Discriminazione in Airbnb

Airbnb è nata con l'obiettivo di rendere più accessibile e flessibile il mercato degli affitti a breve termine, promuovendo un modello basato sulla fiducia tra utenti. Per facilitare questa fiducia, la piattaforma spinge gli utenti, sia host che guest, a completare profili personali con nome, foto e informazioni aggiuntive. Questi elementi sono pensati per aumentare la trasparenza e favorire un clima di autenticità. Tuttavia, questa esposizione dell'identità può produrre effetti collaterali significativi, soprattutto sul piano delle diseguaglianze sociali.

Diversi studi hanno evidenziato come la visibilità dell'identità dell'host possa attivare bias inconsci da parte dei potenziali ospiti. Edelman at al. (2014) [2] sono stati tra i primi a documentare questo fenomeno, dimostrando che gli host afroamericani a New York tendono a guadagnare circa il 12% in meno rispetto ai colleghi bianchi anche a parità di qualità dell'alloggio, posizione e valutazioni. Gli autori fanno notare che questa differenza di prezzo non è dovuta a differenze nei servizi offerti, ma riflette una forma di discriminazione da parte dei guest, che attribuiscono un valore inferiore agli annunci gestiti da persone di colore.

Una ricerca più recente condotta da Marchenko (2019) [1] su un campione di host statunitensi ha mostrato che sia gli host afroamericani sia quelli asiatici, ricevono in media meno prenotazioni e applicano prezzi più bassi. L'autrice ipotizza che questa differenza di prezzo sia dovuta ad una preferenza di host bianchi da parte dei guest, gli host non bianchi notando una diminuzione della domanda sono costretti ad abbassare i prezzi.

La discriminazione, però, non si limita solo sul dell'offerta. Edelman, Luca e Svirsky (2017) [12] hanno condotto un esperimento sul campo per misurare l'effetto dell'etnia percepita del guest sulla probabilità di ricevere una conferma da parte dell'host. L'esperimento è stato svolto inviando richieste di prenotazione identiche, ma firmate con nomi distintivi afroamericani o bianchi. Gli autori hanno scoperto

che i profili con nomi percepiti come afroamericani venivano accettati nel 42% dei casi, contro il 50% per quelli bianchi. Questo divario di 8 punti percentuali, non essendo dovuto a caratteristiche degli account, in quanto volutamente identici, deve necessariamente essere dovuto al nome del guest e dalla etnia percepita.

Inoltre è stato osservato che la discriminazione era più marcata da parte di host che non avevano mai ospitato guest afroamericani, mentre si riduceva in modo significativo nel caso di host che li avevano già ospitati.

Il tentativo della piattaforma di rendere i profili degli utenti più comunicativi, dando la possibilità di aggiungere foto e nome, diventa un arma a doppio taglio e rende possibili dei comportamenti discriminatori. Alcuni autori hanno però hanno riscontrato alcuni casi che mitigano la discriminazione. Park et al. (2023) [11] hanno evidenziato fenomeni di same-race endorsements, ossia un comportamento osservato analizzando i dati della città di New York che porta un guest a scegliere con maggiore probabilità host che abbiano ricevuto recensioni positive da parte di guest della sua stessa etnia. Ad esempio, gli host neri ottengono più prenotazioni da parte di guest bianchi se tra le recensioni sono presenti commenti positivi lasciati da altri utenti bianchi.

Questi risultati suggeriscono che la discriminazione non è solo dovuta ad atteggiamenti individuali, ma è anche influenzata dal modo in cui le piattaforme vengono progettate. Il modo in cui vengono mostrati i nomi e le foto degli host, come anche l'ordine delle recensioni che vengono mostrate rendono il design delle piattaforme sempre meno neutrali. Questo fenomeno può portare ad agevolare disuguaglianze e rende quindi fondamentale porsi delle domande su come le interfacce sono progettate e su come potrebbero essere modificate per ridurre al minimo il rischio di incappare in comportamenti discriminatori.

Capitolo 3

Ipotesi di ricerca

La ricerca è finalizzata a verificare la presenza di eventuali differenze sistematiche nei risultati ottenuti dagli host, cercando di porre l'attenzione agli effetti dei segnali visivi sulla fiducia e sui comportamenti relativi alla prenotazione, attraverso un approccio empirico. In particolare vengono poste le seguenti domande di ricerca:

- **RQ1.** La nazionalità e i segnali visivi delle foto profilo (visual cues) dell'host generano un effetto significativo sulla domanda di prenotazioni?
- **RQ2.** La nazionalità e i segnali visivi delle foto profilo (visual cues) dell'host generano un effetto significativo sul rating dell'annuncio?
- **RQ3.** La nazionalità e i segnali visivi delle foto profilo (visual cues) dell'host generano un effetto significativo sui ricavi per notte (RevPAN) degli host?

A cui vengono associate le seguenti ipotesi di ricerca:

- **HP1.** Gli host percepiti come più affidabili (italiani, segnali visivi positivi) generano una maggiore domanda di prenotazioni.
- **HP2.** Gli host percepiti come più affidabili (italiani, segnali visivi positivi) ottengono rating più alti.
- **HP2.** Gli host percepiti come più affidabili (italiani, segnali visivi positivi) ottengono ricavi maggiori (RevPAN).

(*"Put on your sunglasses and smile: The secret of Airbnb hosts' profile photos"* [3] Banerjee et al. 2022, *"The impact of host race and gender on prices on Airbnb"* [2] Edelman et al. 2014)

Lo studio si allinea con la letteratura scientifica esistente e la ampia, andando ad analizzare il contesto italiano che risulta ad oggi poco attenzionato.

Capitolo 4

Metodologie

Lo scopo del seguente capitolo è quello di presentare il percorso fatto per ottenere i dati utilizzati per lo sviluppo delle regressioni. Il punto di partenza del lavoro è il database panel fornito AirDNA, che è stato arricchito dei nomi degli host ottenuti tramite uno scraping del sito Airbnb. Lo scraping è stato eseguito tramite lo sviluppo di un programma Python. Ottenuti i nomi degli host è stato possibile dedurre la nazionalità e il genere tramite il sistema di onomastica AI di NamSor. La fase di raccolta è stata poi seguita da una fase di pulizia del dataset. Tutte le attività eseguite sono spiegate nelle sezioni successive.

4.1 Analisi del database AirDNA

Lo studio si basa su un dataset fornito da AirDNA , "il principale fornitore di dati e analisi per il settore degli affitti a breve termine. Esiste per fornire a host, investitori e aziende gli strumenti più accurati per avere successo nel settore delle case vacanze. AirDNA tiene traccia delle prestazioni giornaliere di oltre 10 milioni di proprietà su Airbnb e Vrbo in 120.000 mercati globali. Raccoglie anche dati da una libreria in crescita di oltre 1 milione di proprietà partner." [55]

Il dataset raccoglie gli annunci Airbnb nella provincia di Torino per il periodo temporale compreso tra il 2017 e il 2024.

Si tratta di un dataset di tipo panel, in quanto i dati sono organizzati in serie temporali e comprendono un'ampia gamma di variabili relative alle performance economiche e alle caratteristiche degli alloggi.

Variabili del dataset

- Caratteristiche dell'annuncio

- **Property ID:** identificativo univoco della proprietà.
- **Reporting Month:** mese e anno di riferimento.

- **Variabili di performance**

- **Revenue USD:** ricavi totali generati dall'annuncio (in dollari statunitensi).
- **Number of Reservations:** numero complessivo di prenotazioni ricevute nel periodo.
- **Reservation Days:** numero di giorni effettivamente prenotati.
- **Available Days:** numero di giorni in cui l'alloggio è rimasto disponibile.

- **Caratteristiche della proprietà**

- **Host ID:** identificativo univoco dell'host.
- **Property Type:** tipologia della proprietà (appartamento, casa indipendente, stanza, ecc.).
- **Listing Type:** tipo di sistemazione offerta (intero appartamento, stanza privata, stanza condivisa).
- **Property Manager:** indicazione di gestione da parte di agenzia.
- **Host Type:** classificazione in base al numero di annunci gestiti.
- **Comune:** comune di ubicazione.
- **Latitude e Longitude:** coordinate geografiche.
- **Bedrooms, Bathrooms:** numero di camere da letto e bagni.
- **Max Guests:** numero massimo di ospiti.
- **Cancellation Policy:** politica di cancellazione (flessibile, moderata, stretta, molto stretta).
- **Minimum Stay:** durata minima del soggiorno.
- **Created Date:** data di creazione dell'annuncio.
- **Response Rate e Response Time:** percentuale e tempo medio di risposta alle richieste.
- **Number of Photos:** numero di fotografie caricate.
- **Instantbook:** prenotabile immediatamente senza approvazione (1 = vero, 0 = falso).
- **Airbnb Superhost:** status di Superhost (1 = vero, 0 = falso).
- **Listing URL e Listing Main Image URL:** link all'annuncio e all'immagine principale.

- Overall Rating: valutazione complessiva su scala 0-100.
- Voti specifici di comunicazione, accuratezza, pulizia, check-in, posizione e valore (Airbnb Communication Rating, Accuracy Rating, Cleanliness Rating, Checkin Rating, Location Rating, Value Rating).

- Dotazioni e servizi

- Has Kitchen: presenza cucina (1 = vero, 0 = falso).
- Has Parking: presenza parcheggio (1 = vero, 0 = falso).
- Amenities: elenco o codifica dei servizi e delle dotazioni (es. piscina, aria condizionata, Wi-Fi).

- Variabili territoriali

- SLL_2011_T: codifica territoriale secondo sistema statistico del 2011.
- DEN_SL2011: densità abitativa riferita al 2011.
- PRO_COM_T: codice del comune o altra divisione amministrativa territoriale.

4.2 Scraping dei nomi degli host e classificazione

Il dataset panel ottenuto da AirDNA conteneva informazioni interessanti dal punto di vista economico per gli annunci della provincia di Torino, ma identificava l'host di pertinenza esclusivamente con un host_id univoco. Per procedere con l'analisi socio-demografica è stato necessario estrapolare il nome di ogni singolo host partendo dagli host_id forniti da AirDNA. L'ottenimento di queste informazioni è stato possibile grazie ad una fase di scraping automatizzato del sito Airbnb, seguita da analisi automatica del nome attraverso strumenti di intelligenza artificiale (API NamSor), con l'obiettivo di indagare nazionalità e gender dei nomi ottenuti.

4.2.1 Estrazione dei nomi host tramite web scraping

L'attività di scraping (Figura 4.1) è stata svolta su ambiente macOS, utilizzando ChromeDriver, componente che consente di controllare in modo automatizzato Google Chrome.

Per evitare problemi legati agli accessi automatizzati e garantire una sessione stabile, lo script è stato configurato per aprire Chrome con il profilo utente locale, disattivando la modalità headless. In questo modo è stato possibile simulare una normale navigazione manuale, accedendo ai contenuti senza incorrere nei blocchi o nelle limitazioni imposte dalla piattaforma.

In una fase iniziale è stata utilizzata l'estensione Airbnb Data Scraper (disponibile su Chrome Web Store), che ha permesso di raccogliere automaticamente gli ID degli host e alcune informazioni di base sugli annunci. Successivamente, con Selenium, è stato implementato un ciclo automatizzato che, partendo da ogni host_id, apriva la pagina del profilo corrispondente, individuava nel codice HTML il nome dell'host ed estraeva il dato. Il nome veniva poi salvato in un file .csv, mantenendo l'associazione con l'host_id di riferimento.

Il risultato di questa fase è stato un primo dataset strutturato su due colonne principali: host_id e nome_host. Questo file ha rappresentato il punto di partenza per le successive attività di pulizia, filtraggio linguistico e analisi dei nomi.

```

from selenium import webdriver
from selenium.webdriver.chrome.options import Options
from selenium.webdriver.common.by import By
from selenium.webdriver.support.ui import WebDriverWait
from selenium.webdriver.support import expected_conditions as EC
import pandas as pd
import re

# === 1. Setup Chrome collegato a sessione già aperta con debug ===
options = Options()
options.debugger_address = "127.0.0.1:9222"
driver = webdriver.Chrome(options=options)
# === 2. Carica gli host_id dal CSV, senza duplicati ===
df = pd.read_csv("host_ids.csv", dtype=str)
host_ids = (
    df[["host_id"]]
    .dropna()
    .drop_duplicates()
    .astype(str)
    .str.replace(r"\.0$", "", regex=True)
)
# === 3. Liste per salvataggio ===
all_results = []
unique_first_names = set()
# === 4. Loop sugli host ===
for host_id in host_ids:
    try:
        url = f"https://www.airbnb.it/users/show/{host_id}"
        driver.get(url)
        # Cerca il bottone con aria-label contenente il nome
        name_element = WebDriverWait(driver, 10).until(
            EC.presence_of_element_located((
                By.XPATH,
                '//button[starts-with(@aria-label, "Scopri di più sullo")]'
            ))
        )
        aria_label = name_element.get_attribute("aria-label")
        raw_name = aria_label.strip().split()[-1]
        cleaned_name = re.sub(r"[^a-zA-ZàèéòùÀÈÉÒÙçç]", "", raw_name)
        first_name = cleaned_name.split()[0]
        # Salva nei risultati
        all_results.append({
            'host_id': host_id,
            'host_name': first_name
        })
        # Aggiungi alla lista unica
        unique_first_names.add(first_name)
        print(f"[v] {host_id} → {first_name}")
    except Exception as e:
        print(f"[x] Errore per host {host_id}: {e}")
        continue
# === 5. Salvataggio file ===
pd.DataFrame(all_results).to_csv("host_names_raw.csv", index=False)
print("Salvato file 'host_names_raw.csv' con tutti gli host con nome")
with open("unique_first_names.txt", "w") as f:
    for name in sorted(unique_first_names):
        f.write(name + "\n")
print(f"Nomi unici trovati: {len(unique_first_names)}")
print("Salvato file 'unique_first_names.txt'")

```

Figura 4.1: Codice Python per lo scraping dei nomi degli host dal sito Airbnb

4.2.2 Pulizia e normalizzazione dei nomi

I nomi ottenuti tramite lo scraping presentavano numerose irregolarità, per questo è stato necessario sviluppare lo script *puliziaNomiHost.py* (Figura 4.2), che ha permesso di uniformare i dati e ottenere un elenco di nomi coerente e privo di rumore.

La procedura di pulizia ha incluso le seguenti operazioni:

- **Rimozione di simboli, numeri e caratteri speciali**, come “_”, “-”, “@”, “#”, che non corrispondono a nomi propri.
- **Filtraggio dei nomi composti da una sola lettera**, in quanto considerati non informativi (ad esempio: “A”, “D”).
- **Mantenimento dei nomi di due lettere**, poiché alcuni risultano validi (es. Jo, Lu).
- **Esclusione dei nomi contenenti parole appartenenti a una blacklist** appositamente elaborata (es. home, guest, stay, house, srl, Torino, Interhome, ecc.), al fine di eliminare i nomi che rappresentano agenzie, marchi o denominazioni generiche anziché persone fisiche.
- **Conservazione dei nomi scritti interamente in maiuscolo**, in quanto potenzialmente rappresentativi di nomi propri o abbreviazioni di uso comune.

Il risultato è stato un dataset contenente esclusivamente nomi validi e utilizzabili per l’analisi di genere, con una significativa riduzione del rumore testuale.

```

import pandas as pd
import re

# Leggi il file Excel
df = pd.read_excel("host_names_completo.xlsx") # Cambia con il tuo nome file

# Blacklist
blacklist = {
    "srl", "home", "casa", "tour", "group", "agency", "company", "host", "italy",
    "rentals", "guest", "short", "stay", "torino", "house", "interhome", "vacanze",
    "holiday", "relax", "pool", "piscina", "agenzia", "hotel", "apartment", "apartments",
    "appartamento", "appartamenti", "flat", "flats", "piemonte", "piedmont", "airbnb",
    "residence", "residenza", "snc", "turin", "vacancy", "università", "university",
    "venaria", "villa", "villaggio", "village", "welcome", "benvenuti", "maison", "reggia",
    "properties", "property", "spa", "ssa", "sharing", "ss", "top", "team", "suite",
    "suites", "travel", "vista", "viaggio"
}

# Funzione per pulire i nomi
def is_valid_name(name):
    if not isinstance(name, str):
        return False
    name_clean = name.strip().lower()

    # Escludi nomi nella blacklist
    if any(term in name_clean for term in blacklist):
        return False

    # Escludi nomi di una sola lettera
    if len(name_clean) == 1:
        return False

    # Escludi nomi con numeri o simboli (solo lettere ammesse)
    if not re.fullmatch(r"[a-zA-Z]+", name_clean):
        return False

    return True

# Applica la funzione di filtro
df_cleaned = df[df['host_name'].apply(is_valid_name)]

# Salva il nuovo file Excel
df_cleaned.to_excel("host_names_completo_pulito.xlsx", index=False)

print(f"Pulizia completata. File salvato come 'host_names_completo_pulito.xlsx'")

```

Figura 4.2: Codice Python per la pulizia dei nomi degli host

4.2.3 Analisi del genere e della nazionalità tramite API NamSor

Dopo aver proceduto con la pulizia dei nomi degli host si è proceduto con la classificazione del genere e della nazionalità tramite l'utilizzo dell'API esterna NamSor, uno strumento di *onomastics-based AI classification* che consente di dedurre caratteristiche demografiche a partire dai nomi propri. Lo script *namsor_request.py* (Figura 4.3) ha inviato richieste automatizzate all'endpoint dell'API NamSor, in lotti di 1000 nomi per volta, rispettando i limiti di chiamata giornalieri previsti dal piano gratuito del servizio. Ogni richiesta restituiva una stima probabilistica del genere e del paese di origine del nome.

L'output restituito da NamSor includeva i campi principali presentati nella Tabella 4.1.

Le chiamate che hanno restituito errori o risultati nulli sono state gestite tramite istruzioni di exception handling, evitando l'interruzione dello script. Il file risultante è stato salvato in formato .csv ed in seguito è stato effettuato un processo di *merge*

Name	Likely Gender	Gender Scale	Country Origin	Probability Calibrated
Anna	female	0.96	IT	0.98
Luca	male	0.95	IT	0.97
Francois	male	0.95	FR	0.96

Tabella 4.1: Output restituito da NamSor

tramite il comando CERCA.VERT di Excel che ha permesso di unire il file ottenuto con quello contenente i dati panel di AirDNA. Il dataset finale contiene oltre 530.000 osservazioni, ciascuna riferita a un annuncio univoco, con informazioni complete su:

- Identificativo annuncio (property_id).
- Identificativo host (host_id).
- Nome host (host_name).
- Genere.
- Nazionalità.
- Caratteristiche dell'annuncio.
- Indicatori di performance economica (es. ADR, RevPAN, OCC).

```

import requests
import pandas as pd
import time

API_KEY = "0934fcb11f0b62de4ea2d5f4b200e2e6"

with open("unique_first_names.txt", "r", encoding="utf-16") as file:
    names = [line.strip() for line in file if line.strip()]

results = []

# === Invia le richieste a NamSor (1 richiesta al secondo) ===
for name in names:
    try:
        response = requests.get(
            f"https://v2.namsor.com/NamSorAPIV2/api2/json/origin/{name}",
            headers={"X-API-KEY": API_KEY}
        )
        if response.status_code == 200:
            data = response.json()
            origin = data.get("likelyOrigin", "unknown")
            results.append({"name": name, "origin": origin})
            print(f"[+] {name} → {origin}")
        else:
            print(f"[x] Errore API per {name} → status {response.status_code}")
    except Exception as e:
        print(f"[x] Errore per {name}: {e}")
    time.sleep(1) # Rispetta il rate limit

# === Salvataggio dei risultati ===
df = pd.DataFrame(results)
df.to_csv("name_origin_namsor.csv", index=False)
print("File 'name_origin_namsor.csv' salvato con successo.")

```

Figura 4.3: Codice Python per la classificazione dei nomi degli host tramite NamSor

4.2.4 Analisi della foto profilo degli host

Nel dataset di 530.000 annunci composto nella fase di scraping si individuano oltre 12.200 differenti host operanti nella provincia di Torino tra il 2017 e il 2024. Di questi si seleziona un campione con livello di confidenza del 95% randomizzato tramite comando Excel, ottenendo un totale di 370 host di cui analizzare le foto profilo. In seguito è stato sviluppato un script Python con Selenium, presentato in Figura 4.4, che, host per host, aprisse il profilo dell'host permettendo di visualizzare la rispettiva foto profilo. Una volta visualizzata l'immagine questa viene analizzata visivamente registrando in un dataset la presenza dei segnali visivi di interesse. Sono stati individuati sei segnali visivi utili per l'analisi:

- **Face:** che indica la visibilità del volto dell'host nella foto.
- **Smile:** che indica che l'host sorride nella foto profilo.
- **Sunglasses:** Indica la presenza di occhiali da sole.
- **Group:** presenza di una foto di gruppo, ossia con più di una persone.
- **Animal:** presenza di animali nella foto profilo.

- **Children:** presenza di bambini nella foto profilo.

Anche in questo caso le informazioni ottenute sono state unite con i dati panel tramite una operazione di merge.

```

from selenium import webdriver
from selenium.webdriver.chrome.options import Options
import pandas as pd
import time

# === Connessione a Chrome già aperto in modalità debug ===
options = Options()
options.debugger_address = "127.0.0.1:9222"
driver = webdriver.Chrome(options=options)

# === Carica solo la prima colonna del CSV (host_id) ===
try:
    df = pd.read_csv("host_ids.csv", sep=";", usecols=[0], names=["host_id"], header=0, dtype=str)
    host_ids = df["host_id"].dropna().drop_duplicates().astype(str).str.replace(r"\.0$","", regex=True)
    print(f"[INFO] Trovati {len(host_ids)} host_id.")
except Exception as e:
    print(f"[ERRORE] Problema con il file host_ids.csv: {e}")
    exit()

# === Visualizza ogni profilo per 20 secondi ===
for i, host_id in enumerate(host_ids, start=1):
    try:
        url = f"https://www.airbnb.it/users/show/{host_id}"
        driver.get(url)
        print(f"[{i}]/[{len(host_ids)}] Visualizzando host {host_id}: {url}")
        time.sleep(20) # Attesa automatica di 20 secondi
    except Exception as e:
        print(f"[{i}] Errore con host {host_id}: {e}")
        continue

print("\nCompletato.")

```

Figura 4.4: Codice Python per l’analisi delle foto profilo degli host

4.2.5 Considerazioni metodologiche e limiti

L’approccio adottato offre diversi vantaggi. Da un lato, la combinazione tra scraping automatizzato e analisi linguistica ha permesso di arricchire un dataset che, in origine, non conteneva alcuna informazione socio-demografica. Questo ha reso possibile introdurre una prospettiva nuova nello studio di Airbnb. Inoltre, la struttura del lavoro è stata pensata in modo flessibile: ogni passaggio – dall’estrazione dei dati alla pulizia, fino alla classificazione e *merge* finale – è organizzato in script Python separati, così da poter essere aggiornato o riutilizzato con facilità.

Vanno però anche riconosciuti alcuni limiti. L’assegnazione del genere basata esclusivamente sul nome può generare errori, soprattutto in presenza di nomi ambigui, pseudonimi o provenienti da contesti culturali molto diversi. Allo stesso modo, il numero di osservazioni finali risulta inferiore rispetto al dataset di partenza, poiché è stato necessario escludere i casi non interpretabili. Anche l’informazione sulla nazionalità si basa su stime probabilistiche e, di conseguenza, non può essere considerata un dato perfettamente certo.

Nonostante queste cautele, il risultato complessivo restituisce comunque un quadro affidabile del profilo socio-demografico degli host attivi nella provincia di Torino tra il 2017 e il 2024, offrendo una base solida per analisi statistiche e modelli di regressione sul divario di genere nelle performance economiche su Airbnb.

4.3 Misure di performance

Per studiare la possibile presenza di differenze performance, viene effettuata un'analisi delle prestazioni degli host presenti sulla piattaforma tramite il calcolo alcuni indicatori di performance (KPIs), ritenuti particolarmente significativi nel settore dell'ospitalità:

1. **Prenotazioni (number of reservations)**: somma delle prenotazioni che l'host ha ricevuto per i suoi annunci.

Utilizzato nel paper "The Impacts of Quality and Quantity Attributes of Airbnb Hosts on Listing Performance" [56] di K. L. Xie e Z. Mao del 2017.

2. **Rating**: media degli overall rating degli annunci dell'host.

Utilizzato nel paper "The Impacts of Quality and Quantity Attributes of Airbnb Hosts on Listing Performance" [56] di K. L. Xie e Z. Mao del 2017.

3. **ADR (Average Daily Rate)**: rappresenta il prezzo medio pagato per notte prenotata, ed è uno degli indicatori maggiormente utilizzati nel settore.

Utilizzata nel paper "Gender earning gap on digital platforms: The Airbnb case" [13] di Ert, Fleischer e Kopolovich del 2024.

L'ADR si calcola con la seguente formula:

$$ADR = \frac{Revenue}{ReservationDays}$$

dove:

- *Revenue*: ricavi complessivi generati dall'host;
- *ReservationDays*: numero totale di notti effettivamente occupate.

4. **OCC (Occupancy Rate)**: misura la percentuale di occupazione delle proprietà, confrontando le notti prenotate con quelle rese disponibili in un certo periodo. Valori più alti indicano una maggiore capacità di occupare gli spazi. Utilizzata nel paper "Gender earning gap on digital platforms: The Airbnb case" [13] di Ert, Fleischer e Kopolovich del 2024 e nel paper "How to react to a shock? Effects of Airbnb hosts' choices and market segmentation at the time of Covid-19" [57] di Buzzacchi et al. del 2023.

La formula di calcolo è:

$$OCC = \frac{ReservationDays}{ReservationDays + AvailableDays}$$

dove:

- *ReservationDays*: notti riservate dagli ospiti;

- *AvailableDays*: notti messe a disposizione per la prenotazione.

5. **RevPAN (Revenue per Available Night)**: esprime il rapporto tra i ricavi e il numero totale di notti disponibili, indipendentemente dal fatto che siano state prenotate o meno. Un valore alto segnala una buona capacità di ottimizzazione dei profitti, mentre un valore basso indica uno scarso sfruttamento delle possibilità di profitti fornite da Airbnb.

Utilizzato nel paper "How to react to a shock? Effects of Airbnb hosts' choices and market segmentation at the time of Covid-19" [57] di Buzzacchi et al. del 2023.

La formula di riferimento è:

$$RevPAN = \frac{Revenue}{ReservationDays + AvailableDays} = ADR \times OCC$$

dove:

- *ADR*: tariffa media per notte;
- *OCC*: percentuale di occupazione delle notti disponibili.

4.4 Strumenti per l'analisi

4.4.1 Excel

Per l'analisi e la gestione del dataset è stato utilizzato Microsoft Excel, che ha reso possibile l'unione dei diversi dataset contenenti informazioni sugli annunci, sugli host e sulle caratteristiche visive. Il software si è rivelato molto utile grazie alle funzionalità integrate per la manipolazione e la pulizia del dato. Ho usato soprattutto Power Query, una funzione già inclusa in Excel che si rivela utilissima quando si lavora con tanti dati sparsi in file diversi. Ti permette di importare, sistemare e combinare tutto in modo quasi automatico, senza dover fare mille passaggi manuali.

Grazie a questo strumento, l'aggiornamento dei dati è diventato molto più rapido. Ogni volta che cambiavo qualcosa nei file originali, bastava un clic per ricaricare tutto e avere le tabelle aggiornate e coerenti tra loro. In più, Power Query ha garantito che le informazioni restassero uniformi e ordinate, evitando discrepanze o errori tra le diverse fonti. Attraverso l'utilizzo di funzioni di unione e ricerca, come CERCA.VERT, è stato possibile collegare le diverse tabelle, ottenendo un unico file coerente e analizzabile. Inoltre, mediante la creazione di tabelle pivot, è stato possibile sviluppare analisi descrittive sintetiche e immediate. Le tabelle pivot, insieme ai grafici a torta e agli histogrammi, hanno consentito di rappresentare i dati in modo chiaro e intuitivo, offrendo sia una visualizzazione statica, sia una

dinamica nel tempo.

Tramite queste descrittive è stato possibile individuare rapidamente le principali tendenze e distribuzioni all'interno del dataset, facilitando la successiva fase di interpretazione dei risultati.

4.4.2 Stata

Per lo sviluppo delle regressioni e l'analisi delle t di Student è stato utilizzato il software Stata, che ha avuto un ruolo centrale non solo nella parte di stima dei modelli, ma anche in una fase fondamentale di pulizia e preparazione dei dati. Prima dell'elaborazione statistica, infatti, è stato necessario uniformare e trasformare numerose variabili, in particolare convertendo alcune in formato booleano (dummy variables) per poterle impiegare correttamente nelle regressioni. Stata si è rivelato particolarmente utile per la gestione di dataset di grandi dimensioni e per la possibilità di combinare in modo efficiente operazioni di pulizia, manipolazione e analisi. Questa fase ha permesso di rendere il dataset coerente e strutturato, garantendo la qualità delle informazioni utilizzate nei modelli successivi.

Dopo la sistemazione dei dati, si è iniziato a costruire diversi modelli di regressione lineare (OLS). L'obiettivo era capire come certe caratteristiche, come la nazionalità e il genere dell'host, oppure le proprietà degli annunci, influenzassero le prestazioni della struttura, misurate in termini di prenotazioni, prezzo medio, tasso di occupazione e valutazioni.

Per alcune di queste variabili dipendenti è stata usata una trasformazione logaritmica. È una scelta molto utile, perché consente di interpretare meglio i risultati in chiave percentuale. Inoltre, aiuta a ridurre l'impatto dei valori estremi, che spesso possono sporcare le analisi e rendere i modelli meno affidabili.

In sostanza, questa fase ha permesso di andare oltre la semplice descrizione dei dati e di capire quali fattori incidono davvero sulle performance, fornendo una base più solida per trarre conclusioni e confrontare i risultati tra vari gruppi di host. I modelli includono inoltre un insieme di effetti fissi temporali, inseriti per controllare l'influenza di fattori stagionali o tendenze comuni nel tempo. Infine, oltre ai modelli di base, sono state considerate anche interazioni tra variabili (ad esempio tra genere e nazionalità, o tra caratteristiche visive e nazionalità) per analizzare eventuali differenze di comportamento o di performance tra gruppi di host. Tutte le stime sono state effettuate con errori standard robusti per assicurare maggiore affidabilità dei risultati.

4.5 Analisi delle descrittive

L'analisi delle statistiche descrittive consente di capire come sono distribuite le variabili più importanti del dataset, così da avere un'idea più chiara delle caratteristiche degli host e di come si comportano su Airbnb.

Il focus è stato posto sulle variabili più utili per capire il fenomeno: da quelle legate al profilo demografico degli host, fino agli indicatori di performance come il tasso di risposta, il tasso di occupazione, l'ADR (cioè la tariffa media per notte), il RevPAN e le valutazioni medie ricevute dagli ospiti. In pratica, questa prima analisi serve a costruire una sorta di mappa generale del campione: permette di vedere subito se ci sono dei modelli che si ripetono o delle differenze interessanti tra gruppi di host diversi.

Questa attività sarà poi utile per andare più a fondo, nelle sezioni successive, dove verranno utilizzati test statistici e modelli di regressione per confermare o approfondire le tendenze osservate finora.

4.5.1 Distribuzione degli annunci per nazionalità e genere

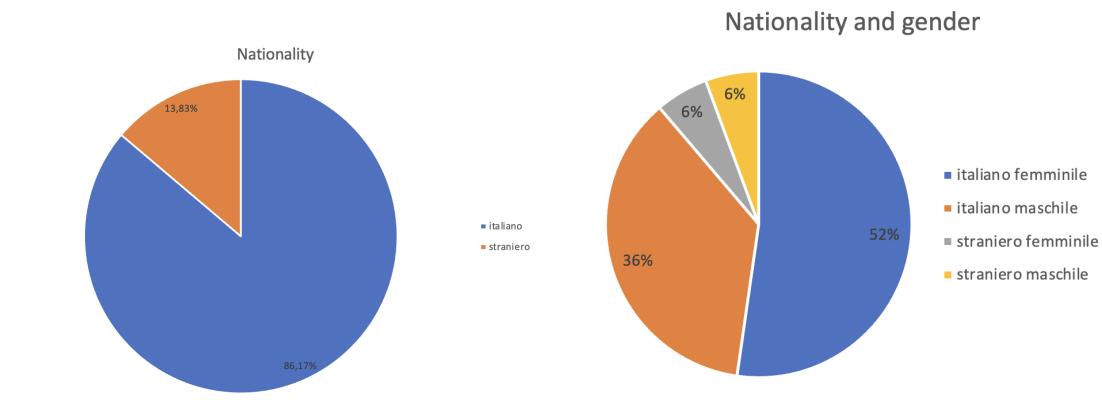


Figura 4.5: Distribuzione annunci

Il dataset analizzato comprende 530.960 osservazioni, corrispondenti ad altrettanti annunci pubblicati sulla piattaforma Airbnb.

La Figura 4.5 mostra la distribuzione per nazionalità: l'86,2% degli annunci appartiene a host italiani, mentre solo il 13,8% è gestito da host stranieri. Ciò conferma la natura fortemente domestica del mercato, dove l'offerta è dominata da operatori locali con una presenza straniera ancora limitata.

Combinando la dimensione della nazionalità con quella del genere, si osserva che oltre la metà degli annunci (52%) è gestita da donne italiane, seguite dagli uomini

italiani (36%), mentre le quote di host stranieri femminili e maschili risultano entrambe marginali, pari rispettivamente al 6% ciascuna.

In sintesi, la maggioranza degli annunci su Airbnb è riconducibile a donne italiane, mentre la componente straniera, seppur presente, incide solo in modo residuale sul totale.

4.5.2 Distribuzione degli host per nazionalità e genere

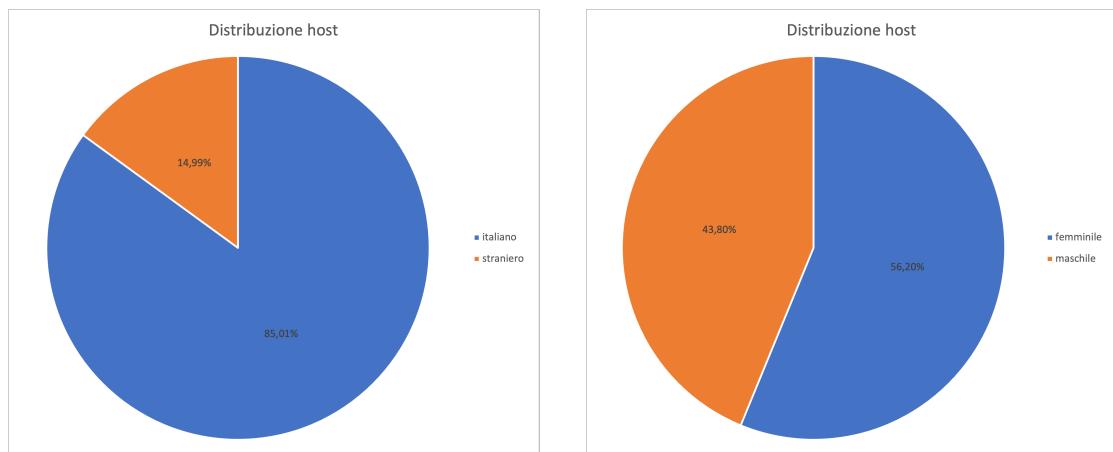


Figura 4.6: Distribuzione host

Il dataset analizzato comprende 12.990 host classificati.

La Figura 4.6 mostra la distribuzione per nazionalità e per genere. La stragrande maggioranza degli host è di nazionalità italiana (85,0%), mentre solo il 15,0% è costituito da host stranieri. Questa composizione riflette la forte presenza di operatori locali nel mercato italiano di Airbnb, dove la gestione diretta da parte di residenti rappresenta ancora la forma predominante di offerta.

La distribuzione per genere evidenzia una leggera prevalenza femminile: il 56,2% degli host è costituito da donne, mentre il 43,8% da uomini.

Questo dato suggerisce una maggiore partecipazione femminile nel mercato della locazione turistica, coerente con l'ipotesi secondo cui le donne risultano più attive in attività che richiedono attenzione all'accoglienza e alla gestione della relazione con gli ospiti.

Nel complesso, la popolazione di host risulta quindi prevalentemente italiana e femminile, un aspetto che sarà utile considerare nell'interpretazione delle analisi successive relative alle performance economiche e reputazionali (ADR, RevPAN, rating, ecc.).

4.5.3 Distribuzione dei ricavi medi



Figura 4.7: Distribuzione dei Ricavi medi

La Figura 4.7 illustra i ricavi medi annuali suddivisi tra host italiani e host stranieri. Nel complesso, si nota che gli host italiani ottengono in media ricavi più alti rispetto agli host stranieri: 664,80 dollari contro 639,87 dollari.

Guardando al loro andamento nel tempo, la distanza tra i due gruppi era piuttosto evidente nei primi anni del campione (2017–2019), ma col passare del tempo si è via via ridotta, fino quasi a scomparire nel periodo più recente, tra il 2022 e il 2024. Questo andamento lascia intendere che il mercato di Airbnb stia trovando un punto di equilibrio, dove le opportunità e le performance economiche degli host risultano sempre più simili, indipendentemente dalla nazionalità.

Per verificare se la differenza nei guadagni osservata è effettivamente significativa o frutto del caso, è stato eseguito un t-test per campioni indipendenti, utile a valutare la validità statistica del confronto.

Il test restituisce un valore di:

$t = -5.191$, $p < 0.001$,

con una differenza media di -24.9 USD, a favore degli host italiani.

Si rifiuta quindi l'ipotesi nulla di uguaglianza delle medie, concludendo che gli host italiani ottengono in media ricavi significativamente superiori rispetto a quelli stranieri.

4.5.4 Distribuzione del Rating medio

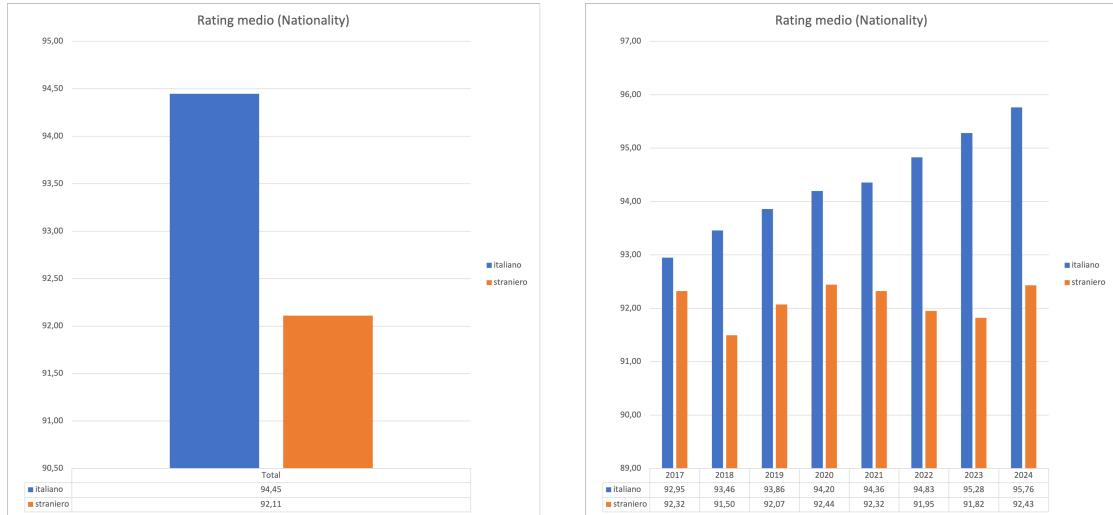


Figura 4.8: Distribuzione del Rating medio

Il grafico successivo Figura 4.8 mostra la distribuzione del rating medio in base alla nazionalità dell'host. Nel complesso, gli host italiani presentano valutazioni più elevate (94,45) rispetto agli host stranieri (92,11). Tale differenza è visibile per tutti gli anni considerati, anche se si osserva un lieve riavvicinamento dei valori negli ultimi anni del periodo.

Il test t condotto per verificare la significatività di tale differenza restituisce:
 $t = -56.9634$, $p < 0.001$

La differenza media stimata è di $-2,34$ punti, a vantaggio degli host italiani.
Si rifiuta pertanto l'ipotesi nulla di uguaglianza dei punteggi medi, concludendo che gli host italiani ottengono valutazioni significativamente superiori rispetto a quelli stranieri.

4.5.5 Distribuzione del numero medio di prenotazioni

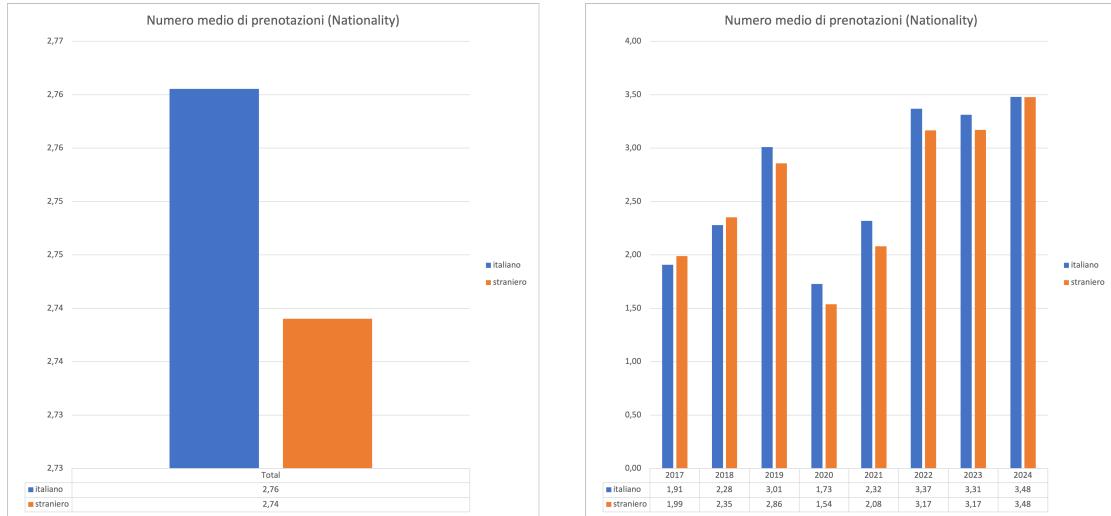


Figura 4.9: Distribuzione del Rating medio

La Figura 4.9 mostra il confronto tra host italiani e host stranieri. In questo caso, le differenze appaiono molto più contenute: i ricavi medi si attestano rispettivamente a 2,76 e 2,74 prenotazioni per annuncio.

Dall'andamento temporale emerge una sostanziale convergenza dei valori, con oscillazioni minime da un anno all'altro e una leggera prevalenza alternata dei due gruppi in periodi differenti.

Il t-test conferma l'assenza di una differenza statisticamente significativa tra le due medie:

$$t = -1.3914, p = 0.1641$$

La differenza stimata è di appena $-0,02$ prenotazioni, non significativa dal punto di vista statistico.

Pertanto, non si rifiuta l'ipotesi nulla di uguaglianza delle medie: si conclude che la nazionalità dell'host non ha un effetto statisticamente rilevante sul numero medio di prenotazioni.

4.5.6 Distribuzione della tipologia di alloggio

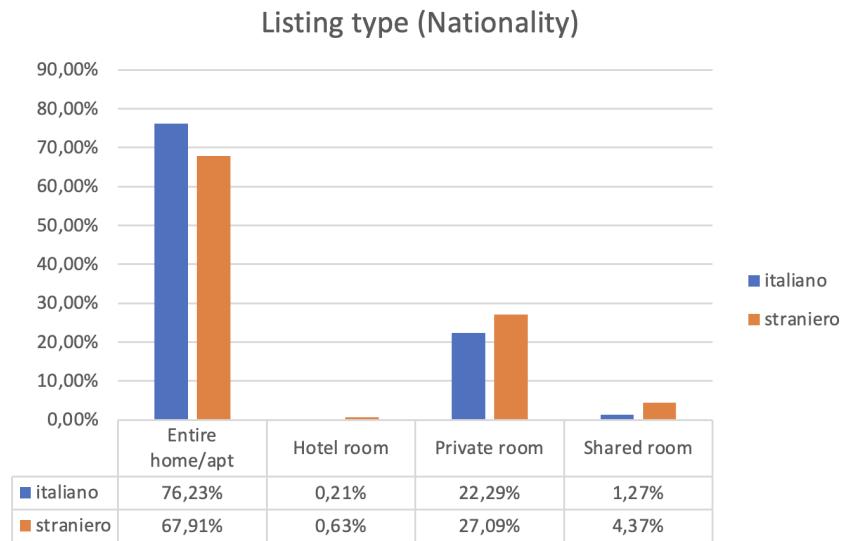


Figura 4.10: Distribuzione dei Listing Type

La Figura 4.10 mette in evidenza come gli host italiani tendano a offrire più spesso case o appartamenti interi, che rappresentano circa il 76,2% delle loro inserzioni, rispetto al 67,9% degli host stranieri.

Al contrario, gli host stranieri sembrano preferire soluzioni più condivise, come stanze private o camere in condivisione, che nel complesso arrivano al 31,5% delle loro offerte (di cui 27,09% stanze private e 4,37% condivise), contro un 23,5% circa nel caso degli host italiani.

In altre parole, mentre gli host italiani puntano maggiormente su strutture indipendenti e complete, gli host stranieri mostrano una tendenza leggermente più forte verso formule di ospitalità condivisa, forse per una diversa strategia di mercato o per caratteristiche abitative differenti.

Anche la quota di camere d'hotel risulta leggermente superiore tra gli host stranieri (0,63% contro 0,21%).

Questi risultati evidenziano una differente strategia di gestione degli alloggi:

- gli host italiani tendono a privilegiare soluzioni più “professionali” o destinate interamente alla locazione turistica, probabilmente legate a proprietà secondarie o investimenti immobiliari dedicati;
- gli host stranieri, invece, appaiono più inclini a condividere spazi della propria abitazione, configurandosi quindi come host occasionali o a carattere domestico.

4.5.7 Distribuzione dello status Superhost

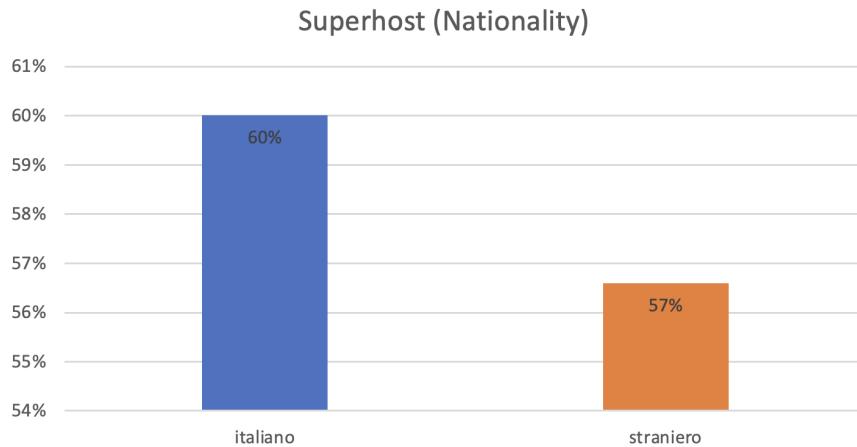


Figura 4.11: Distribuzione dello status Superhost

La Figura 4.11 illustra la distribuzione dello status di Superhost tra host italiani e stranieri.

In questo caso, la differenza è più marcata: il 60% degli host italiani risulta Superhost, contro appena il 57% degli host stranieri.

Questa discrepanza suggerisce che gli host italiani mostrano maggiori livelli di performance complessiva, riuscendo più frequentemente a soddisfare i requisiti qualitativi richiesti dalla piattaforma.

Il t-test per campioni indipendenti conferma la significatività statistica di questa differenza:

$$t = -44.8565, p < 0.001$$

La differenza media stimata è pari a -0,081 punti percentuali, a vantaggio degli host italiani.

Si rifiuta quindi l'ipotesi nulla, concludendo che gli host italiani sono significativamente più rappresentati nella categoria Superhost rispetto agli stranieri.

Tale risultato potrebbe essere legato a diversi fattori, tra cui:

- una maggiore continuità nell'attività di hosting da parte degli italiani, che può favorire il mantenimento dei requisiti richiesti (numero minimo di soggiorni, tasso di risposta, recensioni positive);
- una migliore conoscenza del mercato locale e delle aspettative dei turisti;

- una maggiore stabilità operativa, legata alla gestione di immobili di proprietà piuttosto che di strutture occasionali.

4.5.8 Distribuzione del numero medio di foto

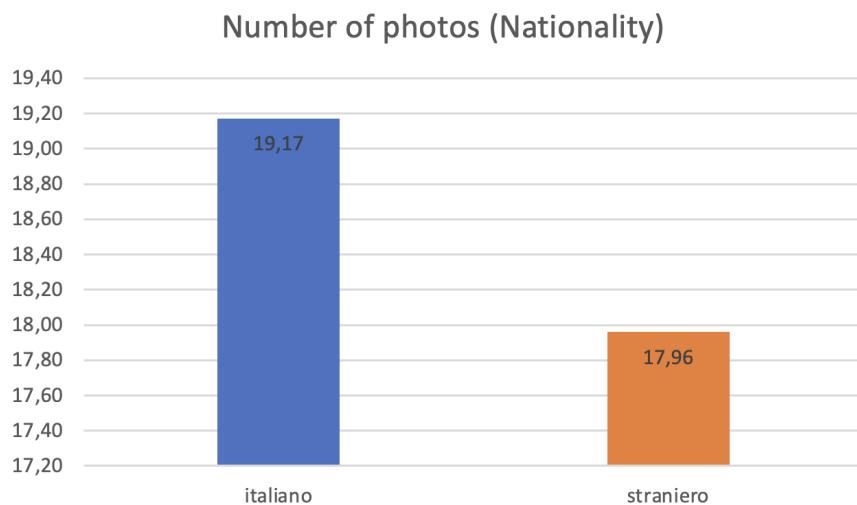


Figura 4.12: Distribuzione del numero medio di foto

Il confronto per nazionalità Figura 4.12 mostra un pattern analogo: gli host italiani pubblicano mediamente 19,17 foto per annuncio, mentre gli host stranieri ne inseriscono 17,96.

Anche in questo caso, la differenza è confermata dal t-test, che restituisce:

$t = -19.9376$, $p < 0.001$

La differenza media è di -1,21 foto, a vantaggio degli host italiani.

Questa evidenza indica che gli host italiani tendono a investire maggiormente nella componente visiva dell'offerta, probabilmente consapevoli del ruolo chiave che le fotografie rivestono nel processo di scelta degli utenti.

Gli host stranieri, invece, potrebbero affidarsi più frequentemente a modelli standardizzati o limitarsi a un numero ridotto di immagini, privilegiando la quantità di annunci alla qualità della presentazione.

4.5.9 Distribuzione della politica di cancellazione

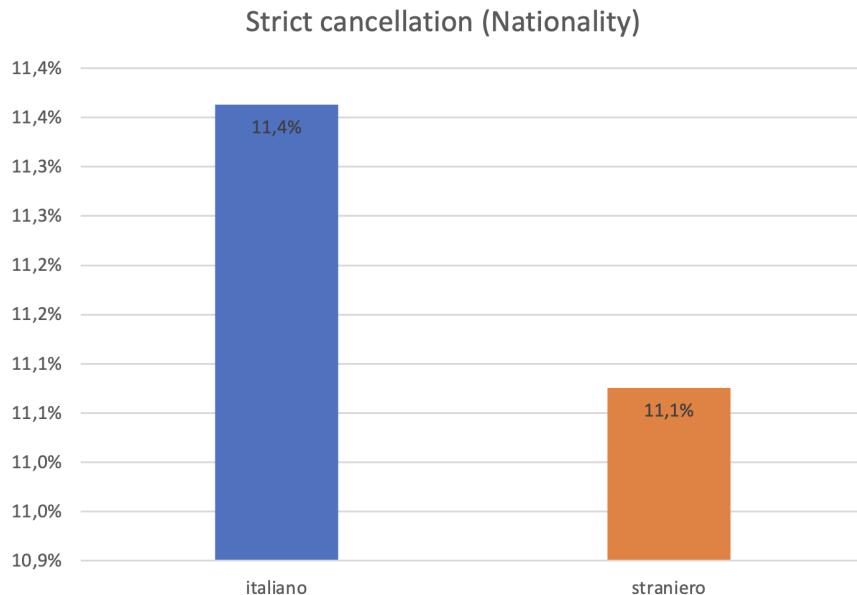


Figura 4.13: Distribuzione della politica di cancellazione

Il secondo confronto Figura 4.13 mostra la distribuzione della politica di cancellazione rigida in base alla nazionalità dell'host.

Gli host italiani presentano una leggera prevalenza nell'adozione di questa policy (11,4%) rispetto agli host stranieri(11,1%).

La differenza, pur minima, risulta statisticamente significativa secondo il test t: $t = -2.0270$, $p = 0.0427$

La differenza media è pari a $-0,0029$, cioè circa lo 0,3% a favore degli host italiani. Questo suggerisce una tendenza marginale degli host italiani verso pratiche più rigide, probabilmente legate a una maggiore stagionalità del mercato turistico locale o a un comportamento più conservativo nella gestione delle prenotazioni.

Tale risultato, seppur di entità modesta, conferma la maggiore prudenza operativa degli host italiani rispetto agli stranieri, che appaiono leggermente più flessibili nel rispondere alla variabilità della domanda.

4.5.10 Distribuzione delle caratteristiche visive

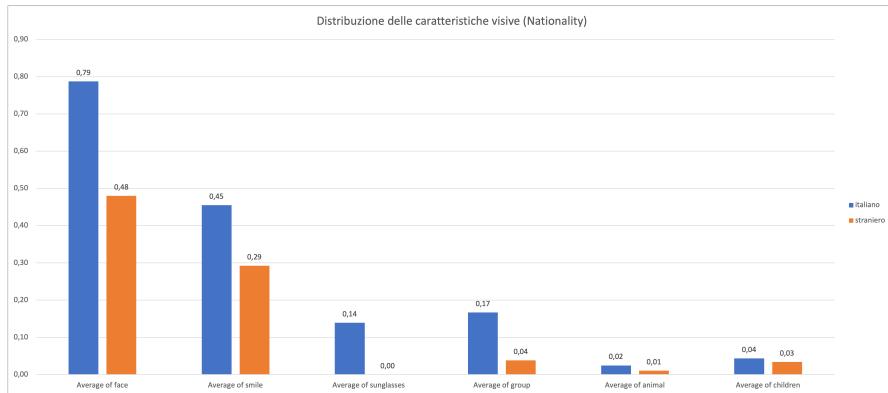


Figura 4.14: Distribuzione delle caratteristiche visive

La Figura 4.14 analizza la distribuzione delle stesse caratteristiche in base alla nazionalità dell'host (italiano vs straniero). Anche in questo caso, la presenza del volto rappresenta l'elemento più ricorrente, ma con una differenza marcata: gli host italiani mostrano il viso nel 79% dei casi, contro il 48% degli host stranieri. Analoga tendenza emerge per il sorriso, più frequente tra gli host italiani (0,45) rispetto agli stranieri (0,29), indicando una maggiore attenzione alla dimensione relazionale e all'espressione di cordialità.

Le foto con occhiali da sole compaiono solo tra gli host italiani (0,14), mentre gli stranieri tendono a evitarle completamente. Anche le foto di gruppo risultano più comuni tra gli host italiani (0,17 contro 0,04), mentre le immagini con animali o bambini presentano valori bassi e poco differenziati tra i due gruppi (inferiori a 0,05).

Nel complesso, gli host italiani sembrano adottare una comunicazione visiva più esplicita e personale, privilegiando la visibilità del volto e il sorriso, mentre gli host stranieri mostrano una tendenza maggiore alla riservatezza o a uno stile più neutro.

4.5.11 Distribuzione dell'ADR medio

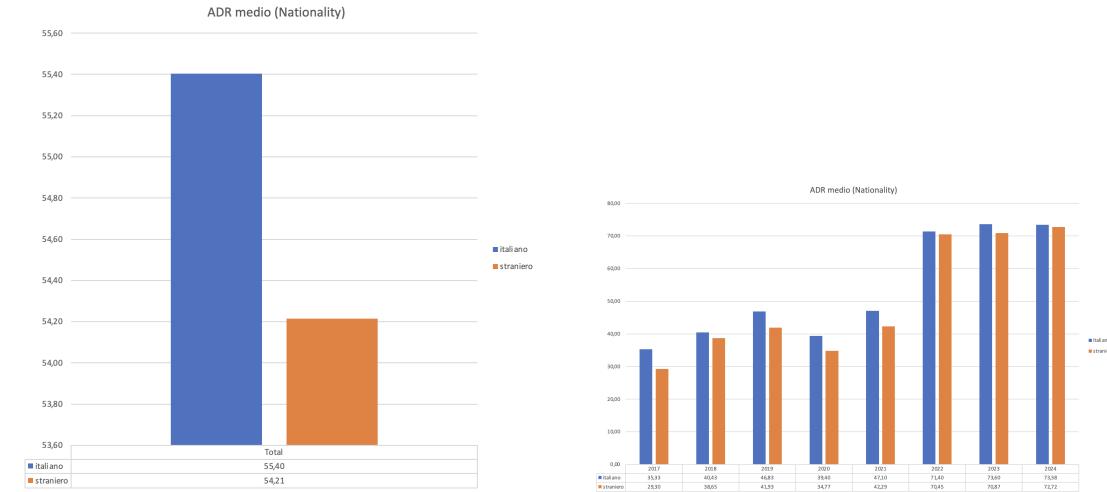


Figura 4.15: Distribuzione dell'ADR medio

La nazionalità incide in modo significativo sul prezzo medio giornaliero.

Gli host italiani applicano un ADR medio di 88,96 USD, leggermente superiore a quello degli host stranieri, pari a 87,23 USD.

Il t-test conferma la significatività della differenza:

$$t = 4.0895, p < 0.001$$

La differenza media è di 1,73 USD, a vantaggio degli host italiani.

Sebbene la distanza tra i due gruppi sia inferiore rispetto a quella osservata per il genere, il risultato suggerisce che gli host italiani tendono a posizionare i propri alloggi su fasce di prezzo marginalmente più alte, probabilmente in relazione a una migliore conoscenza del mercato locale e a un maggiore controllo diretto sulla gestione delle tariffe.

Il grafico temporale riportato in Figura 4.15 mostra chiaramente che, negli ultimi anni, il divario tra host italiani e stranieri si è progressivamente ridotto. Questa evoluzione suggerisce una convergenza dei prezzi medi, probabilmente legata alla crescente competizione nel mercato e alla diffusione di strumenti di pricing automatizzato, che contribuiscono a rendere le tariffe più allineate e flessibili.

4.5.12 Distribuzione dell'OCC medio

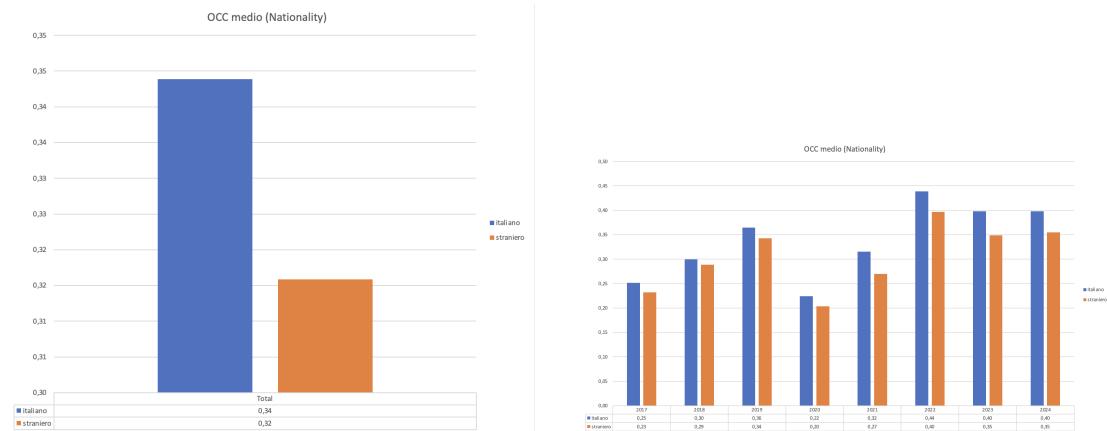


Figura 4.16: Distribuzione dell'OCC medio

In Figura 4.16, invece, si osserva che gli host italiani registrano un tasso medio di occupazione pari a 0,3439, leggermente superiore rispetto a quello degli host stranieri, che si ferma a 0,3157.

La differenza è pari a 0,028 punti percentuali e risulta statisticamente significativa $t = -17.2087$, $p < 0.001$

Anche in questo caso si rifiuta l'ipotesi nulla, concludendo che la nazionalità dell'host ha un effetto significativo sull'occupazione degli alloggi.

Gli host italiani sembrano riuscire a ottenere un maggiore tasso di occupazione, probabilmente grazie a una migliore conoscenza del contesto locale, della domanda turistica stagionale e di pratiche di gestione più mirate al mercato nazionale.

Il grafico temporale mostra un vantaggio costante per gli host italiani lungo tutto il periodo 2017–2024, pur con un lieve ridimensionamento della distanza negli anni più recenti, segnale di un processo di convergenza legato alla crescente professionalizzazione anche degli host stranieri.

4.5.13 Distribuzione del RevPAN medio

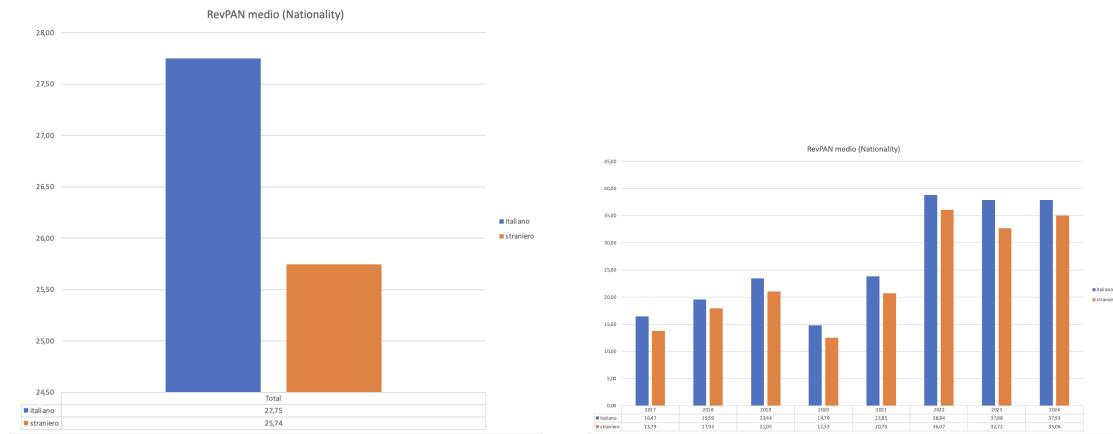


Figura 4.17: Distribuzione del RevPAN medio

La nazionalità dell'host incide significativamente anche sul RevPAN medio.

Gli host italiani mostrano un RevPAN medio pari a 43,69 USD, contro i 42,24 USD degli host stranieri.

Ciò suggerisce che i proprietari italiani riescano, in media, a ottenere rendimenti più elevati dai propri annunci rispetto agli host non italiani. Il t-test conferma la significatività della differenza:

$$t = -5.6176, p < 0.001$$

La differenza media è di 1,44 USD a favore degli host italiani.

Anche in questo caso, l'ipotesi nulla di uguaglianza delle medie viene rifiutata, indicando un effetto statisticamente rilevante della nazionalità sui ricavi medi.

Osservando l'evoluzione temporale, il gap tra italiani e stranieri appare persistente ma tende a ridursi negli anni più recenti. Questo può riflettere un processo di convergenza nelle performance economiche tra i due gruppi, favorito dalla crescente professionalizzazione degli host stranieri e dall'adozione di strumenti di revenue management e pricing dinamico che riducono le differenze legate alla conoscenza locale del mercato.

4.5.14 Andamento delle metriche chiave

La Tabella 4.2 mostra, anno per anno, come sono cambiate le tre metriche selezionate per l'analisi dell'attività su Airbnb dal 2017 al 2024: quante prenotazioni in media vengono fatte, i ricavi per notte attiva (RevPAN) e che voto medio ricevono gli host.

Nel complesso, si nota un miglioramento costante nel numero medio di prenotazioni:

si parte da 1,88 nel 2017 e si arriva a 3,47 nel 2024. C'è però un calo evidente nel 2020, probabilmente a causa della pandemia, che ha rallentato il settore turistico. Dopo quella parentesi complicata, l'attività è tornata a crescere con un ritmo regolare.

Anche i ricavi per notte seguono una linea simile. Nel 2020 si vede un calo, ma poi riprendono rapidamente a salire, toccando il massimo nel 2022 con 56,38 dollari. Negli ultimi due anni, 2023 e 2024, i guadagni si mantengono stabili, attorno ai 53 dollari, segno di una situazione ormai consolidata.

Per quanto riguarda i voti complessivi lasciati dagli ospiti, la storia è diversa: non ci sono grandi variazioni. Le valutazioni restano alte, tra 92,5 e 94,3 su 100. Questo suggerisce che, anche nei periodi difficili, gli host hanno continuato a offrire un servizio di buona qualità, mantenendo un livello di soddisfazione alto e costante. In sintesi, la tabella racconta un percorso chiaro: un momento di crisi nel 2020 seguito da una solida ripresa, con risultati economici in crescita e una qualità dell'esperienza che non ha mai smesso di essere positiva.

Tabella 4.2: Statistiche descrittive per anno

	Year	Mean	Std. Dev.	Min	Max
Number of Reservations					
	2017	1.88	2.93	0	24
	2018	2.29	3.25	0	25
	2019	2.99	3.72	0	27
	2020	1.77	2.84	0	31
	2021	2.35	3.24	0	24
	2022	3.41	3.70	0	30
	2023	3.32	3.61	0	27
	2024	3.47	3.72	0	26
Revenues per Active Night (RevPAN)					
	2017	\$32.36	\$27.72	\$0.00	\$526.20
	2018	\$34.52	\$30.58	\$0.00	\$706.27
	2019	\$35.99	\$31.34	\$0.00	\$1,134.00
	2020	\$28.69	\$31.52	\$0.00	\$751.00
	2021	\$43.82	\$85.88	\$0.00	\$4,556.74
	2022	\$56.38	\$80.38	\$0.00	\$4,858.00
	2023	\$53.60	\$50.19	\$0.00	\$2,570.00
	2024	\$53.18	\$46.91	\$0.00	\$1,601.33
Overall Rating					
	2017	92.56	8.44	0	100
	2018	92.88	8.82	20	100
	2019	93.24	9.06	0	100
	2020	93.42	8.97	0	100
	2021	93.27	9.38	0	100
	2022	93.56	9.47	0	100
	2023	93.69	9.40	0	100
	2024	94.31	9.01	0	100

Capitolo 5

Risultati

Nel seguente capitolo è possibile trovare il fulcro della ricerca, esso ha l'obiettivo di tradurre in analisi quantitativa tutte le riflessioni teoriche fatte nei capitoli precedenti. In questa sezione ci si focalizzerà sui dati reali di Airbnb nella provincia di Torino con un'analisi volta ad indagare la presenza o meno di differenze di performance tra host italiani e host stranieri e individuare i fattori che le determinano. La presente ricerca è caratterizzata da due principali dimensioni di analisi:

- La prima è quella riferita alla stima di modelli econometrici che permettono di valutare l'impatto di genere e nazionalità propri dell'host su una lista di indicatori di performance, sia di base del dataset sia derivate da variabili base al suo interno: il numero di prenotazioni, la tariffa media giornaliera (Average Daily Rate ADR), sia il tasso di occupazione (Occupancy Rate OCC), il ricavo per notte disponibile (Revenue per Available Night, RevPAN) e la valutazione complessiva (Overall Rating).
L'obiettivo dell'analisi è quello di comprendere se e in che misura, a parità di caratteristiche strutturali e gestionali dell'alloggio, il genere e la nazionalità influenzano significativamente i risultati economici ottenuti dagli host;
- La seconda dimensione invece introduce un elemento quasi innovativo e cioè si concentra sul ruolo dei segnali visivi (*visual cues*) delle foto profilo degli host come possibili determinanti delle performance economiche e reputazionali.
Le fotografie sono uno strumento comunicativo capace di influenzare la percezione della fiducia e la propensione alla prenotazione da parte dei possibili ospiti. Per l'analisi sono state considerate caratteristiche visive come: la presenza del volto (face), il sorriso (smile), gli occhiali da sole (sunglasses), più persone (group), di bambini (children) o di animali (animal). Con queste variabili si può analizzare in che modo i segnali visivi della foto dell'host incidono sulle valutazioni ricevute e sulle performance economiche e anche sull'eventuale interazione con il genere.

Le analisi si basano su modelli di regressione multipla con effetti fissi nel tempo. Questo approccio permette di considerare anche le differenze che non si vedono direttamente tra i vari casi analizzati, così da capire con maggiore precisione come le diverse variabili si influenzino a vicenda.

Attraverso lo studio di caratteristiche personali e visive, il capitolo cerca di offrire una panoramica completa, fondata sui dati, delle dinamiche che si sviluppano all'interno della piattaforma Airbnb. L'idea è mostrare come elementi apparentemente secondari, come il genere, la nazionalità o persino la foto del profilo, possano avere un impatto concreto sui risultati economici degli host.

In questo modo, lo studio contribuisce ad arricchire il dibattito sulla discriminazione nelle piattaforme digitali, offrendo nuove prove riferite al contesto italiano. Allo stesso tempo, sottolinea quanto la reputazione e l'immagine personale contino nei mercati online, dove la fiducia gioca un ruolo decisivo.

5.1 Analisi dati panel

5.1.1 Regressione multivariata logaritmica

In questa sezione viene presentata la regressione individuata per lo studio delle differenze di performance tra host italiani e non italiani e host uomini e host donne, in base alle informazioni dedotte dai nomi.

Output della regressione

Tabella 5.1: Modelli logaritmici - Diretti

	(1) M1: ln(Prenotazioni)	(2) M2: ln(ADR)	(3) M3: OCC	(4) M4: ln(RevPAN)	(5) M5: ln(Rating)
Italiano	-0.015*** [0.005]	0.049*** [0.003]	0.021*** [0.002]	0.094*** [0.006]	0.022*** [0.001]
Uomo	0.042*** [0.003]	-0.028*** [0.002]	0.003** [0.001]	-0.030*** [0.004]	-0.011*** [0.000]
Instantbook	0.277*** [0.004]	0.050*** [0.002]	0.074*** [0.001]	0.214*** [0.004]	-0.014*** [0.000]
Response rate	0.002*** [0.000]	-0.001*** [0.000]	0.001*** [0.000]	0.003*** [0.000]	0.000*** [0.000]
Superhost	0.192*** [0.004]	-0.079*** [0.002]	0.155*** [0.001]	0.120*** [0.004]	0.034*** [0.000]
Multiproperty	0.065*** [0.003]	0.084*** [0.002]	-0.021*** [0.001]	0.034*** [0.004]	-0.013*** [0.000]
Bathrooms	-0.043*** [0.004]	0.101*** [0.003]	-0.030*** [0.001]	0.042*** [0.006]	-0.002*** [0.001]
Bedrooms	-0.108*** [0.004]	0.063*** [0.003]	-0.027*** [0.001]	0.010** [0.004]	0.002*** [0.000]
Number of photos	0.005*** [0.000]	0.004*** [0.000]	0.002*** [0.000]	0.007*** [0.000]	0.001*** [0.000]
Accommodates	0.033*** [0.001]	0.111*** [0.001]	0.001 [0.000]	0.101*** [0.002]	-0.001*** [0.000]
Luxury amenities	-0.036*** [0.004]	-0.048*** [0.002]	-0.026*** [0.001]	-0.125*** [0.004]	0.001*** [0.000]
Quality amenities	0.011*** [0.001]	0.035*** [0.001]	0.034*** [0.000]	0.096*** [0.001]	0.008*** [0.000]
Strict cancellation	-0.111*** [0.005]	0.029*** [0.003]	-0.033*** [0.002]	-0.045*** [0.006]	-0.015*** [0.001]
Shared room	-0.101*** [0.015]	-0.709*** [0.009]	-0.032*** [0.004]	-0.842*** [0.019]	-0.024*** [0.003]
Constant	0.490*** [0.012]	3.334*** [0.007]	0.026*** [0.003]	1.741*** [0.015]	4.472*** [0.002]
Year FE	YES	YES	YES	YES	YES
Month FE	YES	YES	YES	YES	YES
R ²	0.107	0.383	0.165	0.231	0.064
N	246521	246506	386735	246506	336825

Standard errors in brackets

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

Note: L'analisi è stata condotta su un campione composto da 530.960 osservazioni.

La Tabella 5.1 presenta i risultati ottenuti da cinque modelli di regressione lineare multipla (M1–M5), sviluppati per identificare i principali fattori che incidono sulla performance degli host su Airbnb. Ogni modello si concentra su un diverso indicatore di successo dell'annuncio, includendo sia metriche economiche sia reputazionali: numero di prenotazioni (ln(Prenotazioni)), tariffa media giornaliera (ln(ADR)), tasso di occupazione (OCC), ricavo medio per annuncio (ln(RevPAN)) e valutazione media (ln(Rating)).

Le variabili esplicative considerate spaziano dalle caratteristiche individuali dell'host, come genere e nazionalità, a quelle strutturali dell'alloggio, quali il numero di stanze e la capienza. Sono incluse anche variabili legate alla reputazione e alla fiducia (es. stato di Superhost, attivazione della prenotazione immediata, qualità dei servizi e delle immagini), nonché variabili di controllo che tengono conto della flessibilità dell'annuncio (es. politica di cancellazione, soggiorno minimo).

Tutti i modelli incorporano effetti fissi per anno e mese, al fine di neutralizzare l'impatto di fattori stagionali e tendenze temporali. Gli errori standard, riportati tra parentesi, sono corretti per eteroschedasticità, garantendo così maggiore robustezza alle stime. I valori di R^2 , compresi tra 0.064 e 0.383, risultano coerenti con la natura microeconomica e trasversale del campione.

Effetto della nazionalità e del genere

La variabile Italiano risulta associata a una leggera diminuzione delle prenotazioni (-0.015), ma con un effetto positivo sulla tariffa media giornaliera (+0.049), sul tasso di occupazione (+0.021) e soprattutto sui ricavi medi (+0.094). Anche il punteggio delle recensioni mostra un effetto positivo e significativo (+0.022).

Questi risultati suggeriscono che, a parità di condizioni, gli host italiani adottano una strategia più redditizia, puntando su prezzi più elevati e strutture di maggiore qualità, pur con una domanda leggermente inferiore.

Quanto al genere, la variabile Uomo mostra un pattern opposto: gli host maschili ottengono più prenotazioni (+0.042), ma applicano tariffe inferiori (-0.028), pur mantenendo un tasso di occupazione positivo (+0.003). Tuttavia, il loro impatto sui ricavi complessivi risulta negativo (-0.030), e anche sul rating medio si osserva un effetto sfavorevole (-0.011).

Ciò potrebbe riflettere differenze gestionali o di percezione del servizio, dove gli host donne riescono a mantenere una reputazione leggermente migliore.

Effetti delle caratteristiche dell'annuncio

L'opzione Instant Book incide positivamente su tutte le principali metriche economiche: aumenta le prenotazioni (+0.277), la tariffa media (+0.050), l'occupazione (+0.074) e i ricavi medi (+0.214). Tuttavia, si rileva un leggero effetto negativo sul rating (-0.014), probabilmente legato alla minore selezione degli ospiti e a esperienze meno controllate.

Lo status di Superhost rappresenta un segnale reputazionale particolarmente forte. Esso mostra un impatto positivo e altamente significativo su prenotazioni (+0.192), occupazione (+0.155), ricavi (+0.120) e valutazioni (+0.034), pur risultando associato a tariffe leggermente inferiori (-0.079). Ciò suggerisce che gli host Superhost

puntano più sulla fiducia e la soddisfazione del cliente che sul prezzo unitario, ottenendo comunque migliori performance complessive.

Gli host Multiproperty presentano risultati contrastanti: effetti positivi su prenotazioni (+0.065), tariffe (+0.084) e ricavi (+0.034), ma negativi su occupazione (-0.021) e rating (-0.013). Questo pattern suggerisce un posizionamento commerciale più strutturato ma meno personalizzato, che può ridurre la qualità percepita dell'esperienza.

Le caratteristiche fisiche dell'alloggio mostrano impatti eterogenei.

Il numero di bagni aumenta significativamente la tariffa media (+0.101) e i ricavi (+0.042), ma riduce le prenotazioni (-0.043) e l'occupazione (-0.030), segnalando un posizionamento verso segmenti premium ma con domanda più ristretta.

Un andamento simile si osserva per le camere da letto, che incidono positivamente su ADR (+0.063) e rating (+0.002), ma negativamente su prenotazioni (-0.108) e occupazione (-0.027).

La capacità ricettiva (Accommodates) mostra effetti positivi e significativi su ADR (+0.111) e RevPAN (+0.101), anche se leggermente negativi sul rating (-0.001), confermando che gli alloggi più capienti attraggono una domanda più ampia e generano maggiori ricavi medi.

Il numero di fotografie pubblicate è positivamente associato a tutti gli indicatori di performance: prenotazioni (+0.005), ADR (+0.004), occupazione (+0.002), ricavi (+0.007) e rating (+0.001). Ciò conferma l'importanza della componente visiva come segnale di qualità e trasparenza.

Le quality amenities risultano anch'esse positivamente correlate con le performance: +0.035 su ADR, +0.034 su OCC e +0.096 su RevPAN a testimonianza del peso dei comfort essenziali nella soddisfazione del cliente.

Al contrario, le luxury amenities mostrano effetti negativi su tutte le metriche (-0.048 su ADR, -0.026 su OCC e -0.125 su RevPAN), indicando che il lusso non rappresenta necessariamente un vantaggio competitivo rispetto alla funzionalità e alla praticità dell'offerta.

La politica di cancellazione strict produce effetti misti: è associata a tariffe più alte (+0.029) ma a un calo di prenotazioni (-0.111), occupazione (-0.033), ricavi (-0.045) e valutazioni (-0.015), suggerendo che la rigidità nelle condizioni riduce la flessibilità percepita e la propensione alla prenotazione.

Infine, la formula shared room mostra impatti fortemente negativi su tutte le dimensioni: prenotazioni (-0.101), ADR (-0.709), occupazione (-0.032), ricavi (-0.842) e rating (-0.024). Ciò conferma la marginalità economica di questa tipologia di offerta e la sua scarsa attrattività per la maggior parte dei viaggiatori.

Sintesi interpretativa

Nel complesso, l'analisi mette in luce come il successo degli host su Airbnb sia influenzato da una combinazione articolata di fattori reputazionali, strutturali e individuali.

Tra questi, i segnali di fiducia, come il badge di Superhost, la prenotazione immediata e la qualità delle immagini, risultano determinanti per la performance economica, coerentemente con le dinamiche della reputation economy che caratterizzano le piattaforme digitali.

Le caratteristiche fisiche dell'alloggio contribuiscono a definire il posizionamento sul mercato, mentre le variabili demografiche dell'host riflettono potenziali differenze gestionali e di percezione da parte dei consumatori.

In sintesi, la performance su Airbnb deriva da un equilibrio tra strategia commerciale, qualità percepita e reputazione, elementi che insieme definiscono il valore complessivo dell'offerta in un ecosistema digitale basato sulla fiducia e sulla trasparenza.

5.1.2 Regressione multivariata logaritmica con interazione

In questa sezione viene presentata la regressione individuata per lo studio delle differenze di performance tra host italiani e non italiani e host uomini e host donne e di come questi aspetti interagiscano tra loro.

Output della regressione

Tabella 5.2: Modelli logaritmici - Diretti e Interaction

	(1) M1: ln(Prenotazioni)	(2) M2: ln(ADR)	(3) M3: OCC	(4) M4: ln(RevPAN)	(5) M5: ln(Rating)
Italiano	-0.020*** [0.007]	0.045*** [0.004]	0.013*** [0.002]	0.094*** [0.008]	0.018*** [0.001]
Uomo	0.034*** [0.010]	-0.036*** [0.006]	-0.012*** [0.003]	-0.029*** [0.011]	-0.017*** [0.002]
Uomo × Italiano	0.009 [0.010]	0.009 [0.006]	0.017*** [0.003]	-0.000 [0.012]	0.007*** [0.002]
Instantbook	0.277*** [0.004]	0.050*** [0.002]	0.074*** [0.001]	0.214*** [0.004]	-0.014*** [0.000]
Response rate	0.002*** [0.000]	-0.001*** [0.000]	0.001*** [0.000]	0.003*** [0.000]	0.000*** [0.000]
Superhost	0.192*** [0.004]	-0.079*** [0.002]	0.155*** [0.001]	0.120*** [0.004]	0.034*** [0.000]
Multiproperty	0.065*** [0.003]	0.084*** [0.002]	-0.021*** [0.001]	0.034*** [0.004]	-0.013*** [0.000]
Bathrooms	-0.043*** [0.004]	0.101*** [0.003]	-0.030*** [0.001]	0.042*** [0.006]	-0.002*** [0.001]
Bedrooms	-0.108*** [0.004]	0.063*** [0.003]	-0.027*** [0.001]	0.010** [0.004]	0.002*** [0.000]
Number of photos	0.005*** [0.000]	0.004*** [0.000]	0.002*** [0.000]	0.007*** [0.000]	0.001*** [0.000]
Accommodates	0.033*** [0.001]	0.111*** [0.001]	0.001 [0.000]	0.101*** [0.002]	-0.001*** [0.000]
Luxury amenities	-0.036*** [0.004]	-0.048*** [0.002]	-0.026*** [0.001]	-0.125*** [0.004]	0.001*** [0.000]
Quality amenities	0.011*** [0.001]	0.035*** [0.001]	0.034*** [0.000]	0.096*** [0.001]	0.008*** [0.000]
Strict cancellation	-0.111*** [0.005]	0.029*** [0.003]	-0.033*** [0.002]	-0.045*** [0.006]	-0.015*** [0.001]
Shared room	-0.101*** [0.015]	-0.709*** [0.009]	-0.031*** [0.004]	-0.842*** [0.019]	-0.023*** [0.003]
Constant	0.494*** [0.012]	3.338*** [0.008]	0.034*** [0.004]	1.740*** [0.015]	4.475*** [0.002]
Year FE	YES	YES	YES	YES	YES
Month FE	YES	YES	YES	YES	YES
R ²	0.107	0.383	0.165	0.231	0.065
N	246521	246506	386735	246506	336825

Standard errors in brackets

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

Note: L'analisi è stata condotta su un campione composto da 530.960 osservazioni.

La Tabella 5.2 mostra i risultati di cinque modelli di regressione lineare logaritmica, stimati tenendo conto degli effetti fissi per anno e mese. Si tratta di un'analisi che prova a capire quali fattori influenzano davvero le performance degli host su Airbnb, eliminando l'impatto delle variazioni legate al tempo.

Le variabili dipendenti prese in esame sono cinque come nei modelli precedenti: il

numero di prenotazioni, il prezzo medio giornaliero (ADR), il tasso di occupazione (OCC), il ricavo per annuncio (RevPAN) e la valutazione media (Rating). Ogni modello include errori standard robusti, indicati tra parentesi, che servono a rendere più affidabili le stime anche quando i dati non sono perfettamente omogenei. Inoltre si tiene conto dei cambiamenti nel tempo che potrebbero incidere sui risultati finali.

Effetto della nazionalità e del genere

Il coefficiente riferito alla variabile “Italiano” risulta positivo e statisticamente significativo in quasi tutti i modelli tranne nel caso delle prenotazioni. In particolare, è negativo per le prenotazioni (-0.020), ma positivo per l’ADR (+0.045), per il tasso di occupazione (+0.01), per il RevPAN (+0.094) e per il rating medio (+0.018). Questo significa quindi che a parità di altre condizioni gli host italiani tendono a fissare prezzi leggermente più alti e ad ottenere ricavi e valutazioni migliori, anche se registrano meno prenotazioni rispetto agli host stranieri. Gli host italiani sembrano puntare più sulla qualità e sul valore, piuttosto che sulla quantità. La variabile “Uomo”, invece, mostra un effetto generalmente negativo e molto significativo su quasi tutte le metriche di performance. Gli host uomini hanno più prenotazioni (+0.034), ma allo stesso tempo prezzi medi più bassi (-0.036 sull’ADR), minore occupazione (-0.012), ricavi inferiori (-0.029) e valutazioni peggiori (-0.017). Quindi gli host uomini attirano più prenotazioni ma i loro risultati economici e di reputazione sono meno favorevoli. Questo potrebbe suggerire un leggero svantaggio competitivo legato al genere, forse per differenze nello stile di gestione o nell’interazione con gli ospiti.

L’interazione tra Uomo × Italiano non è fortemente significativa in tutti i modelli, ma mostra effetti positivi moderati, specialmente sull’occupazione (+0.017, $p<0.01$) e sul rating (+0.007, $p<0.01$). Ciò indica che essere un host italiano uomo può attenuare parzialmente gli svantaggi associati al genere, migliorando leggermente la visibilità e la soddisfazione dei clienti, anche se senza influire in modo sostanziale sui ricavi.

Caratteristiche dell’annuncio

Tra le variabili di controllo, la funzione Instantbook mostra un impatto chiaramente positivo e statisticamente significativo nella maggior parte dei modelli. In particolare, aumenta le prenotazioni di +0.277, l’ADR di +0.050, il tasso di occupazione di +0.074 e il RevPAN di +0.214 (tutti con $p<0.01$). L’unico effetto negativo si osserva sul rating medio che scende di poco di -0.014. Questo risultato suggerisce che la possibilità di prenotare subito, senza attendere la conferma dell’host, rende l’annuncio più attraente per gli utenti e favorisce sia la domanda che i ricavi complessivi. Allo stesso tempo, però, sembra esserci un piccolo effetto collaterale: una leggera diminuzione della soddisfazione media. Probabilmente ciò accade perché,

con l'Instantbook attivo, gli host hanno meno controllo sulla selezione degli ospiti. Il response rate ha un effetto positivo e significativo su tutte le metriche, seppur di piccola entità (tra lo +0.001 e il +0.003), evidenziando che una maggiore reattività dell'host migliora le performance e le valutazioni.

Lo status di Superhost risulta altrettanto importante: gli host certificati ottengono +0.192 prenotazioni, +0.155 di occupazione, +0.120 di RevPAN e +0.034 di rating (tutti $p < 0.01$), ma tendono a praticare prezzi leggermente più bassi (-0.079 sull'ADR). In altre parole, essere Superhost migliora la fiducia e la domanda, ma spinge a una politica di prezzo più competitiva.

Le caratteristiche strutturali mostrano effetti misti ma coerenti. Il numero di bagni mostra un'associazione negativa con le prenotazioni (-0.043), ma allo stesso tempo aumenta il prezzo medio giornaliero (ADR) di +0.101 e il ricavo per annuncio (RevPAN) di +0.042 (entrambi $p < 0.01$). In pratica, gli alloggi più ampi e confortevoli, dotati di più bagni, tendono a costare di più, ma vengono prenotati meno frequentemente.

Una dinamica simile si osserva anche per il numero di camere da letto. Qui, infatti, l'effetto sulle prenotazioni è leggermente negativo (-0.108), mentre l'ADR cresce di +0.063. Questo significa che più stanze equivalgono a prezzi medi più alti, ma a un minore volume di prenotazioni, probabilmente perché gli annunci diventano meno accessibili per viaggiatori singoli o coppie. Nel complesso, questi risultati mostrano un trade-off tra quantità e valore: più comfort e spazio portano a tariffe superiori e guadagni unitari maggiori, ma anche a una domanda leggermente più limitata. Gli annunci che ospitano più persone (Accommodates) ottengono risultati positivi e significativi in tutti i modelli (es. +0.033 sulle prenotazioni), indicando che la capacità di ospitare più persone resta un forte driver di performance. Tra i servizi extra, gli amenities di lusso mostrano un effetto negativo e significativo su quasi tutte le principali misure di performance. Gli impatti più evidenti si notano sull'ADR (-0.048) e sul RevPAN (-0.125). In modo particolare quindi aggiungere dotazioni di lusso non porta sempre in maniera automatica guadagni più alti. Spesso infatti questi servizi aumentano i costi di gestione oppure non vengono realmente apprezzati dagli ospiti, che tendono a scegliere l'alloggio in base ad altri aspetti più pratici. Al contrario, gli amenities di qualità (quelli legati alla comodità e alla funzionalità) producono effetti positivi e significativi: +0.035 sull'ADR e +0.096 sul RevPAN (entrambi $p < 0.01$). Ciò significa che gli ospiti danno più valore a comfort ben curati, spazi accoglienti e attenzioni utili, piuttosto che a elementi di lusso fine a sé stessi.

Gli host con più proprietà (Multiproperty) mostrano effetti contrastanti: +0.065 sulle prenotazioni e +0.084 sull'ADR, ma -0.021 sull'occupazione e -0.013 sul rating, tutti significativi. Ciò implica che i proprietari multipli riescono a mantenere prezzi più alti e un volume di prenotazioni maggiore, ma tendono ad avere una minore

occupazione media e valutazioni leggermente inferiori, forse per minore attenzione individuale.

Le stanze condivise (Shared room) risultano fortemente penalizzate: -0.709 sull'ADR, -0.842 sul RevPAN, e valori negativi anche su tutte le altre metriche, confermando che la domanda privilegia l'alloggio intero rispetto alle soluzioni condivise.

Infine, la politica di cancellazione restrittiva (Strict cancellation) mostra un effetto negativo su prenotazioni (-0.111), occupazione (-0.033), ricavi (-0.045) e anche sul rating (-0.015), mentre risulta positivo sull'ADR (+0.029).

Questo indica che una maggiore severità nella gestione delle prenotazioni tende a ridurre il volume di domanda e la soddisfazione media degli ospiti, ma consente di applicare prezzi più elevati, probabilmente per via della percezione di maggiore serietà o selettività dell'host.

Capacità esplicativa

I valori di R^2 variano tra 0.065 e 0.383, indicando una discreta capacità esplicativa dei modelli, con la migliore aderenza ai dati osservata per il modello ADR (M2), in cui le caratteristiche strutturali e reputazionali spiegano una parte maggiore della variabilità dei risultati.

Tutti i modelli includono effetti fissi per anno e mese (Year FE e Month FE) per controllare le variazioni stagionali e temporali del mercato turistico.

Sintesi interpretativa

Nel complesso, i risultati confermano che le performance economiche degli host su Airbnb dipendono in modo significativo sia da caratteristiche individuali (genere, nazionalità) sia da fattori legati all'annuncio (dimensione, servizi, reputazione e modalità di prenotazione).

L'evidenza empirica mostra un divario di genere a sfavore degli uomini in termini di prezzo, ricavi e valutazioni, ma con effetti mitigati per gli host italiani, nei quali il legame tra nazionalità e performance appare più articolato e influenzato da fattori culturali o reputazionali.

Inoltre, variabili operative come Instant book, Superhost e response rate emergono come leve chiave per incrementare la visibilità e la performance economica, mentre la gestione multipla e le stanze condivise risultano associate a una percezione di minor qualità e, di conseguenza, a performance inferiori.

Comando Stata

A titolo di esempio per mostrare la modalità utilizzate per ottenere le regressioni è stato scelto di presentare il comando Stata sviluppato per le regressioni di

interazione. Il comando, mostrato in seguito nel dettaglio, stima 5 modelli di regressione lineare (OLS) e li sintetizza nella tabella comparativa vista in precedenza. Attraverso il comando reg, vengono eseguite cinque regressioni le diverse variabili dipendenti (Prenotazioni, ADR, OCC, RevPAN e Rating), includendo controlli individuali e fixed effects di anno e mese (i.year i.month). Gli errori standard robusti (vce(robust)) correggono l'eteroschedasticità, mentre eststo e estadd servono a salvare e annotare ciascun modello. Infine, esttab combina i risultati in una tabella formattata, mostrando coefficienti, errori standard tra parentesi, livelli di significatività e statistiche riassuntive (R^2 , N, Year FE, Month FE).

```

1 capture drop uomo_italiano
2 gen byte uomo_italiano = uomo*italiano
3 label var uomo_italiano "Uomo x Italiano"
4
5 eststo clear
6
7 eststo: reg ln_reservations italiano uomo uomo_italiano
    ↪ instantbook response_rate airbnb_superhost multiproperty
    ↪ bathrooms bedrooms number_of_photos accommodates
    ↪ luxury_amenities quality_amenities strict_cancellation
    ↪ shared_room i.year i.month, vce(robust)
8 estadd local YearFE "YES"
9 estadd local MonthFE "YES"
10
11 eststo: reg ln_adr italiano uomo uomo_italiano instantbook
    ↪ response_rate airbnb_superhost multiproperty bathrooms
    ↪ bedrooms number_of_photos accommodates luxury_amenities
    ↪ quality_amenities strict_cancellation shared_room i.year
    ↪ i.month, vce(robust)
12 estadd local YearFE "YES"
13 estadd local MonthFE "YES"
14
15 eststo: reg occ italiano uomo uomo_italiano instantbook
    ↪ response_rate airbnb_superhost multiproperty bathrooms
    ↪ bedrooms number_of_photos accommodates luxury_amenities
    ↪ quality_amenities strict_cancellation shared_room i.year
    ↪ i.month, vce(robust)
16 estadd local YearFE "YES"
17 estadd local MonthFE "YES"
18
19 eststo: reg ln_revpan italiano uomo uomo_italiano instantbook
    ↪ response_rate airbnb_superhost multiproperty bathrooms
    ↪ bedrooms number_of_photos accommodates luxury_amenities
    ↪ quality_amenities strict_cancellation shared_room i.year
    ↪ i.month, vce(robust)
20 estadd local YearFE "YES"
21 estadd local MonthFE "YES"
22
23 eststo: reg ln_rating italiano uomo uomo_italiano instantbook
    ↪ response_rate airbnb_superhost multiproperty bathrooms
    ↪ bedrooms number_of_photos accommodates luxury_amenities
    ↪ quality_amenities strict_cancellation shared_room i.year
    ↪ i.month, vce(robust)
24 estadd local YearFE "YES"
25 estadd local MonthFE "YES"
26
27 esttab, b(%9.3f) se(%9.3f) brackets star(* 0.10 ** 0.05 *** 0.01)
    ↪ label compress nogaps

```

```

28 keep(italiano uomo uomo_italiano instantbook response_rate
      ↪ airbnb_superhost multiproperty bathrooms bedrooms
      ↪ number_of_photos accommodates luxury_amenities
      ↪ quality_amenities strict_cancellation shared_room _cons)
29 order(italiano uomo uomo_italiano instantbook response_rate
      ↪ airbnb_superhost multiproperty bathrooms bedrooms
      ↪ number_of_photos accommodates luxury_amenities
      ↪ quality_amenities strict_cancellation shared_room _cons)
30 coeflabels(italiano "Italiano" uomo "Uomo" uomo_italiano "Uomo"
      ↪ Italiano")
31 mtitles("M1: ln(Prenotazioni)" "M2: ln(ADR)" "M3: OCC" "M4:
      ↪ ln(RevPAN)" "M5: ln(Rating)")
32 stats(YearFE MonthFE r2 N, fmt(%9s %9s 3 0) labels("Year FE"
      ↪ "Month FE" "R^2" "N"))

```

Listing 5.1: Comandi Stata per la stima dei modelli M1–M5 con Italiano, Uomo e interazione Uomo × Italiano

5.1.3 Interaction plot

Di seguito sono presentati gli interaction plot, grafici che rappresentano le interazioni tra variabili e consentono di visualizzare in che modo la relazione tra una variabile indipendente e quella dipendente cambia in funzione dei valori assunti da una variabile moderatrice.

In particolare, si esamina l'effetto di una variabile moderatrice chiave: la nazionalità dell'host (Italiano vs Straniero). Le seguenti Figure forniscono una rappresentazione grafica degli effetti moderatori significativi riscontrati in questo studio, analizzando come la nazionalità dell'host influenzi diverse metriche di performance (Rating, RevPAN, Occupancy rate, Numero di Prenotazioni e ADR) in modo differenziato a seconda del genere.

Prenotazioni

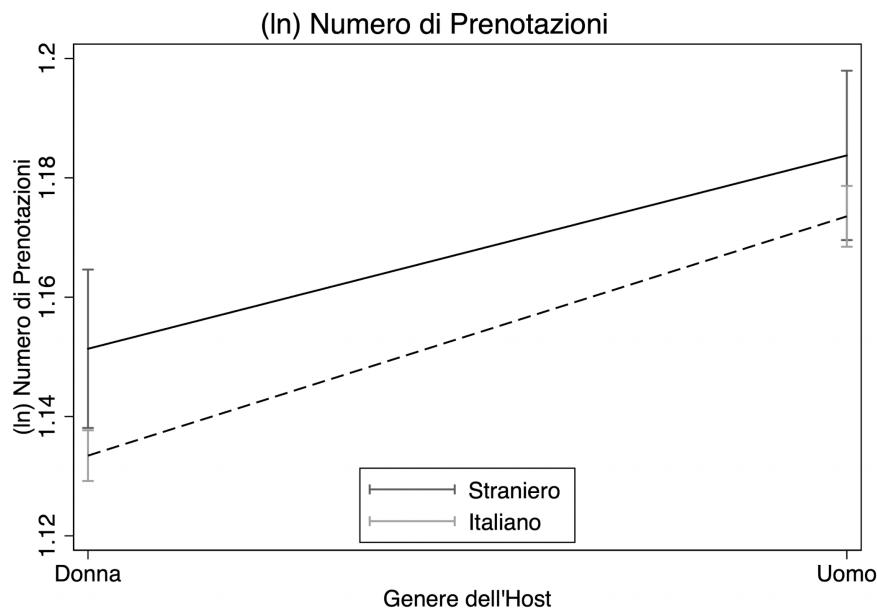


Figura 5.1: Diagramma di interazione per l'effetto moderatore della nazionalità nella relazione tra Genere dell'host e Numero di prenotazioni

L'interaction plot relativo al numero di prenotazioni presenta un pattern interessante, differenziandosi dalle altre metriche analizzate. Entrambe le linee mostrano una pendenza positiva, indicando che gli host maschi ricevono più prenotazioni rispetto alle host donne, indipendentemente dalla nazionalità. Gli host stranieri (linea continua) partono da circa 1.15 prenotazioni per le donne e raggiungono circa 1.18 per gli uomini, mostrando un incremento moderato. Gli host italiani (linea tratteggiata) seguono un pattern simile ma con valori assoluti inferiori: partono da circa 1.13 prenotazioni per le donne e arrivano a circa 1.17 per gli uomini.

Le due linee sono sostanzialmente parallele, indicando che il gender gap nel numero di prenotazioni è simile per italiani e stranieri. In questo caso, contrariamente alle altre metriche, essere uomo rappresenta un vantaggio in termini di volume di prenotazioni ricevute. Questo risultato è particolarmente significativo perché suggerisce che, sebbene le donne possano ottenere rating e RevPAN leggermente superiori, gli uomini riescono ad attrarre un maggior numero di prenotazioni. Ciò potrebbe indicare l'esistenza di bias impliciti dei guest nella scelta iniziale dell'alloggio, che favorirebbero gli host maschi.

Allo stesso modo, contrariamente da quanto ci si sarebbe aspettato, gli host stranieri registrano un numero maggiore di prenotazioni, indipendentemente dal genere.

ADR (Average Daily Rate)

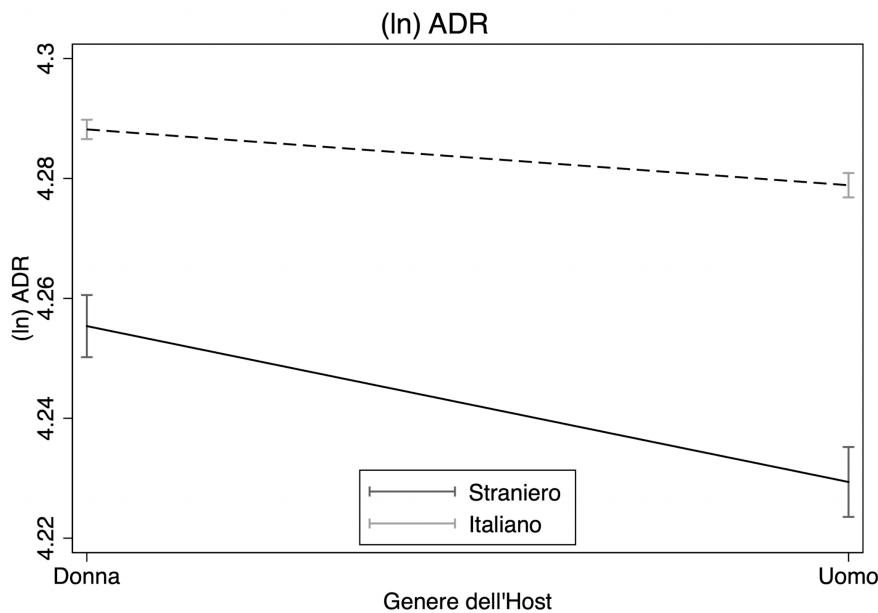


Figura 5.2: Diagramma di interazione per l'effetto moderatore della nazionalità nella relazione tra Genere dell'host e ADR

L'ADR mostra pattern che riflettono parzialmente quanto osservato per il RevPAN, con alcune specificità rilevanti:

- Gli host italiani (linea tratteggiata) mantengono un ADR costantemente superiore, partendo da circa 4.29 per le donne e scendendo leggermente a circa 4.28 per gli uomini. La pendenza negativa è molto contenuta, indicando un gender gap minimo.
- Gli host stranieri (linea continua) registrano valori inferiori: le donne partono da circa 4.26 mentre gli uomini si attestano intorno a 4.23. La pendenza negativa è più accentuata rispetto agli italiani, suggerendo un gender gap più marcato nell'ADR per gli host stranieri.
- Le due linee, pur rimanendo distinte, tendono leggermente a convergere verso il lato sinistro del grafico (donna), indicando che il vantaggio degli italiani sull'ADR si riduce lievemente quando si considerano gli host femminili.
- Questo pattern suggerisce che le host donne, sia italiane che straniere, riescono a praticare prezzi medi giornalieri leggermente superiori rispetto ai

colleghi maschi. Tuttavia, gli host stranieri mantengono un prezzo più basso indipendentemente dal genere.

OCC (Occupancy Rate)

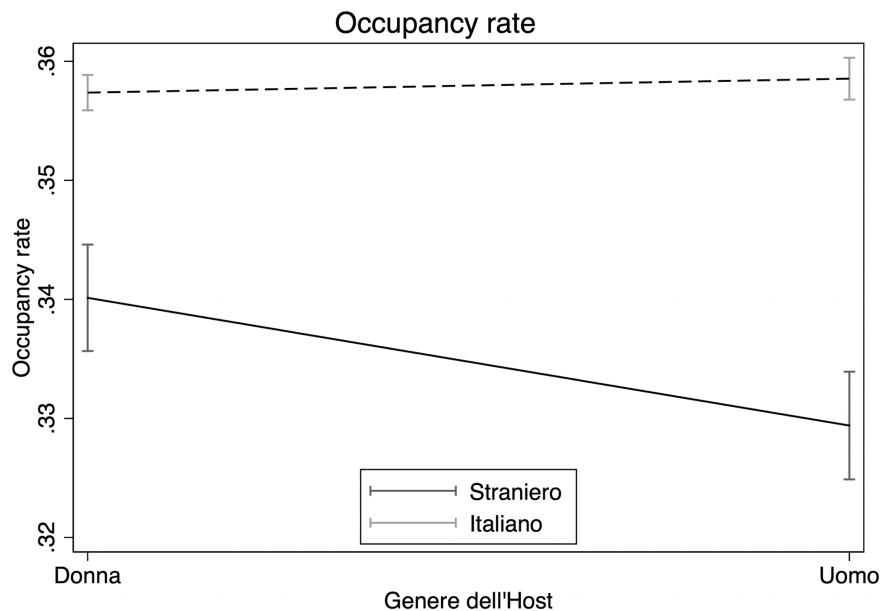


Figura 5.3: Diagramma di interazione per l'effetto moderatore della nazionalità nella relazione tra Genere dell'host e OCC

Gli host italiani (linea tratteggiata) mantengono un tasso di occupazione quasi costante e superiore, attestandosi intorno a 35.8-35.9% sia per le donne che per gli uomini. La linea è quasi piatta, indicando l'assenza di un gender gap significativo per questa categoria. Gli host stranieri (linea continua) mostrano invece un tasso di occupazione inferiore che parte da circa 34.0% per le donne e scende a circa 33.0% per gli uomini. La pendenza negativa indica la presenza di un gender gap più marcato rispetto agli italiani.

Le due linee sono molto distanziate, con gli host italiani che mantengono un vantaggio competitivo di circa 2 punti percentuali sull'occupancy rate. Questo pattern suggerisce che la nazionalità dell'host può avere un effetto moderatore sul gender gap nell'occupancy rate: mentre per gli italiani il genere sembra irrilevante, per gli stranieri essere donna rappresenta un vantaggio in termini di tasso di occupazione, seppur partendo da livelli mediamente inferiori rispetto agli italiani.

RevPAN

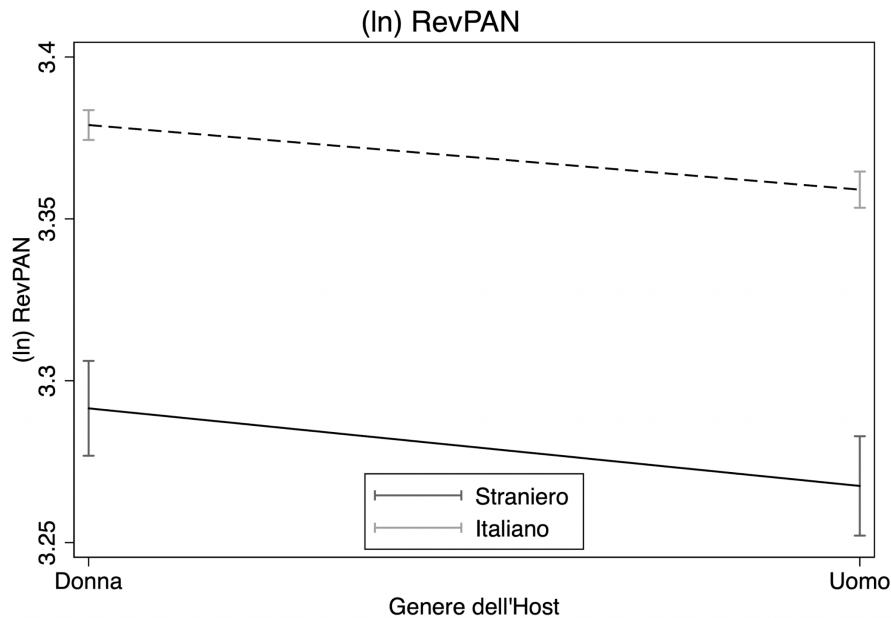


Figura 5.4: Diagramma di interazione per l'effetto moderatore della nazionalità nella relazione tra Genere dell'host e RevPAN

L'analisi del RevPAN (Revenue per Available Night) rivela pattern simili a quelli osservati per il rating, con alcune specificità interessanti. Gli host italiani (linea tratteggiata) mostrano valori di RevPAN sistematicamente superiori rispetto agli stranieri, partendo da circa 3.37 per le donne e attestandosi a circa 3.35 per gli uomini. Il gender gap, seppur presente, risulta modesto. Gli host stranieri (linea continua) registrano valori inferiori: le donne partono da circa 3.30 mentre gli uomini si attestano intorno a 3.26, evidenziando anche qui un gender gap a sfavore degli uomini, sebbene di entità limitata.

Entrambe le linee presentano una pendenza negativa, indicando che gli host maschi tendono a ottenere un RevPAN inferiore rispetto alle host donne, indipendentemente dalla nazionalità. Questo suggerisce che le donne riescono a generare ricavi per notte disponibile leggermente superiori. La distanza tra le due linee rimane pressoché costante lungo l'asse del genere, confermando che la nazionalità dell'host agisce come fattore che influenza il livello generale di performance, ma non altera significativamente il gender gap nel RevPAN.

Rating

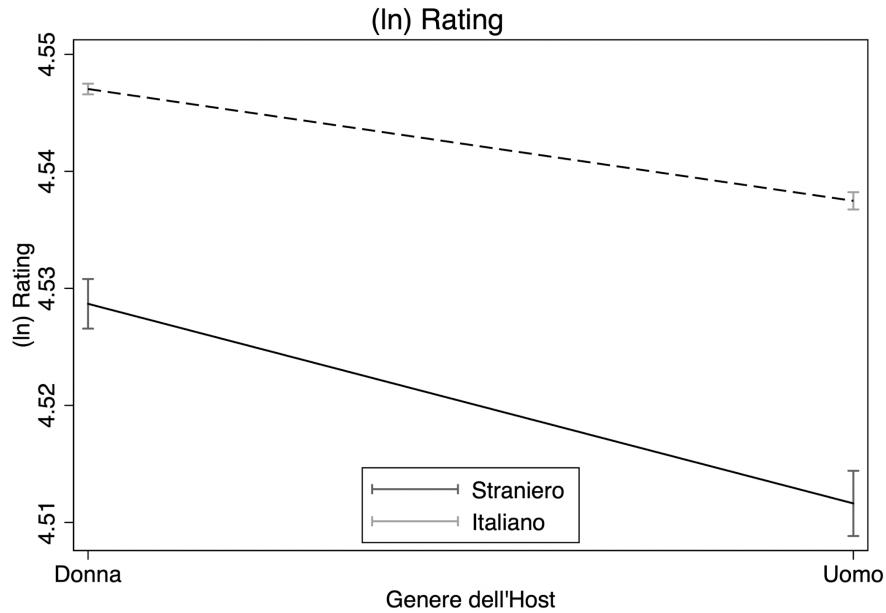


Figura 5.5: Diagramma di interazione per l'effetto moderatore della nazionalità nella relazione tra Genere dell'host e Rating

Il gender gap nel rating degli alloggi mostra un pattern interessante quando si considera la nazionalità dell'host. Come illustrato nella Figura, per gli host italiani (linea tratteggiata) le donne partono da un rating medio superiore a 4.54, mentre gli uomini registrano un rating pari a 4.54. Il divario è minimo, indicando una sostanziale parità nelle valutazioni ricevute da host di sesso diverso tra gli italiani. Per gli host stranieri (linea continua) emerge un pattern differente: le donne hanno un rating iniziale di circa 4.53, mentre gli uomini registrano un rating di circa 4.51, sebbene anche in questo caso il gender gap risulti contenuto.

La caratteristica più rilevante è che entrambe le linee mostrano una pendenza negativa, con il rating che diminuisce passando dalle donne agli uomini. Tuttavia, il divario rimane relativamente stabile per entrambe le nazionalità, suggerendo che la nazionalità dell'host non modifica sostanzialmente l'effetto del genere sul rating ricevuto. Gli host stranieri ottengono in generale rating leggermente inferiori rispetto agli italiani, indipendentemente dal genere, ma il gender gap rimane comparabile tra le due categorie.

5.2 Analisi foto profilo degli host

5.2.1 Regressione multivariata logaritmica

In questa sezione viene presentata la regressione sviluppata per indagare la correlazione tra le caratteristiche visive (*visual cues*) della foto profilo degli host con gli indicatori di performance individuati.

Output della regressione

Tabella 5.3: Modelli logaritmici - Foto del profilo host

	(1) M1: ln(Prenotazioni)	(2) M2: ln(ADR)	(3) M3: OCC	(4) M4: ln(RevPAN)	(5) M5: ln(Rating)
Face	-0.071*** [0.020]	0.012 [0.015]	-0.010 [0.009]	0.050* [0.029]	0.019*** [0.002]
Smile	0.100*** [0.017]	-0.050*** [0.011]	0.057*** [0.008]	0.098*** [0.022]	0.002 [0.002]
Sunglasses	-0.028 [0.024]	0.058*** [0.016]	-0.024** [0.010]	0.036 [0.031]	0.016*** [0.002]
Group	-0.221*** [0.024]	0.068*** [0.017]	-0.074*** [0.011]	-0.098*** [0.033]	0.001 [0.003]
Children	0.273*** [0.041]	-0.041 [0.026]	0.121*** [0.019]	0.260*** [0.055]	0.016*** [0.004]
Animal	0.016 [0.044]	-0.076* [0.040]	0.002 [0.021]	-0.132* [0.070]	-0.023*** [0.007]
Instantbook	0.173*** [0.019]	0.035*** [0.012]	0.074*** [0.008]	0.289*** [0.023]	-0.004* [0.002]
Response rate	0.004*** [0.000]	-0.000 [0.000]	0.002*** [0.000]	0.004*** [0.001]	0.000*** [0.000]
Superhost	0.402*** [0.021]	0.055*** [0.012]	0.177*** [0.009]	0.301*** [0.023]	0.040*** [0.001]
Multiproperty host	-0.089*** [0.015]	0.060*** [0.010]	-0.039*** [0.007]	-0.031 [0.019]	-0.024*** [0.001]
Bathrooms	-0.187*** [0.018]	0.189*** [0.016]	-0.051*** [0.008]	0.125*** [0.030]	0.021*** [0.002]
Bedrooms	-0.160*** [0.016]	0.069*** [0.011]	-0.056*** [0.007]	-0.009 [0.022]	0.011*** [0.002]
Number of photos	0.011*** [0.001]	0.002*** [0.000]	0.003*** [0.000]	0.003*** [0.001]	-0.000*** [0.000]
Accommodates	0.077*** [0.006]	0.079*** [0.005]	0.019*** [0.003]	0.107*** [0.009]	-0.003*** [0.001]
Luxury amenities	0.025 [0.021]	0.065*** [0.014]	0.008 [0.009]	0.079*** [0.026]	0.012*** [0.002]
Quality amenities	0.103*** [0.033]	0.311*** [0.020]	0.031** [0.014]	0.304*** [0.053]	0.013*** [0.005]
Strict cancellation	0.012 [0.021]	-0.008 [0.016]	0.020** [0.010]	0.045 [0.034]	-0.027*** [0.003]
Shared room	-0.345*** [0.046]	-0.594*** [0.041]	-0.169*** [0.019]	-0.977*** [0.120]	-0.080*** [0.014]
Constant	0.192*** [0.049]	3.192*** [0.044]	0.026 [0.021]	1.540*** [0.093]	4.471*** [0.007]
Year FE	YES	YES	YES	YES	YES
Month FE	YES	YES	YES	YES	YES
R ²	0.212	0.404	0.173	0.276	0.195
N	12098	7844	12098	7844	10344

Standard errors in brackets

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

Note: L'analisi è stata condotta su un campione composto da 12.380 osservazioni.

La Tabella 5.3 riporta i risultati di cinque modelli di regressione logaritmica stimati con effetti fissi per anno e mese, che analizzano l'impatto dei segnali visivi presenti nelle foto profilo degli host sulle principali metriche di performance su Airbnb. Le variabili dipendenti rappresentano differenti dimensioni del successo economico e reputazionale: il numero di prenotazioni (M1), il prezzo medio giornaliero o ADR (M2), il tasso di occupazione (M3), il ricavo per annuncio o RevPAN (M4) e il rating medio ricevuto (M5). Tutti i modelli includono errori standard robusti (riportati tra parentesi) e controllano per caratteristiche strutturali, reputazionali e di servizio dell'annuncio, isolando così l'effetto specifico dei tratti visivi associati al volto e alla presentazione personale dell'host.

Effetto dei segnali visivi e del volto

La presenza del volto dell'host nella foto profilo (face) mostra un effetto negativo e statisticamente significativo sulle prenotazioni (-0.071), suggerendo che gli host che mostrano il proprio volto ottengono, a parità di altre condizioni, un numero inferiore di prenotazioni. Tuttavia, il segno positivo sul ricavo medio per annuncio (+0.050) indica che questi host possono compensare parzialmente tale svantaggio praticando prezzi più elevati o attirando ospiti disposti a pagare di più. In altri termini, la visibilità del volto può ridurre la quantità di transazioni ma migliorarne la qualità economica, in linea con l'idea che una maggiore riconoscibilità personale accresca la fiducia e la percezione di autenticità.

Tra tutti i segnali emotivi, il sorriso è quello che rappresenta la fiducia e la simpatia. Dai dati gli host che sorridono ricevono più prenotazioni (+0.100) quindi chi mostra un volto aperto e accogliente sembra riuscire a conquistare più facilmente chi cerca un alloggio. Però al tempo stesso il sorriso è legato anche ad una leggera diminuzione delle tariffe (-0.050) fa pensare che gli host puntano verso prezzi un po' più bassi per ampliare la propria clientela. Alla fine comunque il RevPAN cresce (+0.098), mostrando come l'espressione amichevole porta comunque più guadagni complessivi. Al contrario, la presenza di occhiali da sole (sunglasses) è associata a: meno fiducia, prezzi più alti (+0.058), forse perché nascondere lo sguardo fa sembrare l'host meno trasparente. Anche le foto di gruppo mostrano effetti negativi: -0.221 sulle prenotazioni, -0.074 sull'occupazione e -0.098 sul RevPAN. Quando in foto ci sono più persone, diventa difficile capire chi sia davvero l'host e questa piccola incertezza riduce la propensione a prenotare.

Quando in una foto compaiono dei bambini, l'effetto è quasi sempre positivo. Dalle analisi in termini numerici deduciamo che: +0.273 sulle prenotazioni, +0.121 sull'occupazione e +0.260 sui ricavi. Quindi mostrare un contesto familiare o un momento di vita quotidiana provoca in chi guarda: empatia o fiducia, una sensazione che spinge più persone a prenotare. Per quanto riguarda gli animali domestici invece i dati mostrano un effetto negativo sull'ADR (-0.076): chi include

un cane o un gatto nella foto tende a praticare prezzi leggermente più alti forse perché la scena appare più autentica o personale. Dall'altro lato, però, emergono alcuni fattori negativi: RevPAN in diminuzione (-0.132) e un rating un po' più basso (-0.023). Si capisce quindi che un animale può far sembrare l'ambiente più accogliente, ma anche sollevare qualche perplessità magari legata all'igiene o alla cura degli spazi.

Effetti delle caratteristiche dell'annuncio

Tra le variabili di controllo, la possibilità di prenotazione immediata (Instantbook) mostra un impatto fortemente positivo in tutti i modelli economici: +0.173 sulle prenotazioni e +0.289 sui ricavi medi. Ciò conferma che la riduzione delle frizioni nel processo di prenotazione migliora l'efficienza e l'attrattività complessiva dell'annuncio.

Anche lo status du Superhost risulta significativamente associato a performance superiori in tutte le metriche (+0.402 sulle prenotazioni, +0.301 sui ricavi), rafforzando l'importanza della fiducia reputazionale nella sharing economy. L'effetto del response rate è positivo ma di piccola entità, indicando che una maggiore reattività dell'host contribuisce solo marginalmente alle performance. Infine, la variabile Multiproperty host ha effetti negativi sul tasso di occupazione (-0.039) e sul rating (-0.024), suggerendo che la gestione di più proprietà può ridurre l'attenzione personale e la qualità percepita del servizio.

Le caratteristiche strutturali (numero di bagni, camere da letto e capacità ricettiva) mostrano, come atteso, un effetto positivo sul prezzo medio e sul ricavo complessivo, indicando che gli annunci più spaziosi e confortevoli attraggono una domanda disposta a pagare di più. Le dotazioni di Quality amenities risultano tra i driver più forti dell'ADR (+0.311), confermando che la qualità percepita dell'esperienza offerta rappresenta una leva determinante per la valorizzazione economica dell'annuncio. Inoltre, la politica di cancellazione restrittiva ha un impatto negativo sulle valutazioni.

Capacità esplicativa

I valori di R^2 variano tra 0.173 e 0.404, con la maggiore capacità esplicativa registrata per il modello relativo all'ADR (M2). Ciò implica che circa il 40% della variazione nei prezzi giornalieri può essere spiegato dalle variabili incluse nel modello, confermando la solidità statistica e la coerenza interna delle stime. Tutti i modelli risultano globalmente significativi e includono effetti fissi per anno e mese, così da controllare le fluttuazioni stagionali della domanda turistica.

Sintesi interpretativa

Nel complesso, i risultati confermano che i segnali visivi contenuti nella foto profilo dell'host esercitano un'influenza significativa e sistematica sulle performance economiche e reputazionali all'interno della piattaforma. Elementi visivi che trasmettono calore, empatia e autenticità, come il sorriso o la presenza di bambini, aumentano la fiducia e migliorano le performance complessive, mentre segnali che riducono la trasparenza o la riconoscibilità (occhiali da sole, foto di gruppo) sono associati a risultati inferiori. Queste evidenze empiriche sono coerenti con la teoria dei "trust cues" nelle piattaforme digitali, secondo la quale gli utenti interpretano le immagini come segnali reputazionali sostitutivi dell'interazione faccia a faccia. In questo contesto, la presentazione visiva dell'host diventa parte integrante della strategia di fiducia e posizionamento economico su Airbnb, influenzando direttamente la probabilità di prenotazione, il prezzo, il ricavo e la valutazione complessiva ottenuta.

5.2.2 Regressione multivariata logaritmica con interazione

In questa sezione viene presentata la regressione sviluppata per indagare la correlazione tra le caratteristiche visive (*visual cues*) della foto profilo degli host con gli indicatori di performance individuati, con l'ulteriore informazione di come queste caratteristiche si comportano interagendo con la nazionalità italiana dell'host.

Output della regressione

Tabella 5.4: Modelli logaritmici - Face Attributes e Interazioni × Italiano

	(1) M1: ln(Prenotazioni)	(2) M2: ln(ADR)	(3) M3: OCC	(4) M4: ln(RevPAN)	(5) M5: ln(Rating)
Italiano	0.123*** [0.037]	-0.361*** [0.035]	0.108*** [0.014]	-0.119* [0.067]	0.028*** [0.008]
Face	-0.357*** [0.053]	-0.227*** [0.062]	-0.074*** [0.021]	-0.391*** [0.133]	0.059*** [0.008]
Smile	0.303*** [0.054]	-0.022 [0.058]	0.052** [0.022]	0.135 [0.127]	-0.027*** [0.004]
Sunglasses	-0.007 [0.023]	0.051*** [0.017]	-0.024** [0.010]	-0.003 [0.030]	0.017*** [0.002]
Group	0.460*** [0.111]	-0.304*** [0.089]	0.307*** [0.065]	0.458** [0.179]	0.073*** [0.011]
Children	0.283** [0.143]	-0.092 [0.105]	0.043 [0.055]	0.158 [0.187]	-0.010 [0.013]
Animal	-0.166 [0.101]	0.822*** [0.096]	-0.066** [0.030]	-0.091 [0.375]	-0.022*** [0.007]
Face × Italiano	0.269*** [0.057]	0.329*** [0.064]	0.040* [0.023]	0.489*** [0.136]	-0.047*** [0.008]
Smile × Italiano	-0.201*** [0.057]	-0.052 [0.059]	0.015 [0.023]	-0.036 [0.129]	0.033*** [0.004]
Sunglasses × Italiano	0.000 [.]	0.000 [.]	0.000 [.]	0.000 [.]	0.000 [.]
Group × Italiano	-0.693*** [0.114]	0.387*** [0.091]	-0.401*** [0.066]	-0.598*** [0.183]	-0.076*** [0.011]
Children × Italiano	-0.042 [0.149]	0.104 [0.109]	0.075 [0.059]	0.147 [0.196]	0.027* [0.014]
Animal × Italiano	0.155 [0.110]	-0.923*** [0.102]	0.059 [0.037]	-0.041 [0.381]	0.000 [.]
Instantbook	0.151*** [0.019]	0.044*** [0.012]	0.065*** [0.008]	0.292*** [0.023]	-0.004** [0.002]
Response rate	0.004*** [0.000]	0.000 [0.000]	0.002*** [0.000]	0.004*** [0.001]	0.000*** [0.000]
Superhost	0.374*** [0.021]	0.023** [0.011]	0.173*** [0.009]	0.279*** [0.022]	0.042*** [0.001]
Multiproperty host	-0.084*** [0.015]	0.065*** [0.010]	-0.038*** [0.007]	-0.042** [0.019]	-0.023*** [0.001]
Bathrooms	-0.191*** [0.018]	0.223*** [0.017]	-0.060*** [0.008]	0.114*** [0.030]	0.017*** [0.002]
Bedrooms	-0.136*** [0.016]	0.033*** [0.011]	-0.042*** [0.007]	-0.016 [0.022]	0.013*** [0.002]
Number of photos	0.010*** [0.001]	0.002*** [0.000]	0.003*** [0.000]	0.003*** [0.001]	-0.000*** [0.000]
Accommodates	0.071*** [0.006]	0.078*** [0.005]	0.018*** [0.003]	0.109*** [0.008]	-0.003*** [0.001]
Luxury amenities	0.034* [0.020]	0.033** [0.013]	0.017* [0.009]	0.046* [0.026]	0.014*** [0.002]
Quality amenities	0.059* [0.032]	0.291*** [0.020]	0.014 [0.013]	0.272*** [0.052]	0.014*** [0.005]
Strict cancellation	0.002 [0.022]	-0.048*** [0.017]	0.018* [0.010]	0.009 [0.035]	-0.025*** [0.003]
Shared room	-0.420*** [0.043]	-0.549*** [0.051]	-0.188*** [0.019]	-1.063*** [0.131]	-0.089*** [0.015]
Constant	0.028 [0.054]	3.399*** [0.052]	-0.059*** [0.022]	1.618*** [0.104]	4.452*** [0.010]
Year FE	YES	YES	YES	YES	YES
Month FE	YES	YES	YES	YES	YES
R ²	0.216	0.411	0.185	0.278	0.205
N	12098	7844	12098	7844	10344

Standard errors in brackets

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

Note: L'analisi è stata condotta su un campione composto da 12.380 osservazioni.

Anche per l'analisi dell'impatto dei segnali visivi (visual cues) è stato sviluppata una regressione per analizzare le interazioni con la nazionalità italiana, vedi tabella. Le variabili dipendenti sono le medesime delle regressioni precedenti, considerate logaritmiche, e la regressione è OLS.

Oltre alle variabili indipendenti dirette per i segnali visivi, sono state aggiunte le dummy di interazione con la caratteristica italiano dell'host, per andare a valutare se l'impatto delle visual cues cambia in base alla nazionalità dell'host. Anche in questo caso sono state inserite le variabili di controllo utilizzate nei modelli precedenti. Come in precedenza i coefficienti possono essere considerati percentuali, dato che le variabili dipendenti sono espresse in forma logaritmica.

Come osservato in precedenza si conferma un effetto positivo della nazionalità italiana per quanto riguarda le prenotazioni (+12%), il tasso di occupazione (+11%) e il rating (+3%). Si conferma anche un forte effetto negativo sui prezzi medi per notte.

Effetto dei segnali visivi e del volto

Per quanto riguarda i segnali visivi si conferma un effetto positivo del sorriso che si traduce in un 30% in più di prenotazioni e in un maggiore tasso di occupazione (+5%). Ciò conferma che i guest percepiscono il sorriso come un segnale di affidabilità e fiducia. Le foto di gruppo restituiscono effetti positivi sul numero di prenotazioni (+46%), su OCC (+31%) e sui ricavi per notte (+46%). La presenza di bambini ha effetto positivo solo sul numero di prenotazioni (+28%), confermando un effetto fiducia.

Dalle interazioni tra la nazionalità e i segnali visivi emergono degli effetti interessanti. Mostrare il volto nella foto profilo ha benefici più marcati per gli host italiani rispetto a quelli stranieri. Si osservano effetti positivi su prenotazioni (+27%), ADR (+33%) e ricavi (+49%). Anche il sorriso ha effetto sulle prenotazioni maggiormente per gli italiani rispetto agli stranieri. Per gli italiani la scelta di mettere una foto di gruppo risulta deleterio in quanto si osservano effetti negativi di oltre il 40% su numero di prenotazioni, tasso di occupazione e ricavi per notte. Analogamente la presenza di animali penalizza i prezzi medi del 92%. Da questa analisi risulta come negli host italiani si ricerchi una presentazione più impostata e professionale, con la visibilità del volto e del sorriso. Segnali più informali e familiari come foto di gruppo e animali lasciano veicolare meno fiducia ed affidabilità. Da questa analisi risulta una differenza di interpretazione dei segnali visivi a seconda

della provenienza culturale, si potrebbe ipotizzare che gli italiani ricercano host più professionali invece di persone troppo informali e domestici.

Effetti delle caratteristiche dell'annuncio

Per quanto riguarda le caratteristiche dell'annuncio la possibilità di prenotare direttamente e la qualifica di superhost hanno effetti positivi e significativi su tutte le metriche analizzate, andando a confermare quanto osservato in precedenza. La presenza di politiche di cancellazione stringenti e di stanze condivise, invece, si confermano significativamente negativi, con un picco di -50% per ADR. Il numero di bagni e di camere da letto risulta significativo sul numero di prenotazioni con un 0,191 di prenotazioni in meno per ogni bagno e un 0,136 di prenotazioni in meno per ogni camera da letto. Il numero di foto ha effetti positivi sulla performance, ma in modo meno marcato rispetto ad altri indicatori.

Capacità esplicativa

I valori di R^2 variano tra 0.185 e 0.411, con la maggiore capacità esplicativa registrata per il modello relativo all'ADR (M2). Tutti i modelli risultano globalmente significativi e includono effetti fissi per anno e mese, così da controllare le fluttuazioni stagionali della domanda turistica.

Sintesi interpretativa

Gli italiani ricevono più prenotazioni, di conseguenza hanno un tasso di occupazione maggiore, e riescono anche a mantenere un rating maggiore. Il sorriso e le foto di gruppo si confermano positivi ai fini del veicolare fiducia, ma non in modo univoco. Infatti si osservano differenze a seconda della nazionalità dell'host. Per gli host italiani la presenza di foto di gruppo o la presenza di animali restituisce performance negative, andando a denotare la presenza di un filtro interpretativo che varia l'effetto dei segnali visivi a seconda della nazionalità dell'host.

Capitolo 6

Conclusioni

In questo capitolo viene presentato un riassunto dei risultati più importanti che sono emersi dall'analisi empirica effettuata nei capitoli precedenti cercando di interpretarli al meglio sia dal punto di vista teorico che da quello pratico.

6.1 Sintesi dei risultati

L'obiettivo è stato quello di verificare se le piattaforme digitali nonostante si presentano come ambienti aperti e meritocratici possono riprodurre o amplificare le disuguaglianze sociali preesistenti. L'analisi empirica, basata su dati AirDNA relativi agli annunci della provincia di Torino, ha fornito evidenze coerenti con le ipotesi formulate.

In particolare:

- RQ1: La nazionalità e i segnali visivi generano un effetto significativo sulla domanda di prenotazioni.
→ I risultati mostrano che gli host percepiti come più affidabili, in particolare i profili con foto chiare, sorridenti e con il volto ben visibile, ricevono un numero più elevato di prenotazioni. Questo risultato conferma l'ipotesi HP1, mostrando che la fiducia percepita che è influenzata dai segnali visivi gioca un ruolo fondamentale nelle scelte degli utenti. Per quanto riguarda gli italiani, risulta che hanno un numero inferiore di prenotazioni rispetto agli stranieri, forse a causa del fatto che gli stranieri hanno prezzi inferiori.
- RQ2: La nazionalità e i segnali visivi influenzano significativamente il rating medio.
→ Anche in questo caso l'ipotesi è confermata: gli italiani ottengono valutazioni più alte, in maniera coerente con l'idea che gli utenti associano loro qualità

relazionali e di affidabilità. Tuttavia, il vantaggio reputazionale non implica necessariamente un miglioramento economico.

- RQ3: La nazionalità e i segnali visivi influenzano i ricavi per notte (RevPAN).
→ I risultati mostrano un effetto più marcato: nonostante il numero di prenotazioni non risulti più alto per gli host italiani, i ricavi medi per notte sono maggiori, forse a causa di un prezzo medio superiore rispetto agli stranieri. Gli italiani sembra che puntino maggiormente sulla soddisfazione del cliente e sulla qualità del servizio, mentre gli stranieri si concentrano sulla quantità.

Ciò che emerge da questa analisi è che gli host italiani hanno un numero inferiore di prenotazioni, ma riescono a mantenere più alti i prezzi e di conseguenza registrano ricavi maggiori. Per mantenere più alti i prezzi, gli host italiani, puntano sulla qualità del servizio e sulla soddisfazione del guest, infatti registrano rating superiori. Gli host stranieri, invece, registrano un numero maggiore di prenotazioni, ma ricavi inferiori a causa dei prezzi più bassi. Il dato di rating più bassi rispetto agli italiani insieme con il dato di un tasso più alto di stanze condivise, lascia pensare che gli host stranieri prediligano maggiori quantità, anche sacrificando la soddisfazione e il comfort del guest.

Nel complesso, i risultati indicano che le piattaforme digitali non sono ambienti neutri, ma ecosistemi in cui la costruzione della fiducia dipende tanto da caratteristiche oggettive dell'offerta quanto da elementi visivi e identitari. I segnali visivi, come un sorriso, lo sguardo diretto o la chiarezza del volto, funzionano un po' come etichette di reputazione. Sono piccoli indizi che trasmettono fiducia o diffidenza. Questi dettagli, spesso percepiti in maniera inconscia, modificano il modo in cui le persone valutano l'affidabilità di chi hanno davanti, influenzando poi anche le loro scelte di acquisto o di interazione.

6.2 Implicazioni manageriali

I risultati offrono spunti pratici per i gestori di piattaforme digitali, i professionisti del settore turistico e i decisori pubblici. Design inclusivo delle piattaforme:

- Design inclusivo delle piattaforme: È importante creare interfacce e sistemi di valutazione che riducono l'influenza dei pregiudizi legati all'aspetto. Ad esempio, nascondere in parte le informazioni dei profili o usare algoritmi progettati per garantire maggiore equilibrio può aiutare a limitare le discriminazioni basate sul genere o sull'immagine personale.
- Educazione e consapevolezza degli utenti: Programmi di educazione digitale e di sensibilizzazione sui pregiudizi nascosti possono aiutare gli utenti a riconoscere l'influenza dei segnali visivi e a prendere decisioni basate su criteri più oggettivi.

- Strategie per gli host: Gli host possono utilizzare consapevolmente i segnali visivi come leva strategica, curando la qualità delle foto profilo e la comunicazione della propria identità digitale in modo coerente con i principi di autenticità e fiducia.
- Politiche pubbliche e trasparenza algoritmica: Le istituzioni possono promuovere controlli indipendenti sugli algoritmi di classificazione e raccomandazione, per assicurarsi che le piattaforme operino in modo equo e trasparente.

6.3 Limiti della ricerca

La ricerca presenta alcuni limiti.

In primo luogo, il campione geografico è limitato alla provincia di Torino. Questo significa che i risultati non possono essere estesi con sicurezza ad altri luoghi poiché in contesti turistici più grandi o con culture diverse, le dinamiche potrebbero cambiare.

In secondo luogo, i dati utilizzati derivano da osservazioni reali e non da esperimenti controllati, quindi non è possibile stabilire relazioni di causa-effetto certe tra i segnali visivi, fiducia e performance economiche, ma solo evidenziare delle associazioni statisticamente significative.

Infine, la nazionalità è stata dedotta dai nomi degli host e non è detto che sia quella effettiva. Potrebbero quindi essere presenti delle distorsioni causate da nomi ambigui, soprannomi o nomi non riconducibili a nessuna nazionalità.

6.4 Considerazioni finali

La tesi dimostra che la discriminazione nelle piattaforme digitali esiste ed è un problema reale. Nonostante la sharing economy possa avere effetti inclusivi, non si può ignorare come i pregiudizi presenti nel mondo reale si riversino anche nelle piattaforme digitali, rendendole degli ambienti sociali complessi. L'obiettivo deve essere quello di rendere le piattaforme digitali degli ambienti più neutri, ma soprattutto più equi, promuovendo un design più inclusivo e degli algoritmi più trasparenti.

Se è vero che Internet connette il mondo, è giusto che connetta tutti allo stesso modo, senza penalizzare nessuno.

Capitolo 7

Sviluppi futuri

Nonostante i risultati della ricerca siano interessanti e sufficienti per sottolineare la presenza di bias all'interno della piattaforma Airbnb, sarebbe comunque interessante ampliare il panorama in analisi. Sarebbe utile estendere l'analisi a tutto il territorio italiano e non solo alla provincia di Torino. In seguito sarebbe interessante confrontare i risultati riscontrati a rivello nazionale con quelli rilevati in altri stati sia europei che extraeuropei. Inoltre sarebbe interessante scendere maggiormente nel dettaglio della nazionalità e non distinguere solo tra italiani e stranieri, ma analizzare le differenti nazionalità per indagare quali siano i cittadini di quale nazione più penalizzati e se lo sono allo stesso modo in tutti gli stati. Un altro fronte di ricerca potrebbe essere quello di indagare le performance degli host in base alle nazionalità dedotte dalle foto profilo. Ossia indagare gli effetti delle caratteristiche fisiche dell'host, ad esempio la pelle nera o gli occhi a mandorla, per avere un riscontro non solo della discriminazione basata sul nome, ma anche sulla foto profilo. Sarebbe inoltre interessante studiare anche la nazionalità dei guest, per indagare eventuali gruppi maggiormente discriminatori o la presenza di endorsement tra utenti della stessa nazionalità. Infine, il panorama delle piattaforme digitali è vasto e non si limita ad Airbnb, quindi sarebbe utile estendere questa tipologia di ricerca anche ad altre piattaforme, ad esempio Booking o Uber.

Bibliografia

- [1] Anya Marchenko. «The impact of host race and gender on prices on Airbnb». In: *Journal of Housing Economics* 46 (2019) (cit. alle pp. 2, 3, 19).
- [2] Benjamin Edelman e Michael Luca. «Digital Discrimination: The Case of Airbnb.com». In: *Harvard Business School Working Paper No. 14-054* (2014) (cit. alle pp. 2, 4, 19, 21).
- [3] Snehasish Banerjee, Alton Y.K. Chua e Irene Yeo. «Put on your sunglasses and smile: Examining the effects of eyewear and smiling on trust and booking intentions on Airbnb». In: *International Journal of Hospitality Management* (2022) (cit. alle pp. 2, 4, 21).
- [4] Gary S. Becker. *The Economics of Discrimination*. Chicago, IL: University of Chicago Press, 1957 (cit. alle pp. 3, 6).
- [5] Kenneth J. Arrow. «The Theory of Discrimination». In: *Discrimination in Labor Markets*. A cura di Orley Ashenfelter e Albert Rees. Princeton, NJ: Princeton University Press, 1973 (cit. alle pp. 3, 6).
- [6] Edmund S. Phelps. «The Statistical Theory of Racism and Sexism». In: *The American Economic Review* (1972) (cit. alle pp. 3, 6).
- [7] Eva Zschirnt e Didier Ruedin. «Ethnic Discrimination in Hiring Decisions: A Meta-Analysis of Correspondence Tests 1990–2015». In: *Journal of Ethnic and Migration Studies* (2016) (cit. alle pp. 3, 6).
- [8] Alexandre Flage. «Ethnic and Gender Discrimination in Rental Housing Market: Evidence from Meta-Analysis of Correspondence Tests, 2006-2017». In: *Working paper* (2018) (cit. alle pp. 3, 8).
- [9] Alberto Alesina, Michela Carlana, Eliana La Ferrara e Paolo Pinotti. «Revealing Stereotypes: Evidence from Immigrants in Schools». In: *Working paper* (2018) (cit. alle pp. 3, 9).
- [10] Alina Köchling e Marius Claus Wehner. «Discriminated by an algorithm: a systematic review of discrimination and fairness by algorithmic decision-making in the context of HR recruitment and HR development». In: *Business Research* (2020) (cit. a p. 3).

- [11] Minsu Park, Chao Yu e Michael Macy. «Fighting bias with bias: How same-race endorsements reduce racial discrimination on Airbnb». In: *Science Advances* 9.6 (2023) (cit. alle pp. 3, 20).
- [12] Benjamin G. Edelman, Michael Luca e Dan Svirsky. «Racial Discrimination in the Sharing Economy: Evidence from a Field Experiment». In: *American Economic Journal: Applied Economics* 9.2 (2017) (cit. alle pp. 4, 19).
- [13] Eyal Ert, Aliza Fleischer e Dana Kopolovich. «Gender earning gap on digital platforms: The Airbnb case». In: *Tourism Economics* (2024) (cit. alle pp. 4, 32).
- [14] Istituto della Enciclopedia Italiana fondata da Giovanni Treccani. *Discriminazione razziale*. Accesso: 19 ottobre 2025. 2025. URL: <https://www.treccani.it/enciclopedia/discriminazione-razziale/> (cit. a p. 6).
- [15] Asad Islam, Debayan Pakrashi, Liang Choon Wang e Yves Zenou. «Determining the Extent of Taste-Based and Statistical Discrimination: Evidence from a Field Experiment in India». In: *Working paper* (2023) (cit. a p. 6).
- [16] Ian Ayres e Peter Siegelman. «Race and Gender Discrimination in Bargaining for a New Car». In: *American Economic Review* (1995) (cit. a p. 6).
- [17] David Neumark, Roy J. Bank e Kyle D. Van Nort. «Sex Discrimination in Restaurant Hiring: An Audit Study». In: *Quarterly Journal of Economics* (1996) (cit. a p. 6).
- [18] Devah Pager. «The Mark of a Criminal Record». In: *American Journal of Sociology* (2003) (cit. a p. 7).
- [19] U.S. Census Bureau. *Poverty in the United States: 2022*. Washington, D.C.: U.S. Department of Commerce, Census Bureau, 2023 (cit. alle pp. 7, 10, 11).
- [20] Marianne Bertrand e Sendhil Mullainathan. «Are Emily and Greg More Employable Than Lakisha and Jamal? A Field Experiment on Labor Market Discrimination». In: *American Economic Review* 94.4 (2004) (cit. a p. 8).
- [21] Francisco B. Galarza e Gustavo Yamada. «Labor Market Discrimination in Lima, Peru: Evidence from a Field Experiment». In: *World Development* (2014) (cit. a p. 8).
- [22] Lieselotte Blommaert, Marcel Coenders e Frank van Tubergen. «Discrimination of Arabic-Named Applicants in the Netherlands: An Internet-Based Field Experiment Examining Different Phases in Online Recruitment Procedures». In: *Social Forces* (2014) (cit. a p. 8).
- [23] Magnus Carlsson e Stefan Eriksson. «Discrimination in the Rental Market for Apartments». In: *Journal of Housing Economics* (2014) (cit. a p. 8).

- [24] Giovanni Busetta, Maria Gabriella Campolo e Demetrio Panarello. «Immigrants and Italian Labor Market: Statistical or Taste-Based Discrimination?». In: *Genus* (2018) (cit. a p. 8).
- [25] Michelle J. Budig, Joya Misra e Irene Boeckmann. «Racial and Gender Pay Disparities: The Role of Motherhood». In: *The Russell Sage Foundation Journal of the Social Sciences* (2021) (cit. a p. 10).
- [26] Daniel Parent e Javier Gardeazabal. «Performance Pay and the White–Black Wage Gap». In: *Journal of Labor Economics* (2012) (cit. a p. 10).
- [27] Francine D. Blau e Andrea H. Beller. «Black–White Earnings over the 1970s and 1980s: Gender Differences in Trends». In: *The Review of Economics and Statistics* (1992) (cit. a p. 10).
- [28] OECD. *An Introduction to Online Platforms and Their Role in the Digital Transformation*. Rapp. tecn. OECD Publishing. Paris: Organisation for Economic Co-operation e Development (OECD), 2019 (cit. a p. 12).
- [29] Jean-Charles Rochet e Jean Tirole. «Platform Competition in Two-Sided Markets». In: *Journal of the European Economic Association* 1.4 (2003), pp. 990–1029 (cit. a p. 13).
- [30] Geoffrey Evans e Richard Schmalensee. *Matchmakers: The New Economics of Multisided Platforms*. Boston, MA: Harvard Business Review Press, 2016 (cit. a p. 13).
- [31] Geoffrey G. Parker, Marshall W. Van Alstyne e Sangeet Paul Choudary. *Platform Revolution: How Networked Markets Are Transforming the Economy*. New York, NY: W. W. Norton & Company, 2016 (cit. a p. 13).
- [32] Geoffrey Parker e Marshall W. Van Alstyne. «Two-Sided Network Effects: A Theory of Information Product Design». In: *Management Science* (2005) (cit. a p. 13).
- [33] Peter C. Evans e Annabelle Gawer. *The Rise of the Platform Enterprise: A Global Survey*. 1. New York, NY, U.S.: The Center for Global Enterprise, 2016 (cit. a p. 14).
- [34] Jana Valant. *A European Agenda for the Collaborative Economy*. Members' Research Service Briefing. Brussels: European Parliamentary Research Service (EPRS), nov. 2016 (cit. a p. 15).
- [35] Bernard Caillaud e Bruno Jullien. «Chicken & Egg: Competition Among Intermediation Service Providers». In: *RAND Journal of Economics* (2003) (cit. a p. 15).
- [36] Steven Tadelis. «Reputation and Feedback Systems in Online Platform Markets». In: *Annual Review of Economics* (2016) (cit. a p. 15).

- [37] Mingchun Sun e Edison Tse. «When Does the Winner Take All in Two-Sided Markets?». In: *Review of Network Economics* (2007) (cit. a p. 16).
- [38] Arun Sundararajan. «From Zipcar to the sharing economy». In: *Harvard Business Review* (2013) (cit. alle pp. 16, 17).
- [39] Ronaldo C. Parente, José-Mauricio G. Geleilate e Ke Rong. «The sharing economy globalization phenomenon: A research agenda». In: *Journal of International Management* (2018) (cit. alle pp. 16, 17).
- [40] Mingming Cheng. «Sharing economy: A review and agenda for future research». In: *International Journal of Hospitality Management* (2016) (cit. alle pp. 16, 17).
- [41] Timm Teubner e Christoph M. Flath. «Privacy in the sharing economy». In: *Journal of the Association for Information Systems* (2019) (cit. a p. 16).
- [42] Mokter Hossain. «Sharing economy: A comprehensive literature review». In: *Technological Forecasting and Social Change* (2020) (cit. alle pp. 16, 17).
- [43] *Oxford English Dictionary – Sharing economy*. Accesso: 13 ottobre 2025. 2025. URL: https://www.oed.com/dictionary/sharing-economy_n?tab=factsheet#1335003350100 (cit. a p. 16).
- [44] Patcharapar Rojanakit, Rui Torres de Oliveira e Uwe Dulleck. «The sharing economy: A critical review and research agenda». In: *Journal of Business Research* (2022) (cit. a p. 17).
- [45] *Airbnb: case vacanze, chalet, alloggi ed esperienze*. Accesso: 12 ottobre 2025. 2025. URL: <https://www.airbnb.it/> (cit. a p. 17).
- [46] *Airbnb — Wikipedia*. Accesso: 12 ottobre 2025. 2025. URL: <https://it.wikipedia.org/wiki/Airbnb> (cit. a p. 17).
- [47] C. Michael Hall, Girish Prayag, Alexander Safonov, Tim Coles, Stefan Gössling e Sara Naderi Koupaei. «Airbnb and the sharing economy». In: *Current Issues in Tourism* (2022) (cit. a p. 17).
- [48] Georgios Zervas, Davide Proserpio e John W. Byers. «The Rise of the Sharing Economy: Estimating the Impact of Airbnb on the Hotel Industry». In: *Journal of Marketing Research* (2017) (cit. a p. 18).
- [49] Tapio Ikkala e Airi Lampinen. «Monetizing network hospitality: Hospitality and sociability in the context of Airbnb». In: *Proceedings of the 13th European Conference on Computer-Supported Cooperative Work (ECSCW)* (2015) (cit. a p. 18).
- [50] Jun Li, Antonio Moreno e Dennis Zhang. «Pros vs Joes: Agent Pricing Behavior in the Sharing Economy». In: *Working Paper / SSRN* (2016) (cit. a p. 18).

- [51] Keren Horn e Mark Merante. «Is Home Sharing Driving Up Rents? Evidence from Airbnb in Boston». In: *Journal of Housing Economics* (2017) (cit. a p. 18).
- [52] David Wachsmuth e Alexander Weisler. «Airbnb and the Rent Gap: Gentrification through the Sharing Economy». In: *Environment and Planning A: Economy and Space* 50.6 (2018), pp. 1147–1170 (cit. a p. 18).
- [53] Andrea Raboni. *Cos'è la gentrificazione, quando interi quartieri si modernizzano e "cambiano" cittadini*. Set. 2023. URL: <https://www.geopop.it/cose-la-gentrificazione-quando-interi-quartieri-si-modernizzano-e-cambiano-cittadini/> (cit. a p. 18).
- [54] Giovanni Quattrone, Davide Proserpio, Daniele Quercia, Licia Capra e Mirco Musolesi. «Who Benefits from the “Sharing” Economy of Airbnb?» In: *WWW Conference (Pre-print)* (2016) (cit. a p. 19).
- [55] *AirDNA: Short-Term Rental Data Analytics*. Accesso: 12 ottobre 2025. 2025. URL: <https://www.airdna.co/> (cit. a p. 22).
- [56] Karen L. Xie e Zhenxing Mao. «The Impacts of Quality and Quantity Attributes of Airbnb Hosts on Listing Performance». In: *International Journal of Contemporary Hospitality Management* (2017) (cit. a p. 32).
- [57] Luigi Buzzacchi, Francesco Luigi Milone, Emilio Paolucci e Elisabetta Raguseo. «How to react to a shock? Effects of Airbnb hosts' choices and market segmentation at the time of Covid-19». In: *Information & Management* (2023) (cit. alle pp. 32, 33).

Elenco delle figure

2.1	Tabella sinottica dei principali studi analizzati	5
2.2	Tasso di povertà per etnia negli Stati Uniti (2023)	7
2.3	Reddito mediano delle famiglie negli Stati Uniti (2023)	10
2.4	Evoluzione storica del reddito negli Stati Uniti (1967–2023)	11
2.5	Tipologie di piattaforme digitali	14
2.6	Visualizzazione delle relazioni in una piattaforma a due parti	15
4.1	Codice Python per lo scraping dei nomi degli host dal sito Airbnb .	26
4.2	Codice Python per la pulizia dei nomi degli host	28
4.3	Codice Python per la classificazione dei nomi degli host tramite NamSor	30
4.4	Codice Python per l’analisi delle foto profilo degli host	31
4.5	Distribuzione annunci	35
4.6	Distribuzione host	36
4.7	Distribuzione dei Ricavi medi	37
4.8	Distribuzione del Rating medio	38
4.9	Distribuzione del Rating medio	39
4.10	Distribuzione dei Listing Type	40
4.11	Distribuzione dello status Superhost	41
4.12	Distribuzione del numero medio di foto	42
4.13	Distribuzione della politica di cancellazione	43
4.14	Distribuzione delle caratteristiche visive	44
4.15	Distribuzione dell’ADR medio	45
4.16	Distribuzione dell’OCC medio	46
4.17	Distribuzione del RevPAN medio	47
5.1	Diagramma di interazione per l’effetto moderatore della nazionalità nella relazione tra Genere dell’host e Numero di prenotazioni	63
5.2	Diagramma di interazione per l’effetto moderatore della nazionalità nella relazione tra Genere dell’host e ADR	64

5.3	Diagramma di interazione per l'effetto moderatore della nazionalità nella relazione tra Genere dell'host e OCC	65
5.4	Diagramma di interazione per l'effetto moderatore della nazionalità nella relazione tra Genere dell'host e RevPAN	66
5.5	Diagramma di interazione per l'effetto moderatore della nazionalità nella relazione tra Genere dell'host e Rating	67

Elenco delle tabelle

4.1	Output restituito da NamSor	29
4.2	Statistiche descrittive per anno	49
5.1	Modelli logaritmici - Diretti	52
5.2	Modelli logaritmici - Diretti e Interaction	56
5.3	Modelli logaritmici - Foto del profilo host	69
5.4	Modelli logaritmici - Face Attributes e Interazioni × Italiano	73