



**Politecnico  
di Torino**

**Politecnico di Torino**

Corso di Laurea Magistrale in Ingegneria Gestionale

Percorso ICT & Data Analytics per il Management

Anno Accademico 2024/2025

Sessione di Laurea: Novembre 2025

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

Relatori:

*Elisabetta Raguseo*

*Francesco Luigi Milone*

Candidata:

*Nicoletta Maria Carnemolla*

“Technology is neither good  
nor bad; nor is it neutral.”  
Melvin Kranzberg

## ABSTRACT

Il lavoro di tesi ha come punto centrale la discriminazione di genere all'interno delle piattaforme digitali, una delle questioni più particolari della trasformazione digitale. Anche se la sharing economy viene spesso vista come un modello capace di creare più uguaglianza e opportunità per tutti, la realtà è un po' 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 il genere e le caratteristiche visive del profilo, in particolare la foto dell'host, possano avere un impatto reale sulle performance economiche. I dati osservati riguardano 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 è legato alle teorie sui pregiudizi di genere e 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 verso il 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.

I capitoli di questa tesi seguono la struttura di un paper scientifico ovvero in ordine: introduzione, revisione della letteratura con riferimenti a gender gap, piattaforme digitali e Airbnb, poi ipotesi e obiettivi della ricerca, segue il capitolo dedicato all'analisi empirica con i modelli econometrici ottenuti e i rispettivi risultati e infine, per concludere, il capitolo finale che riguarda le conclusioni della tesi con le possibili implicazioni future.

# SOMMARIO

<b>CAPITOLO 1 - INTRODUZIONE .....</b>	<b>1</b>
<b>CAPITOLO 2 – REVISIONE DELLA LETTERATURA.....</b>	<b>2</b>
<b>2.1 Contributo alla letteratura.....</b>	<b>2</b>
<b>2.2 Discriminazione di genere .....</b>	<b>5</b>
2.2.1 Fondamenti teorici .....	6
2.2.2 Principali teorie economiche della discriminazione di genere .....	7
2.2.3 Evidenze sintetiche sul gender pay gap .....	9
2.2.4 Le istituzioni sociali e le radici strutturali della disuguaglianza: l'indice SIGI dell'OECD.....	11
2.2.5 Implicazioni per i contesti digitali.....	15
<b>2.3 Introduzione alle piattaforme .....</b>	<b>16</b>
2.3.1 Caratteristiche strutturali e classificazione delle piattaforme .....	17
2.3.2 Modello economico e dinamiche di mercato.....	18
2.3.3 Strategie competitive e gestione delle piattaforme .....	19
2.3.4 Impatto economico e settori di applicazione.....	20
<b>2.4 Sharing Economy .....</b>	<b>21</b>
2.4.1 Definizioni e inquadramento teorico .....	21
2.4.2 Caratteristiche e dinamiche fondamentali .....	22
2.4.3 Considerazioni conclusive e prospettive critiche.....	22
<b>2.5 Airbnb .....</b>	<b>23</b>
2.5.1 La piattaforma Airbnb .....	23
2.5.2 Evoluzione e impatto della piattaforma.....	24
2.5.3 Gender Gap in Airbnb.....	25
<b>CAPITOLO 3 – IPOTESI E OBIETTIVI DELLA RICERCA.....</b>	<b>27</b>
<b>CAPITOLO 4 – METODOLOGIA .....</b>	<b>28</b>
<b>4.1 - Analisi database AirDNA.....</b>	<b>29</b>
<b>4.2 – Scraping dei nomi degli host e classificazione.....</b>	<b>32</b>
4.2.1 Estrazione dei nomi host tramite web scraping .....	32
4.2.2 Pulizia e normalizzazione dei nomi .....	35
4.2.3 Analisi del genere e della nazionalità tramite API NamSor.....	38
4.2.4 Creazione del dataset finale .....	41
4.2.5 Analisi delle foto profilo degli host .....	41
<b>4.3 - Misure di performance .....</b>	<b>47</b>

<b>4.4 - Strumenti per l'analisi .....</b>	<b>49</b>
4.4.1 Excel .....	49
4.4.2 Stata .....	49
<b>4.5 - Analisi descrittive .....</b>	<b>50</b>
4.5.1 - Distribuzione degli annunci per genere e nazionalità .....	51
4.5.2 - Distribuzione degli host per genere e nazionalità .....	52
4.5.3 - Ricavi medi per genere .....	53
4.5.4 - Rating medio per genere .....	54
4.5.5 - Numero medio di prenotazioni per genere.....	55
4.5.6 - Tipologia di alloggio per genere.....	56
4.5.7 - Status di Superhost per genere.....	57
4.5.8 - Numero medio di foto per genere .....	58
4.5.9 - Politica di cancellazione per genere.....	58
4.5.10 - ADR (Average Daily Rate) per genere .....	59
4.5.11 - Occupancy Rate (OCC) per genere .....	60
4.5.12 - RevPAN per genere .....	61
4.5.13 – Distribuzione delle caratteristiche visive.....	62
4.5.14 – Andamento delle metriche chiave .....	63
<b>CAPITOLO 5 - RISULTATI.....</b>	<b>65</b>
5.1 Analisi dei modelli logaritmici – Diretti per genere e nazionalità. ....	65
5.2 Analisi dei modelli logaritmici – Diretti e di Interazione per genere e nazionalità .....	70
5.3 Analisi dei modelli logaritmici – Diretti per i segnali visivi delle foto profilo degli host. ....	77
5.4 Analisi dei modelli logaritmici – Diretti e di Interazione per i segnali visivi delle foto profilo degli host con la variabile Uomo. ....	81
5.5 – Interaction Plot.....	86
<b>CAPITOLO 6 – CONCLUSIONI E IMPLICAZIONI FUTURE.....</b>	<b>92</b>
6.1 Sintesi dei risultati e risposte alle domande di ricerca .....	92
6.2 Implicazioni manageriali e di policy .....	93
6.3 Limiti della ricerca .....	94
6.4 Direzioni per ricerche future .....	94
6.5 Considerazioni conclusive .....	95
<b>BIBLIOGRAFIA .....</b>	<b>1</b>

***ELENCO DELLE FIGURE..... 8***

***ELENCO DELLE TABELLE ..... 10***

## CAPITOLO 1 - INTRODUZIONE

In questo primo capitolo della tesi viene introdotto il tema della ricerca con il relativo contesto teorico, le ipotesi che ne hanno guidato lo sviluppo e gli interrogativi a cui si intende fornire una risposta. In secondo luogo, viene indicata il contributo dato dalla letteratura scientifica analizzata per portare avanti le analisi effettuate.

La trasformazione digitale che è stata al centro di discussioni in questi ultimi anni ha portato numerose innovazioni, prima tra tutte l'introduzione del fenomeno chiamato sharing economy ovvero ciò che ha combinato le dinamiche del mercato tradizionale con delle nuove possibilità di intermediazione che sono offerte proprio dalle piattaforme digitali. Queste ultime garantiscono maggiore accessibilità, trasparenza e inclusione però allo stesso tempo la struttura algoritmica che ne è alla base può allo stesso tempo generare delle distorsioni che sono legate a delle percezioni soggettive o anche da bias cognitivi.

Airbnb analizzato proprio in questa ricerca rappresenta un caso particolare di piattaforma digitale. Questo ambiente online che viene fondato nel 2008 rappresenta ciò che ha rivoluzionato il settore dell'ospitalità introducendo quello che viene chiamato modello peer to peer basato sulla fiducia, sulla reputazione e sulla visibilità digitale degli host. Diversi studi mostrano che delle differenze legate al genere e alla presentazione visiva ( come ad esempio la viabilità del volto o l'espressione del sorriso) 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. (1) (2) (3).

Infatti, la letteratura internazionale ha evidenziato come nei mercati digitali la costruzione della fiducia non dipende solo da fattori oggettivi ma anche da segnali sociali e visivi che vanno ad agire come indicatori sulla reputazione (4). In particolare si nota che gli host donna tendono ad essere percepiti come più accoglienti e affidabili ottenendo di conseguenza valutazioni più alte e una migliore reputazione online (5) (6). Questo vantaggio però non si traduce sempre in performance economiche superiori perché esistono dei meccanismi di mercato e degli algoritmi per la visibilità che possono penalizzare i profili femminili indirettamente (7) (8).

La tesi si fonda sull'ipotesi che la fiducia degli utenti non dipende solo dalla qualità oggettiva dell'offerta ma anche dal punto di vista soggettivo riferito a fattori identitarie e visivi che vanno ad

## CAPITOLO 2 – REVISIONE DELLA LETTERATURA

influenzare la percezione dell'affidabilità. In questa ottica gli host percepiti come più affidabili in particolare le donne e quelli che hanno un volto visibile nella foto profilo tendono a generare una domanda più elevata e a ricevere valutazioni migliori. Nonostante questo, come viene mostrato da alcuni studi recenti i vantaggi reputazionali possono non tradursi in ricavi superiori a causa di bias di mercato che incidono sulla performance economica degli annunci. (1) (2) (5)

Analizzando sia genere che rappresentazione visiva dei profili degli host questa ricerca si propone di esplorare il legame tra la fiducia, la reputazione e i risultati economici andando a verificare se le piattaforme digitali, anche se vengono definite come spazi neutri, possono contribuire alla creazione di disuguaglianze sociali preesistenti. L'obiettivo finale si articola in due punti, da un lato quello di fornire un'evidenza empirica sul contesto italiano poco indagato nella letteratura sul gender gap nelle piattaforme digitali e dall'altro quello di offrire spunti per migliorare la gestione dei profili su Airbnb individuando degli elementi che favoriscono una rappresentazione più efficace degli host. (4) (8)

In questa prospettiva, la tesi contribuisce al dibattito accademico sui bias di genere nei mercati digitali e ad una riflessione manageriale sul design più inclusivo delle piattaforme. Capire come i segnali visivi e identitarie influenzano la percezione e le performance economiche portare ad orientare le strategie di comunicazione e di interfaccia verso una sharing economy più equa e trasparente.

## CAPITOLO 2 – REVISIONE DELLA LETTERATURA

### 2.1 Contributo alla letteratura

Per un'accurata analisi della letteratura per la seguente tesi partiamo dagli studi classici di Becker, Arrow e Phelps che hanno posto le basi teoriche per comprendere la discriminazione andando a distinguere tra approcci basati su delle preferenze individuali definito come taste-based discrimination e tra quelle fondati su delle aspettative statistiche. In seguito, alcuni studi più recenti come quelli di Blinder e Oaxaca hanno cambiato il modo di analizzare le differenze di salario tra uomini e donne. Hanno introdotto strumenti statistici piuttosto utili, capaci di scomporre e capire meglio da dove nascono queste disparità retributive. Queste analisi sono diventate un punto di riferimento per tantissime ricerche successive, che cercano di spiegare in modo più chiaro quello che viene definito gender wage gap all'interno di mercati del lavoro tradizionali. (10) (11) (12) (13) (14)



Degli studi hanno mostrato come, nonostante i grandi passi avanti fatti dalle donne in termini di istruzione e partecipazione al lavoro, le differenze economiche tra uomini e donne continuano. Anche se oggi le donne studiano di più e lavorano di più il divario nei guadagni rimane evidente e la cosa interessante è anche che solo una parte di questo divario si può spiegare con fattori visibili come l'esperienza, il tipo di impiego o le ore lavorate, il resto sembra essere determinato da una rete di meccanismi strutturali, istituzionali e culturali che continua a influenzare le opportunità delle donne, non solo dal punto di vista economico ma anche sociale. Il *gender gap* è quindi il risultato di un sistema dove certi ruoli, aspettative e comportamenti si rinforzano nel tempo. (15) (16) (17)

I contesti digitali sono entrati a far parte del dibattito accademico con la diffusione delle piattaforme digitali portando nuove domande sulle disuguaglianze nei mercati mediati da algoritmi digitali. L'economia delle piattaforme veniva presentata come un modello accessibile a tutti ma successivamente ha rivelato dinamiche più complesse di esclusione e differenziazione che sono collegate a bias impliciti e meccanismi di reputazione non neutrali. Le piattaforme digitali basano la costruzione della fiducia su sistemi di recensioni, di punteggi e di foto dei profili che nonostante aumentino la trasparenza possono anche portare dei pregiudizi legati al genere. (18) (19) (4) (20)

Airbnb rappresenta, in questo senso, un caso emblematico. La ricerca accademica ha indagato la piattaforma da prospettive diverse, analizzando la costruzione della fiducia, le strategie di prezzo e l'impatto economico sui mercati locali. (1) (21) (22) (23) Parallelamente, studi più recenti hanno evidenziato come anche in questo contesto emergano forme di discriminazione. Edelman e Luca hanno documentato penalizzazioni economiche per gli host afroamericani, mentre lo studio di Koh, Li, Livan e Capra hanno mostrato la presenza di meccanismi di omofilia di genere ed etnica che influenzano i processi di matching. (24) (25)

Il tema del *gender gap* su Airbnb ha ricevuto attenzione crescente negli ultimi anni. Davidson e Gleim hanno quantificato un divario di guadagni a favore degli uomini, riconducibile a strategie di prezzo più elevate e a un approccio più imprenditoriale nella gestione degli annunci. Su e Mattila hanno invece evidenziato l'effetto della congruità di genere tra host e guest: le donne tendono a preferire host femminili, percepite come più affidabili, suggerendo che la fiducia sia mediata anche da fattori identitari. A tali contributi quantitativi si affiancano analisi qualitative come quella di Maier e Gilchrist, che interpretano l'esperienza di hosting femminile in chiave intersezionale, mettendo in luce il legame tra genere, vulnerabilità economica e precarietà abitativa. (5) (6) (8)

La letteratura suggerisce quindi che le piattaforme digitali non sono spazi neutrali, ma contesti socio-tecnici in cui interagiscono elementi economici, culturali e algoritmici. Le differenze di genere

possono tradursi in differenze nelle performance economiche, nella visibilità degli annunci e nella costruzione della reputazione, anche in assenza di discriminazioni esplicite. Gli studi più recenti hanno inoltre richiamato l'attenzione sui segnali visivi dei profili come l'espressione del volto o il sorriso che influenzano la fiducia percepita e la propensione alla prenotazione. (1) (9)

Con tutti questi contributi la ricerca si propone quindi di estendere le evidenze esistenti al contesto italiano, ancora poco esplorato, analizzando la relazione tra genere, rappresentazione visiva e performance economiche degli host all'interno di Airbnb.

Studio	Descrizione sintetica
<b>Banerjee, S., Lens, M., &amp; Pal, A. (2022)</b> <i>Put on your sunglasses and smile: The secret of Airbnb hosts' profile photos. International Journal of Hospitality Management.</i>	<b>Obiettivo:</b> Analizzare come genere, espressione facciale e uso degli occhiali nelle foto profilo degli host influenzino fiducia e intenzione di prenotazione. <b>Variabili indipendenti (VI):</b> Genere (maschio/femmina); espressione facciale (positiva/neutra); occhiali da sole (presente/assente). <b>Variabili dipendenti (VD):</b> Intenzione di fidarsi; intenzione di prenotare. <b>Variabili di controllo:</b> Età e genere del partecipante; percezione di realismo dello scenario sperimentale. <b>Fonti teoriche:</b> Teoria della percezione affettiva; teoria della self-presentation online; studi su trust e facial cues in Airbnb (Ert et al., 2016; Su & Mattila, 2020). <b>Metodologia:</b> Esperimento online 2x2x2 con 524 partecipanti; immagini dal database Cohn-Kanade; analisi tramite regressioni multiple e modelli di mediazione (Hayes, 2013). <b>Risultati:</b> 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. <b>Conclusioni:</b> Le caratteristiche visive influenzano significativamente le decisioni dei guest; preferibili foto con volto visibile e sorriso.
<b>Ert, E., Fleischer, A., &amp; Kopelovich, D. (2024)</b> <i>Gender earning gap on digital platforms: The Airbnb case. Tourism Economics, 31(3), 453-476.</i>	<b>Obiettivo:</b> Analizzare se esista un gender earnings gap tra host Airbnb e determinarne la direzione e le cause. <b>Variabili indipendenti (VI):</b> Genere dell'host (maschio/femmina). <b>Variabili dipendenti (VD):</b> RevPAR (revenue per available room); ADR (average daily rate); occupancy rate. <b>Variabili di controllo:</b> Caratteristiche dell'appartamento (dimensione, distanza dal mare/centro, rating, foto, Airbnb Plus, pets, ecc.); caratteristiche dell'host (esperienza, Superhost, listings, verifica, ecc.). <b>Fonti teoriche:</b> Gender wage gap (Blau & Kahn, 2017); ICT empowerment (Goldin, 2014); fiducia e reputazione (Ert & Fleischer, 2020). <b>Metodologia:</b> 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. <b>Risultati:</b> 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. <b>Conclusioni:</b> Si osserva un'inversione del gender earnings gap, con un vantaggio economico per le donne dovuto alla percezione di fiducia.
<b>Edelman, B. G., &amp; Luca, M. (2014)</b> <i>The impact of host race and gender on prices on Airbnb. Harvard Business School Working Paper.</i>	<b>Obiettivo:</b> Esaminare se razza e genere dell'host influenzino i prezzi degli alloggi, indicando la presenza di discriminazioni implicite. <b>Variabili indipendenti (VI):</b> Razza percepita dell'host (caucasico / afroamericano / asiatico / altro); genere (maschio / femmina). <b>Variabili dipendenti (VD):</b> Prezzo medio dell'annuncio. <b>Variabili di controllo:</b> Caratteristiche dell'appartamento (posizione, tipologia, recensioni, valutazione, dimensione, foto, ecc.); caratteristiche del quartiere (reddito medio, diversità etnica, ecc.). <b>Fonti teoriche:</b> Discrimination economics (Becker, 1971); digital discrimination e trust nelle piattaforme (Edelman & Luca, 2014; Ert et al., 2016). <b>Metodologia:</b> Analisi quantitativa OLS su dataset Airbnb di New York City; classificazione razziale tramite valutatori indipendenti; controlli di robustezza geografica. <b>Risultati:</b> Gli host afroamericani applicano prezzi inferiori (12%) rispetto a host caucasici per proprietà comparabili; differenze non spiegate da qualità o posizione. <b>Conclusioni:</b> Evidenza di discriminazione razziale nei prezzi Airbnb; la trasparenza visiva (foto profilo) amplifica bias impliciti.

Tabella 1 - Tabella sinottica della letteratura sul gender gap in Airbnb

### 2.2 Discriminazione di genere

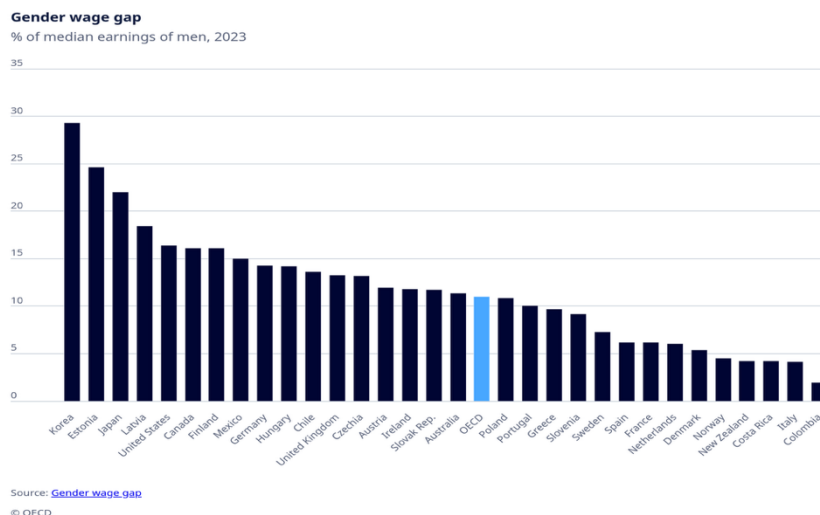
La tematica della discriminazione di genere e della differenza per quanto riguarda la retribuzione tra uomini e donne è stata oggetto di molte indagini economiche e sociologiche. Si è registrata col tempo una riduzione del gender pay gap nei paesi avanzati ma come viene documentato da alcuni studi questo progresso non ha eliminato completamente le disuguaglianze.

Grazie alle ricerche di Blau e Kahn si è potuto mostrare come le differenze per quanto riguarda l'istruzione e l'esperienza lavorativa, che in passato spiegavano una parte significativa del divario oggi hanno un peso più limitato. Fattori strutturali come la segregazione occupazionale, le interruzioni di carriera legate alla maternità e le forme di discriminazione diretta e indiretta purtroppo persistono ancora. Gli autori con il loro lavoro volevano evidenziare quindi di come il gender wage gap si sia ridotto grazie all'aumento del capitale umano femminile ma, persiste ancora una componente non spiegata che può essere attribuita a meccanismi istituzionali e culturali radicati. A questo è anche legato l'ampia meta-analisi internazionale condotta da Weichselbaumer e Winter-Ebmer i quali confermano che, nonostante l'aumento del capitale umano femminile una quota ragionevole del divario rimane ancora non spiegata e riconducibile all'esistenza di barriere strutturali e a dinamiche discriminatorie nel mondo del lavoro. (15) (16)

Ricerche più recenti hanno approfondito l'analisi in particolari contesti. Iwasaki e Satogami hanno mostrato che nei mercati emergenti europei il gender gap salariale è ancora rilevante e influenzato da istituzioni nazionali e norme culturali che portano le donne verso ruoli occupazionali più penalizzanti. Contemporaneamente, Usui con il suo studio ha sottolineato come la segregazione di genere sia l'effetto di interazione tra preferenze individuali (ad es. maggiore avversione da parte delle donne verso orari lunghi a lavoro), frizioni di mercato (ad es. offerta di pacchetti specifici di salario e orario che vanno ad attrarre o più uomini o più donne) e anche elementi discriminatori che determinano una femminilizzazione di settori a basso livello retributivo. Strittmatter e Wunsch hanno infine posto l'accento sulla complessità di misurare correttamente il gender gap in un ambiente caratterizzato da forte segregazione, sottolineando come la comparabilità tra dipendenti e lavoratrici risulti spesso compromessa andando a sovrastimare la componente non spiegata del divario. (17) (26) (27)

Un'altra dimensione di ricerca riguarda le barriere sistemiche che determinano la presenza femminile sul lavoro. Il Social Institutions and Gender Index evidenzia come le disuguaglianze di genere abbiano radici nelle norme sociali e culturali, che incidono sull'accesso all'istruzione, alla salute, alla

rappresentanza politica e al mercato del lavoro. Risulta una visione del gender gap come fenomeno complesso che non può essere analizzato solo in chiave economica ma che si realizza come conseguenza di fattori strutturali, istituzionali e culturali. (28)



*Figura 1 – Gender Wage Gap – Dati del 2023 – Fonte:OECD*

Il grafico sopra va ad evidenziare il gender wage gap nei paesi OCSE nel 2023. L'indicatore viene espresso come percentuale del salario mediano maschile. Si possono notare notevoli differenze internazionali: paesi asiatici ed europei orientali presentano i divari più elevati, mentre paesi come Italia e Colombia registrano valori più contenuti. Ciò conferma che il divario di genere è influenzato da fattori istituzionali e culturali propri di ciascun contesto nazionale.

### 2.2.1 Fondamenti teorici

Un ruolo importante viene rivestito anche dal punto di vista economico e sociale, negli ultimi anni, per la discriminazione. Si ha una crescente consapevolezza riguardo alla disparità nell'occupazione e nei guadagni di uomini e donne che non possono essere spiegate solo da tratti personali come istruzione, esperienza o settore occupazionale.

Da una ricerca recente di Blau e Kahn viene indicato come ci siano ancora ampi divari in termini di retribuzione, opportunità di carriera e accesso alle posizioni più in alto nonostante il grande aumento della partecipazione femminile al mondo del lavoro. (15)

Guardando le stime indicate dall' OECD in merito al gender pay gap si nota come nelle economie più avanzate il divario retributivo medio tra uomini e donne rimane al 12%, con una significativa differenza tra paesi in proporzione ai propri sistemi di welfare e alle loro politiche di uguaglianza. Il Global Gender Gap Report mostra anche che la piena parità economica non potrà essere raggiunta nel breve periodo. (28)

Dal punto di vista analitico, la discriminazione di genere viene indicata attraverso il gender pay gap che rappresenta la differenza media di retribuzione tra lavoratori e lavoratrici. Tale differenza viene calcolata in due modi: o come *raw gap* ovvero come misura della semplice differenza nei salari medi, oppure, come *adjusted gap* che va a controllare per fattori come età, titolo di studio, esperienza e settore lavorativo. La componente residua che non può essere spiegata da variabili osservabili viene indicata come *unexplained gap* ed essa viene interpretata come la parte del differenziale salariale dovuta a comportamenti o meccanismi discriminatori. (15)

La letteratura inoltre distingue tra discriminazione diretta e indiretta. La discriminazione diretta usata quando individui con le stesse qualifiche vengono trattati diversamente in base al genere, mentre la indiretta emerge da pratiche che sembrano neutrali ma presentano poi effetti sfavorevoli per uno dei due sessi. Si parla anche di discriminazione strutturale e cioè quella che si riferisce all'insieme di meccanismi sociali, culturali e istituzionali che limitano l'accesso delle donne alle stesse opportunità economiche degli uomini nel lungo periodo.

Un altro elemento da attenzionare sono i bias cognitivi e gli stereotipi di genere, essi sono dei pregiudizi impliciti e delle rappresentazioni sociali che vanno a influenzare la valutazione delle competenze e delle performance. Dallo studio di Ridgeway viene mostrato come questi fattori contribuiscono a mantenere differenze di status e riconoscimento anche in contesti formalmente paritari. (29)

### **2.2.2 Principali teorie economiche della discriminazione di genere**

A partire dalla metà del XX secolo viene sviluppata l'analisi economica riguardo alla discriminazione di genere. L'analisi aveva l'obiettivo di spiegare le ragioni per cui le donne continuano a ricevere trattamenti economici differenti rispetto agli uomini nonostante la parità di produttività. Grazie a numerose teorie vengono elaborati, con vari approcci, dei modelli che condividono l'idea che la

discriminazione di genere sia il risultato di determinati comportamenti individuali, di aspettative collaborative e di meccanismi di mercato che vanno ad auto-rinforzarsi nel tempo.

Una delle prime formulazioni sistematiche è proposta da Gary Becker nella cosiddetta teoria del gusto per la discriminazione. Secondo questo approccio, alcuni datori di lavoro, colleghi o consumatori provano una disutilità nel lavorare o interagire con individui appartenenti a determinati gruppi ovvero come quello delle donne. Tale preferenza che non ha una natura economica si esprime in decisioni di tipo discriminatorio: i datori di lavoro disposti a “pagare” un costo pur di evitare di assumere donne tenderanno a offrire loro salari più bassi o minori opportunità di impiego. In questo modello, la discriminazione comporta inefficienze di mercato, poiché riduce la produttività complessiva e la competitività dell’impresa. (10)

Un secondo filone teorico introduce la teoria della discriminazione statistica. In un contesto di informazione imperfetta i datori di lavoro non conoscono esattamente la produttività individuale dei candidati e possono utilizzare caratteristiche osservabili come il genere come proxy per stimare la produttività attesa. In tal modo, le donne possono ricevere offerte salariali inferiori non per una minore produttività effettiva, ma perché i datori di lavoro gli attribuiscono, in media, una minore probabilità di permanenza o di disponibilità oraria. Questi meccanismi, se ripetuti nel tempo, generano un equilibrio dove le aspettative discriminatorie si alimentano da sole e influenzano le scelte di investimento in capitale umano delle lavoratrici. (12) (11)

Da ricerche più recenti si nota come vengano estesi questi modelli andando a incorporare anche le frizioni di ricerca (search frictions) e le preferenze individuali. Particolare dimostrazione di ciò viene data dallo studio di Usui che mostra come in mercati del lavoro caratterizzati da vincoli di tempo, informazione incompleta e preferenze eterogenee, le donne possono accettare occupazioni con minore remunerazione ma caratterizzate da orari più contenuti o condizioni maggiormente compatibili con le proprie preferenze rispetto al tempo di lavoro. In questo caso, il gender pay gap non deriva unicamente da vera e propria discriminazione ma anche da scelte adattive determinate da vincoli sociali e strutturali che limitano le opportunità di mercato.

Per concludere queste teorie contribuiscono a delineare un quadro interpretativo articolato: la discriminazione di genere può derivare da preferenze soggettive (10), da meccanismi informativi imperfetti (12) (11), o da interazioni tra vincoli di mercato e norme sociali (26). L’evidenza empirica mostra come nessuna di queste spiegazioni da sola è sufficiente a giustificare la persistenza del divario

di genere, suggerendo che il fenomeno deriva quindi da una combinazione di fattori economici, culturali e istituzionali che si rinforzano reciprocamente nel tempo.

### 2.2.3 Evidenze sintetiche sul gender pay gap

In contemporanea all'elaborazione delle teorie economiche per il gender gap viene sviluppata anche l'analisi empirica per quanto riguarda il gender pay gap. L'analisi aveva l'obiettivo di quantificare e decomporre le differenze salariali tra uomini e donne andando a distinguere la parte che andava attribuita a caratteristiche osservabili da quella legata ai comportamenti discriminatori o a dei meccanismi di mercato che non potevano essere misurati direttamente.

In merito a questo scopo uno degli strumenti più utilizzati per analizzare le differenze di salario che va menzionato è la decomposizione di Blinder-Oaxaca. L'approccio indicato consente la separazione del divario salariale complessivo in due componenti; la prima è quella "spiegata" che è riconducibile a delle differenze riguardo a caratteristiche produttive che possono essere osservate (ad es. istruzione, esperienza, settore e tipologia di impiego), la seconda è invece quella "non spiegata" che riflette delle differenze nei rendimenti di queste caratteristiche e viene interpretata come indicatore della discriminazione o anche di fattori che non vengono osservati. (13) (14)

Tale metodologia è oggi la base standard per l'analisi empirica delle disuguaglianze retributive di genere. (15)

Purtroppo, l'applicazione della tecnica menzionata precedentemente incontra dei limiti metodologici alquanto rilevanti. Da studi più recenti sono stati riscontrati dei problemi, e cioè che quando si è di fronte a un mercato caratterizzato da una forte segregazione occupazionale, non è sempre possibile paragonare uomini e donne riferite alle stesse caratteristiche. Riprendendo lo studio citato precedentemente condotto da Strittmatter e Wunsch, per esempio, esso mostra che in numerosi contesti occupazionali non esistono uomini e donne con caratteristiche davvero paragonabili e di conseguenza, il divario retributivo rischia di essere interpretato come discriminazione (sovrastima della componente non spiegata) anche quando deriva da differenze strutturali nella composizione del mercato. Questo suggerisce la necessità di approcci di stima più robusti, basati sull'applicazione di criteri di common support e su metodologie più flessibili rispetto alle tradizionali decomposizioni parametriche. (27)

Allo stesso tempo, sono state condotte numerose meta-analisi che hanno sintetizzato i risultati di centinaia di studi empirici. Va ricordato per la sua particolare importanza il lavoro citato precedentemente di Weichselbaumer e Winter-Ebmer, i quali hanno analizzato oltre 260 stime relative al divario retributivo provenienti da più di sessanta studi condotti in diversi contesti nazionali. Con questo, gli autori, volevano mostrare che il raw gap tende a ridursi nel tempo mentre l'adjusted gap rimane persistentemente significativa, attestandosi mediamente tra il 10% e il 20%. I risultati della loro ricerca suggeriscono che una parte consistente del divario salariale rimane collegabile a fattori discriminatori o istituzionali, anche a parità di capitale umano e caratteristiche lavorative. (16)

Uno studio più recente da citare è quello di Iwasaki e Satogami, i quali hanno condotto una meta-analisi sui paesi dell'Europa centro-orientale e dell'ex Unione Sovietica, andando ad evidenziare come l'adesione all'Unione Europea e l'introduzione di politiche di uguaglianza hanno contribuito a ridurre il raw gap, pur senza eliminare completamente la componente non spiegata. Da ciò l'idea del gender pay gap come un fenomeno complesso che riflette sia differenze strutturali nei mercati del lavoro sia la persistenza di meccanismi culturali e istituzionali che penalizzano la forza lavoro femminile. (17)

Per concludere quindi in sintesi la letteratura empirica converge su tre risultati principali:

- Il divario retributivo grezzo si è ridotto negli ultimi decenni, grazie all'aumento del capitale umano femminile e alla maggiore partecipazione al lavoro.
- La componente che non si riesce a spiegare resta comunque significativa e ciò che ci siano ancora forme di discriminazione diretta o indiretta.
- Il gender pay gap varia tra paesi e settori, riflettendo l'influenza delle istituzioni, delle norme sociali e delle politiche pubbliche di conciliazione tra vita professionale e familiare.

### **Politiche istituzionali contemporanee: la pay transparency nei paesi OECD**

Al crescente numero di evidenze empiriche sul gender pay gap si è affiancato un aumento delle istituzioni internazionali volte a misurare il divario e a introdurre strumenti per ridurre la persistenza del gender gap. Tra questi una grande importanza va data al OECD per l'elaborazione di linee guida sulla pay transparency che sono orientate a rendere più visibili e verificabili le differenze retributive tra uomini e donne all'interno delle imprese. Dalla lettura del report "Reporting Gender Pay Gaps in OECD Countries" si può vedere come oltre metà dei paesi membri richiede oggi alle aziende di



pubblicare o comunicare i propri dati salariali disaggregati per genere, con l'obiettivo di aumentare la responsabilità organizzativa e ridurre le asimmetrie informative che ostacolano l'eguaglianza retributiva. In alcuni momenti queste misure sono anche integrate da obblighi di audit interni ed eventuali sanzioni, configurandosi come una forma di regolazione attiva del mercato del lavoro.

L'OECD sottolinea come le politiche di trasparenza retributiva nonostante rappresentano un passo significativo verso una maggiore uguaglianza esse però non sono sufficienti a eliminare completamente il divario visto che intervengono solo sui sintomi e non sulle cause della disuguaglianza. Il gender pay gap rimane radicato nelle norme sociali o nei processi culturali che precedono l'ingresso delle donne nel mondo del lavoro.

(30)

### **2.2.4 Le istituzioni sociali e le radici strutturali della disuguaglianza: l'indice SIGI dell'OECD**

In ambito economico e occupazionale si osservano delle disuguaglianze di genere che non possono essere interpretate solamente come esito del mercato del lavoro ma bensì sono legate a sistemi di tipo sociale, culturale e normativo che hanno definito ruoli, aspettative e delle opportunità diverse per uomini e donne. In questo caso va menzionato il Social Institutions and Gender Index (SIGI) del Centro di Sviluppo dell'OCSE che misura la discriminazione nei confronti delle donne all'interno delle istituzioni sociali in 179 paesi. Tenendo conto di leggi, norme sociali e pratiche che limitano i diritti delle donne e delle ragazze e il loro accesso alle opportunità e alle risorse, il SIGI riesce a catturare i fattori alla base delle disuguaglianze di genere.

L'obiettivo del SIGI è quello di sostenere il processo decisionale fornendo dei dati necessari ai responsabili politici e decisionali, esperti e ricercatori, organizzazioni internazionali, quindi anche al pubblico in generale. È una delle fonti di dati ufficiali per il monitoraggio dell'indicatore SDG che afferma che “Se siano in atto o meno quadri giuridici per promuovere, far rispettare e monitorare l'uguaglianza di genere e l'emancipazione delle donne” insieme a UN Women e The World Bank Group's Women Business and the Law. (31).

### **Il quadro concettuale del SIGI**

Il SIGI viene determinato poiché c'era l'esigenza di superare la visione limitata del gender gap che riferita agli indicatori salariali o occupazionali rischiava di trascurare le dimensioni invisibili ma

determinanti come ad esempio le aspettative familiari, le pratiche culturali e la protezione dei diritti fondamentali. L'idea centrale dell'indice SIGI è quella di una disuguaglianza di genere non come il prodotto di differenze individuali ma come quella di strutture istituzionali che vanno a condizionare le scelte e le possibilità di vita delle donne fin da quando sono bambine. Esso integra un approccio più socio-istituzionale che va a porsi come completamento alle analisi economiche sul gender pay gap.

### Le quattro dimensioni dell'indice SIGI

L'indice SIGI si articola in quattro dimensioni, ciascuna delle quali rappresenta una specifica area di potenziale discriminazione istituzionale. Di seguito vengono elencate e descritte le 4 dimensioni, i dati presi si riferiscono a Country = Italia.

#### *a) Discriminazione nella famiglia*

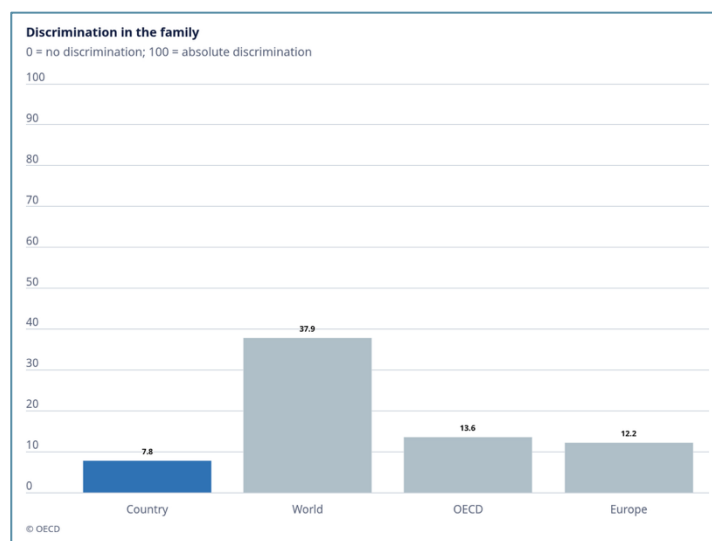
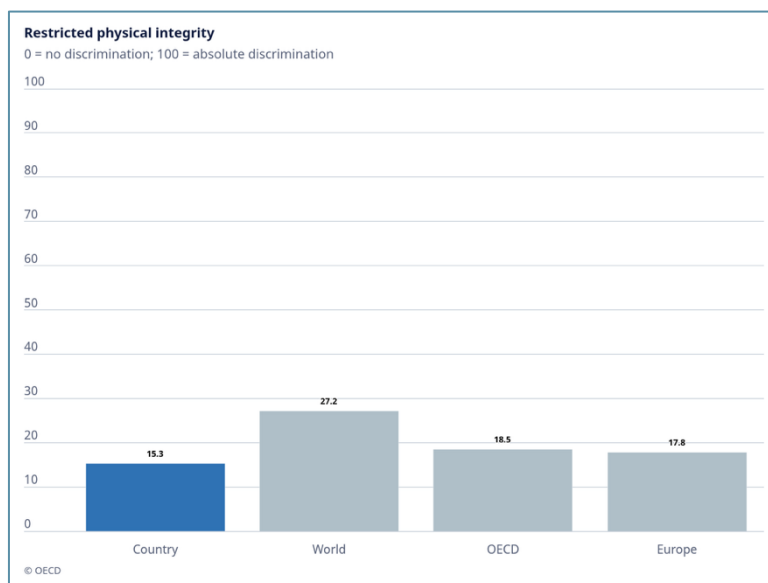


Figura 2 – Grafico che mostra la discriminazione della famiglia – Fonte: OECD

Questa dimensione misura la presenza di norme o pratiche che limitano i diritti delle donne in ambito familiare, riferito ad aspetti quali età minima per il matrimonio, diritti ereditari e potere decisionale domestico. Anche in contesti giuridicamente avanzati rimangono delle disuguaglianze legate alla divisione del lavoro di cura e alle aspettative di ruolo, che continuano a gravare in modo sproporzionato sulle donne.

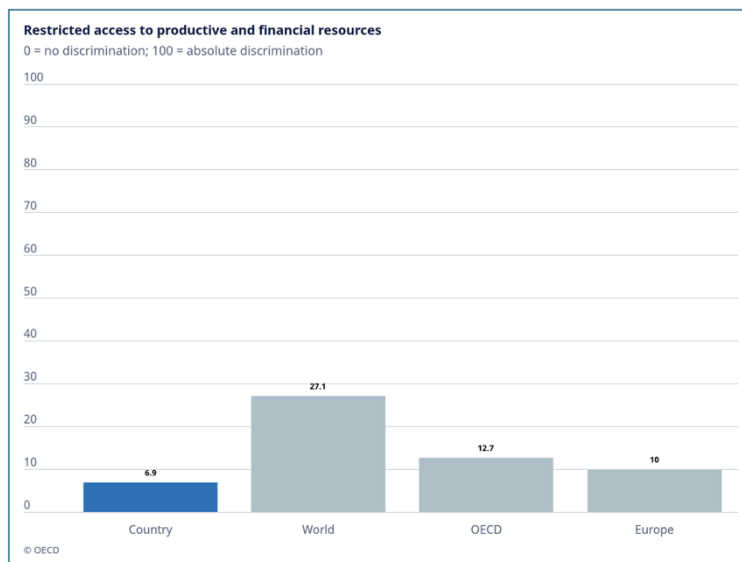
### *b) Integrità fisica e autonomia corporea*



*Figura 3 – Grafico che mostra l'integrità fisica per le donne - Fonte: OECD*

Questa dimensione valuta la protezione contro la violenza di genere, gli abusi domestici e le limitazioni all'autonomia riproduttiva. Le poche ed adeguate politiche di prevenzione e protezione incidono non solo sulla sicurezza individuale, ma anche sulla partecipazione economica e pubblica delle donne, evidenziando il legame tra libertà personale e inclusione sociale.

### *c) Accesso alle risorse produttive ed economiche*

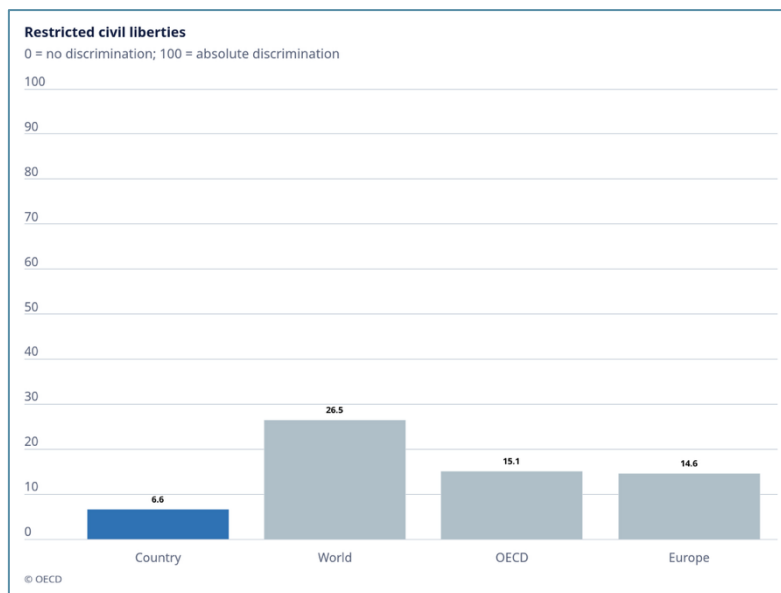


*Figura 4 – Mostra l'accesso alle risorse produttive – Fonte: OECD*

Questa dimensione analizza il diritto delle donne ad accedere e controllare risorse economiche, come proprietà, credito e opportunità lavorative. Anche in presenza di pari diritti legali, rimangono delle

barriere informali e pratiche discriminatorie che limitano la capacità femminile di accumulare capitale economico e negoziare condizioni eque nel mercato del lavoro.

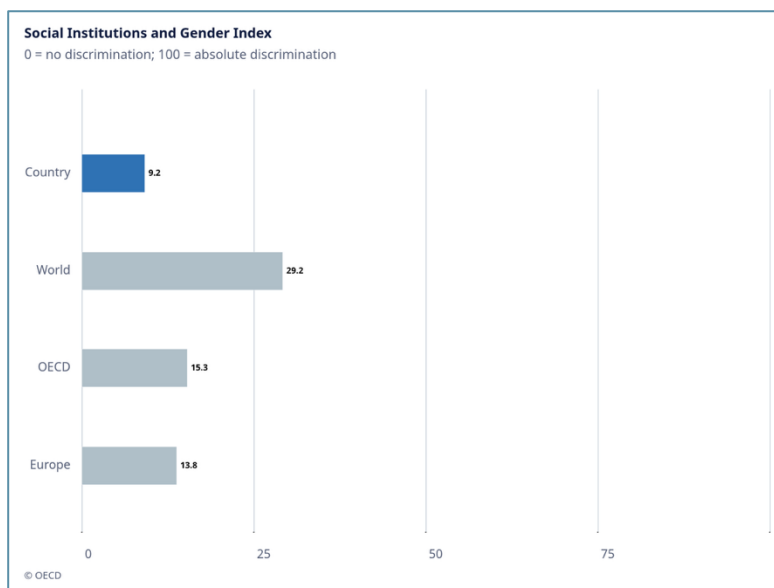
### *d) Libertà civili e partecipazione pubblica*



*Figura 5 – Mostra le libertà civili per le donne – Fonte: OECD*

Questa dimensione riguarda i diritti di cittadinanza, partecipazione politica, accesso alla giustizia e mobilità. La sottorappresentanza femminile nei processi decisionali e nelle cariche pubbliche limita la capacità delle donne di influenzare le politiche che le riguardano, riproducendo un divario di potere che precede quello economico.

### **Il caso italiano nel confronto internazionale**



*Figura 6 – Mostra il caso Italiano confrontato internazionalmente – Fonte: OECD*

Nel confronto internazionale, l'Italia occupa una posizione intermedia nell'indice SIGI. Il quadro legislativo nazionale garantisce formalmente l'uguaglianza dei diritti, ma continuano delle criticità legate alla sfera familiare e alle opportunità economiche. Le aspettative sociali sulla divisione del lavoro domestico, la carenza di servizi per la conciliazione vita-lavoro e la persistente sottorappresentanza nelle cariche pubbliche evidenziano che le disuguaglianze non derivano da mancanza di norme, ma da modelli culturali tradizionali.

### **Implicazioni per l'analisi delle disuguaglianze di genere**

L'evidenza fornita dal SIGI conferma che le disuguaglianze economiche e occupazionali non possono essere affrontate esclusivamente attraverso politiche di mercato o misure di trasparenza salariale. Le base del gender gap sono da collocarsi anche all'interno delle istituzioni sociali dove persistono norme e pratiche che vanno ad influenzare le scelte educative, professionali e familiari delle donne. Di conseguenza la riduzione del divario di genere richiede interventi economici e anche di trasformazioni culturali e istituzionali più orientate al riconoscimento vero e proprio dell'autonomia e della partecipazione delle donne in tutte le sfere della vita sociale.

(32)

### **2.2.5 Implicazioni per i contesti digitali**

I meccanismi di disuguaglianza osservati nel mercato del lavoro tendono a riprodursi anche in contesti digitali anche se con modalità differenti.

Le tecnologie dell'informazione e della comunicazione (ICT) sono spesso presentate come strumenti di empowerment femminile in quanto riducono le barriere di accesso al mercato, favoriscono la flessibilità lavorativa e limitano il ruolo dell'intermediazione diretta, andando ad attenuare le forme di discriminazione legate al datore di lavoro. (33) (28)

Alcune ricerche mostrano però in realtà che anche la digitalizzazione non elimina necessariamente i pregiudizi già preesistenti ma introduce purtroppo nuovi meccanismi attraverso cui si manifesta la discriminazione.

Un importante studio da menzionare in merito a questo discorso è quello di Edelman e Luca che va ad evidenziare come nella piattaforma Airbnb gli host afroamericani ricevono meno prenotazioni e guadagni inferiori rispetto agli altri host con caratteristiche simile a causa della visibilità delle foto del profilo. Da questo risultato si può comprendere come alcuni elementi apparentemente neutrali

come, ad esempio, sistemi reputazionali e profili utente possono diventare veicolo di bias impliciti che vanno ad amplificare le disuguaglianze anziché attenuarle. (24)

Le piattaforme digitali rappresentano un laboratorio per osservare l'evoluzione delle disuguaglianze di genere. Esse combinano elementi di mercato tradizionale come, ad esempio, la competizione o anche la valutazione delle performance con nuove logiche di interazione mediate da algoritmi e meccanismi reputazionali. I successivi capitoli consentiranno appunto di comprendere se e in che misura il progresso in ambito tecnologico stia realmente contribuendo a ridurre il gender gap o, al contrario, lo stiano riproducendo sotto nuove forme.

### **2.3 Introduzione alle piattaforme**

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 che forse erano già esistenti come quelle di genere.

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. (34) (35) (36)

### **2.3.1 Caratteristiche strutturali e classificazione delle piattaforme**

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. (37)

TIPOLOGIA	DESCRIZIONE	ESEMPI
Piattaforme di <b>transazione</b>	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 <b>innovazione</b>	Tecnologie, prodotti o servizi che rendono possibile a un gruppo di imprese – potenzialmente identificabili come partner di un ecosistema innovativo* – di sviluppare tecnologie, prodotti o servizi complementari tra loro.	Microsoft, Linux, Oracle.
Piattaforme di <b>integrazione</b>	Tecnologie, prodotti o servizi che integrano una funzione di <b>transazione</b> e una di <b>innovazione</b> .	Apple, Alphabet, Amazon, Facebook.
Piattaforme di <b>investimento</b>	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 7 - Tipologie di piattaforme digitali - (38)

Come si vede dall'immagine sopra 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 finanziano nuove imprese ad alto potenziale (es. Softbank, Naspers, Priceline). (38)

### 2.3.2 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 subsidiare 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 stimolarne la crescita. In questo modo la massa critica viene raggiunta progressivamente e ciò consente alla piattaforma di diventare autosostenibile.

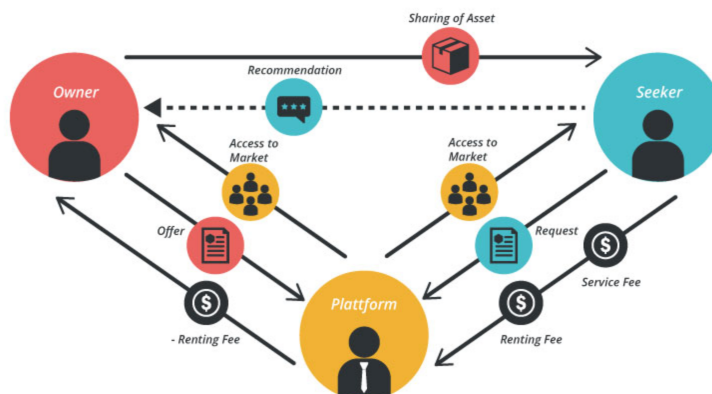


Figura 8 - Visualizzazione delle relazioni in una piattaforma a due parti Fonte: *A European Agenda for the Collaborative Economy* (39)

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. (40)

### 2.3.3 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. (4)

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. (41)

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.

### **2.3.4 Impatto economico e settori di applicazione**

Di seguito un sottocapitolo sui diversi settori di applicazione per le piattaforme digitali. Partendo dal discorso sulla mobilità urbana vanno indicati i nomi delle piattaforme Uber e Lyft, ovvero quelle che hanno modificato i servizi di trasporto portando più flessibilità e accessibilità.

Continuiamo con il mondo dell’ospitalità dove citiamo le realtà di Airbnb e HomeExchange, 2 piattaforme che hanno trasformato il modo di concepire l’alloggio temporaneo creando un nuovo mercato in parallelo a quello degli alloggi tradizionali (ad es. gli hotel). Va riconosciuta un’importanza anche al commercio elettronico dove ambienti online come Amazon o Alibaba hanno centralizzato i processi di scambio. Per concludere indichiamo anche un riferimento all’ambito culturale e di intrattenimento con le realtà di YouTube e Spotify che hanno cambiato il rapporto tra produttori e consumatori di contenuti portando modelli più partecipativi e interattivi.

Ci sono inoltre da ricordare anche la nascita di nuove forme occupazionali e la diffusione di quella che viene definita gig economy. Questa espressione viene utilizzata per indicare il fenomeno in cui i lavoratori autonomi offrono dei servizi personali tramite le piattaforme digitali. Con questa idea si è ottenuta una maggiore valorizzazione di alcune risorse che già esistevano e che erano sottoutilizzate, andando a favorire una transizione da modelli produttivi tradizionali a quelli più flessibili. Quindi, la

platform economy può essere indicata come una nuova fase dell'organizzazione contemporanea a livello economico che viene basata su infrastrutture digitali secondo quanto riportato dall'analisi di Kenney e Zysman. (18)

### **2.4 Sharing Economy**

L'economia collaborativa, definita come Sharing Economy, è un modo diverso di intendere il mercato e cioè si basa sulla condivisione e sullo scambio di risorse tra persone, rese possibili dalle piattaforme digitali di cui abbiamo parlato nel capitolo precedente. (42) (43)

#### **2.4.1 Definizioni e inquadramento teorico**

Ci sono vari modi per capire cos'è davvero la Sharing Economy, e questo fa capire quanto sia un fenomeno complicato da studiare. La condivisione di beni o servizi può assumere forme molto diverse, a seconda del contesto in cui ci si trova ovvero quello economico, sociale o anche tecnologico. La base di questa particolare economia è quella di mettere a disposizione delle risorse che possono essere di tipo materiale (es. un'auto o una casa) o di tipo immateriale (ad es. competenze o tempo), che senza resterebbero non utilizzate. (44)

Una possibile definizione viene fornita dal Oxford English Dictionary ossia: un sistema economico in cui beni o servizi sono condivisi tra individui attraverso internet, spesso con remunerazione. Andando a mettere in luce la dimensione economica e non solo altruistica di questi scambi. (45)

Con lo studio di Botsman e Rogers viene introdotto quello che viene chiamato collaborative consumption che esprime la dimensione sociale della condivisione e il ruolo della tecnologia come strumento chiave di nuovi comportamenti degli utenti. Sundararajan invece ha determinato un'altra espressione ancora ovvero quella che viene definita crowd-based capitalism che interpreta la Sharing Economy come una forma di capitalismo dove la partecipazione economica diffusa viene resa possibile grazie all'infrastruttura digitale. (42) (43)

La Sharing Economy quindi non è rappresentata come un nuovo semplice modello di consumo, ma come cambiamento istituzionale nella configurazione dei mercati, che interseca realtà economiche, tecnologiche e sociali. (42)

### 2.4.2 Caratteristiche e dinamiche fondamentali

Ci sono una serie di caratteristiche che determinano la logica economica e il potenziale di trasformazione dello Sharing Economy. Il primo elemento riguarda la centralità dell'accesso rispetto alla proprietà. Il valore non è più nel possesso di un bene, ma nella possibilità di usufruirne in modo flessibile e temporaneo, riducendo sprechi e favorendo un uso più sostenibile delle risorse (46) (47).

Un secondo elemento distintivo è rappresentato dal ruolo delle piattaforme digitali elemento già attenzionato nel capitolo precedente.

La flessibilità organizzativa costituisce un elemento chiave di questi modelli. L'assenza di consistenti asset fisici e la centralità delle tecnologie digitali consentono un elevato grado di adattarsi ai diversi contesti locali (48).

### 2.4.3 Considerazioni conclusive e prospettive critiche

La Sharing Economy rappresenta dunque un fenomeno centrale dell'economia contemporanea, essa riesce a collegare innovazione tecnologica, nuove modalità di consumo e trasformazioni socio-economiche. Essa porta avanti un utilizzo più efficiente delle risorse e incoraggia modelli di partecipazione più flessibili e inclusivi.

Bisogna considerare ad oggi anche la crescente attenzione verso il ruolo delle piattaforme digitali e le implicazioni della *sharing economy*. All'inizio essa viene presentata come un modello collaborativo che era fondato sulla condivisione di risorse poco utilizzate, ma con il tempo si è evoluta in un sistema sempre più orientato alla logica economica e quindi caratterizzato da una grande attività di mercato e da concentrazione del potere economico nelle mani di poche piattaforme di dimensione globale sempre più crescente.

Oskam e Boswijk (49) hanno mostrato come il caso di Airbnb che verrà studiato meglio nei capitoli successivi, nato come strumento di condivisione domestica, si sia rapidamente evoluto in un sistema orientato alla professionalizzazione degli host, generando tensioni significative con i mercati immobiliari locali. Allo stesso modo, Hall et al. evidenziano come la promessa di inclusività e

democratizzazione tipica delle piattaforme digitali coesista con effetti distributivi regressivi e con nuove forme di precarietà lavorativa. (19)

La letteratura afferma che la possibilità di partecipare con successo alla Sharing Economy non dipende solo dall'impegno individuale, ma presuppone il possesso di un insieme di risorse economiche, sociali e digitali che risultano distribuite in modo non uguale tra la popolazione.

La Sharing Economy quindi può essere interpretata come un fenomeno che può essere espresso in due modi diversi, da un lato costituisce un motore di innovazione e di efficienza mentre dall'altro rischia di aumentare le differenze.

### **2.5 Airbnb**

Airbnb 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. (50)

#### **2.5.1 La piattaforma Airbnb**

Airbnb viene fondata a San Francisco nell'ottobre del 2007 da Chesky, Gebbia e Blecharczyk. Nasce come progetto sperimentale di ospitalità condivisa e diventa subito un attore centrale nei settori del turismo e dell'immobiliare. Come espresso da Hall et al., il successo della piattaforma si deve alla sua capacità di rispondere alla crescente domanda di esperienze di viaggio personalizzate, affermandosi come un'alternativa più flessibile all'ospitalità tradizionale. Airbnb ha permesso quindi una connessione diretta tra host e ospiti. (51) (19)

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, questo modello ha introdotto un nuovo tipo di concorrenza per il settore alberghiero, contribuendo a modificare profondamente il mercato turistico. (22)

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 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. (52)

Oltre ai sistemi reputazionali, la piattaforma adotta algoritmi di pricing dinamico e criteri di visibilità personalizzata. Gli host possono scegliere da soli 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 elementi come la frequenza di risposta, il numero di recensioni, la qualità del profilo e l'utilizzo di funzionalità come la prenotazione immediata. Secondo Li, questi strumenti premiano chi adotta strategie più sofisticate di gestione, creando asimmetrie tra host occasionali e host professionali. (21)

### **2.5.2 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) 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. (53)

Wachsmuth e Weisler (2018) evidenziano la presenza di un effetto di gentrificazione guidato dalla sharing economy. (54)

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). (55)

Infatti come osservato da Quattrone et al. 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. (23)

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.

### 2.5.3 Gender Gap in Airbnb

Airbnb è una piattaforma digitale ben lontana dall'essere definita, come citato nei capitoli precedenti, uno spazio neutrale e anzi tende a riprodurre e in alcuni casi ad amplificare le disuguaglianze di genere esistenti. Il gender gap sulla piattaforma si manifesta in forme multiple, che includono disparità nei guadagni, asimmetrie reputazionali, vincoli di accesso e discriminazioni implicite.

Davidson e Gleim mostrano come gli host uomini ottengano mediamente ricavi annui più elevati rispetto alle donne, anche a parità di caratteristiche strutturali di alloggio e di localizzazione. Questa differenza non è legata alla qualità dell'offerta, ma piuttosto a fattori operativi e strategici e cioè gli uomini tendono a fissare prezzi più alti, a gestire un numero maggiore di prenotazioni e ad avere proprietà con maggiore disponibilità di posti letto. Le donne, pur ottenendo valutazioni mediamente migliori, non riescono a tradurre questo vantaggio reputazionale in un equivalente rendimento economico, a causa di strategie di pricing più caute, vincoli di tempo e carichi familiari che limitano la loro flessibilità operativa. Queste disuguaglianze non derivano soltanto da discriminazioni esplicite, ma emergono dall'interazione tra i meccanismi della piattaforma e le condizioni sociali e individuali degli host. (5)

Anche la costruzione della fiducia risulta influenzata dal genere: Su e Mattila mostrano che le donne tendono a preferire host dello stesso sesso, percepiti come più affidabili, mentre l'effetto è assente tra i consumatori uomini. (6)

Parallelamente, Ert, Fleischer e Kopolovich evidenziano come le fotografie del profilo e in particolare l'espressione facciale delle host incidano significativamente sulla percezione di affidabilità e sull'intenzione di prenotazione, senza tuttavia tradursi in guadagni equivalenti. (1)

Sul piano strutturale, Davidson e Gleim rilevano che gli uomini sono più attivi nella gestione di interi appartamenti o di proprietà multiple, mentre le donne si concentrano su stanze private o soluzioni abitative meno professionalizzate, spesso collegate alla propria residenza. Tale differenziazione si

traduce in margini economici inferiori per le host, poiché gli interi appartamenti generano rendimenti più elevati. (5)

Maier e Gilchrist, con un approccio intersezionale, mostrano inoltre che molte donne intraprendono l'attività di hosting da posizioni di vulnerabilità economica, utilizzandola come strategia di sostentamento più che come scelta imprenditoriale, evidenziando come il gender gap rifletta disuguaglianze di genere, classe e precarietà abitativa. (8)

A queste dinamiche si aggiungono i meccanismi algoritmici della piattaforma: come mostrato da Haldar et al., il sistema di ranking di Airbnb privilegia gli annunci con maggiore probabilità di prenotazione, qualità e frequenza di aggiornamento, premiando chi dispone di più tempo e competenze digitali. In questo modo, gli algoritmi, pur formalmente neutrali, generano effetti cumulativi che rafforzano le disuguaglianze di genere, favorendo profili più attivi e professionalizzati. Una prospettiva femminista critica, infine, invita a leggere queste dinamiche non solo in chiave economica ma come espressione del rentier capitalism, in cui il lavoro emozionale e relazionale delle donne host rimane spesso invisibile e sottovalutato. (3)

In sintesi, il gender gap in Airbnb è un fenomeno multidimensionale che scaturisce dall'interazione di fattori economici, reputazionali, strutturali e tecnologici: la piattaforma, lungi dall'essere neutrale, agisce come dispositivo socio-tecnico che filtra e premia gli utenti in modo diseguale, rendendo necessaria una riflessione integrata tra analisi empirica, critica sociale e intervento regolatorio.



## CAPITOLO 3 – IPOTESI E OBIETTIVI DELLA RICERCA

In questo capitolo vengono descritti gli obiettivi della ricerca determinati dopo un'attenta analisi della letteratura da cui si è preso spunto per formulare delle domande di ricerca e di conseguenza delle ipotesi legate a queste ultime.

La presente ricerca ha come obiettivo quello di analizzare quanto le caratteristiche di genere e le componenti visive del profilo degli host possono influenzare la domanda, la reputazione e le performance economiche nella piattaforma Airbnb.

Partendo da un dataset dettagliato che viene fornito da AirDNA relativo agli annunci presenti nella provincia di Torino lo studio attraverso un approccio empirico prova 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.

In particolare vengono poste le seguenti domande di ricerca:

- **RQ1.** Il genere e i segnali visivi delle foto profilo (visual cues (9)) dell'host generano un effetto significativo sulla domanda di prenotazioni?
- **RQ2.** Il genere e i segnali visivi delle foto profilo (visual cues) dell'host generano un effetto significativo sul rating dell'annuncio?
- **RQ3.** Il genere 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 (donne, segnali visivi positivi) generano una maggiore domanda di prenotazioni.
- **HP2.** Gli host percepiti come più affidabili (donne, segnali visivi positivi) ottengono rating più alti.
- **HP3.** Gli host percepiti come più affidabili (donne, segnali visivi positivi) ottengono ricavi maggiori (RevPAN).

(1)

## CAPITOLO 4 – METODOLOGIA

Nei capitoli seguenti è possibile invece trovare il cuore della ricerca, essi hanno l'obiettivo di tradurre in analisi quantitativa tutte le riflessioni più teoriche fatte nei capitoli precedenti. Dopo la parte più teorica i capitoli seguenti si concentrano invece sul caso di Airbnb andando ad analizzare i dati relativi a Torino con l'obiettivo di verificare se esistono delle differenze di performance tra gli host uomini e donne e individuare se esistono dei fattori che le determinano.

L'analisi effettuata è basata su un dataset di dati panel fornito da AirDNA, i dati sono riferiti all'orizzonte temporale che va dal 2017 al 2024 e va a raccogliere delle informazioni sugli annunci Airbnb come per esempio le caratteristiche delle proprietà, politiche adottate dagli host e alcune performance economiche da cui siamo partiti per il calcolo dei KPI utilizzati per i modelli econometrici.

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 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 domestici (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 metodologie usate nella ricerca si basano su dei modelli di regressione multipla con effetti fissi nel tempo in modo da considerare anche le differenze che non si vedono direttamente tra i vari casi analizzati per capire con maggiore precisione come le diverse variabili si influenzano 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 che possono sembrare secondari, come il genere, la nazionalità o anche la foto del profilo, possono avere un impatto concreto sui risultati economici degli host.

In questo modo, lo studio contribuisce ad arricchire il dibattito sul gender gap 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.

### **4.1 - Analisi 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. (56)

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.

Nello specifico le variabili disponibili sono le seguenti:

### *1. Caratteristiche dell'annuncio*

- Property ID: stringa identificativa unica di una singola proprietà, utilizzata per distinguerla nel database.
- Reporting Month: mese e anno di riferimento.

### *2. Variabili di performance*

- Revenue USD: ricavi totali generati dall'annuncio, espressi in dollari statunitensi.
- Number of Reservations: numero complessivo di prenotazioni ricevute nell'intervallo temporale considerato.
- Reservation Days: numero di giorni in cui l'alloggio è stato effettivamente prenotato.
- Available Days: numero di giorni in cui l'alloggio è rimasto disponibile e non prenotato.

### *3. Caratteristiche della proprietà*

- Host ID: codice univoco identificativo dell'host.
- Property Type: tipologia di proprietà, che può includere appartamento, casa indipendente, stanza, ecc.
- Listing Type: tipo di sistemazione offerta, ad esempio intero appartamento, stanza privata, stanza condivisa.
- Property Manager: indicazione se la proprietà è gestita da un'agenzia.
- Host Type: classificazione dell'host in base al numero di annunci gestiti.
- Comune: comune in cui si trova l'alloggio.
- Latitude: coordinata geografica di latitudine dell'alloggio.
- Longitude: coordinata geografica di longitudine dell'alloggio.
- Bedrooms: numero di camere da letto presenti nell'alloggio.
- Bathrooms: numero di bagni presenti nell'alloggio.
- Max Guests: numero massimo di ospiti che la proprietà può ospitare.

- Cancellation Policy: politica di cancellazione adottata dall'host (flessibile, moderata, stretta, molto stretta).
- Minimum Stay: durata minima del soggiorno richiesta.
- Created Date: data di creazione dell'annuncio nella piattaforma.
- Response Rate: percentuale di risposte date dall'host alle richieste di prenotazione.
- Response Time: tempo medio di risposta alle richieste.
- Number of Photos: numero di fotografie caricate nell'annuncio.
- Instantbook: indica se l'alloggio è prenotabile immediatamente senza necessità di approvazione da parte dell'host ("1 = vero", "0 = falso").
- Airbnb Superhost: indica se l'host ha lo status di Superhost su Airbnb ("1 = vero", "0 = falso").
- Listing URL: link diretto all'annuncio sulla piattaforma.
- Listing Main Image URL: link all'immagine principale dell'annuncio.
- Overall Rating: valutazione media complessiva data dagli ospiti su una scala da 0 a 100.
- Airbnb Communication Rating: valutazione specifica relativa alla comunicazione con l'host.
- Airbnb Accuracy Rating: valutazione sull'accuratezza della descrizione dell'annuncio.
- Airbnb Cleanliness Rating: valutazione sulla pulizia dell'alloggio.
- Airbnb Checkin Rating: valutazione sull'esperienza di check-in.
- Airbnb Location Rating: valutazione sulla posizione dell'alloggio.
- Airbnb Value Rating: valutazione del rapporto qualità/prezzo.

### *4. Dotazioni e servizi*

- Has Kitchen: indica se l'alloggio dispone di una cucina ("1 = vero", "0 = falso").
- Has Parking: indica se l'alloggio dispone di un parcheggio ("1 = vero", "0 = falso").
- Amenities: elenco o codifica dei servizi e dotazioni presenti nell'alloggio (es. piscina, aria condizionata, Wi-Fi, ecc.).

### *5. Variabili territoriali*

- SLL\_2011\_T: codifica territoriale, relativa a un sistema statistico o censuario del 2011.
- DEN\_SL2011: densità abitativa o altra misura territoriale 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 9) è stata svolta su ambiente macOS, utilizzando ChromeDriver, componente che consente di controllare in modo automatizzato Google Chrome.

Per non avere problemi dovuti agli accessi automatici e mantenere la sessione stabile, lo script è stato impostato per avviare Chrome usando il profilo utente locale, con la modalità headless disattivata. Con questa soluzione è stato possibile simulare una navigazione manuale potendo accedere ai contenuti senza blocchi o limitazioni imposti 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.

## CAPITOLO 4 – METODOLOGIA

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 come si può vedere dalla Tabella 1.

host_id	host_name
123456789	Anna
987654321	Luca

*Tabella 2 - Dataset con host-id e host-name*

```

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 = []          # Tutti gli ID con nome
unique_first_names = set() # Solo i nomi unici per Genderize.io

# === 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-ZàèéòìùÀÈÉÌÒÙçÇ\']+", "",
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"[✓] {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"Nomini unici trovati: {len(unique_first_names)}")
print("Salvato file 'unique_first_names.txt'")

```

Figura 9 - 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 10), 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 poiché non davano alcuna informazione sul nome (ad esempio: “A”, “D”).
- Mantenimento dei nomi di due lettere poiché alcuni risultano validi (es. Jo, Lu).

## CAPITOLO 4 – METODOLOGIA

- Esclusione dei nomi contenenti parole appartenenti a una blacklist appositamente elaborata (es. home, guest, stay, house, srl, Torino, Interhome, ecc.) per 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("unique_first_names.txt")

# 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("unique_first_names_pulito.txt", index=False)

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

Figura 10 - Codice Python per la pulizia del database

#### 4.2.3 Analisi del genere e della nazionalità tramite API NamSor

Per la determinazione del genere e della nazionalità associati ai nomi filtrati, è stata utilizzata l'API esterna NamSor, uno strumento di *onomastics-based AI classification* che consente di interpretare caratteristiche demografiche a partire dai nomi propri.

Lo script `namsor_request.py`, in figura 11, 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 dove era possibile del paese di origine del nome.

L'output restituito da NamSor includeva i seguenti campi principali:

name	likelyGender	genderScale	countryOrigin	probabilityCalibrated
Anna	female	0.96	IT	0.98
Luca	male	0.95	IT	0.97

Tabella 3 - Output restituito da NamSor

Per garantire una classificazione più interpretabile sono state determinate quattro classi di genere:

- **“femminile”**, se *likelyGender* = *female*;
- **“maschile”**, se *likelyGender* = *male*;
- **“ambiguo”**, per nomi con probabilità simili tra maschio e femmina (es. *Andrea*, *Ali*);
- **“indeterminato”**, per i casi in cui NamSor non ha fornito un risultato affidabile.

## CAPITOLO 4 – METODOLOGIA

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.

```

import requests
import pandas as pd
import time

# === 1. Inserisci la tua API Key di NamSor ===
API_KEY = "0934fcb11f0b62de4ea2d5f4b200e2e6"

# === 2. Carica i nomi dal file filtrato ===
with open("unique_first_names.txt", "r", encoding="utf-16") as file:
    names = [line.strip() for line in file if line.strip()]

# === 3. Prepara la lista dei risultati ===
results = []

# === 4. 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/gender/{name}",
            headers={"X-API-KEY": API_KEY}
        )
        if response.status_code == 200:
            data = response.json()
            gender = data.get("likelyGender", "unknown")
            results.append({"name": name, "gender": gender})
            print(f"[✓] {name} → {gender}")
        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

# === 5. Salva i risultati ===
df = pd.DataFrame(results)
df.to_csv("name_gender_namsor.csv", index=False)
print("File 'name_gender_namsor.csv' salvato con successo.")

```

*Figura 11 - Codice Python per la richiesta Namsor per osservare il genere associato al nome degli host*

### 4.2.4 Creazione del dataset finale

La fase successiva ha riguardato l'integrazione delle informazioni anagrafiche ottenute con i dati economici e di performance degli host.

Il file risultante dalle fasi descritte precedentemente è stato salvato in formato .csv ed in seguito è stato effettuato un processo di merge 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).

### 4.2.5 Analisi delle 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 uno script Python con Selenium, presentato in Figura 12, 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 del dataset finale definito nel paragrafo precedente tramite una operazione di merge.

```
# -*- coding: utf-8 -*-  
"""
```

```
analisi_foto_host.py
```

```
- Si aggancia a Chrome già aperto in remote-debugging (porta 9222).
```



- Legge host\_id dal CSV (es. randomized\_completo.csv).
- Ordina e rimuove duplicati.
- Tiene traccia dell'avanzamento in progress.txt.
- Apre ogni profilo host e attende TIME\_PER\_PROFILE secondi (modificabile).
- Se trova la scritta "There was an error processing this image." salta subito.

## Istruzioni:

- 1) Avvia Chrome in debug e fai login su Airbnb.
- 2) python3 analisi\_foto\_host.py
- 3) Puoi interrompere quando vuoi (Ctrl+C). Alla prossima esecuzione riparte dal punto salvato.

```
"""
```

```
import os
import sys
import time
import re
from pathlib import Path

import pandas as pd
from selenium import webdriver
from selenium.webdriver.chrome.options import Options
from selenium.webdriver.chrome.service import Service
from selenium.webdriver.common.by import By
from selenium.webdriver.support.ui import WebDriverWait
from selenium.webdriver.support import expected_conditions as EC
from webdriver_manager.chrome import ChromeDriverManager

# ===== CONFIG =====
CSV_PATH = "random_Id.csv" # metti qui il tuo file
HOST_ID_COLUMN_CANDIDATES = ["host_id", "Host_ID", "id_host", "ID"] # prova automatica
PROGRESS_PATH = "progress.txt"
REMOTE_DEBUG_PORT = 9222
TIME_PER_PROFILE = 12 # secondi di attesa per la tua annotazione manuale
WAIT_PAGE_READY = 10 # timeout per caricamento minimo della pagina
# =====

def attach_to_debug_chrome(port: int = REMOTE_DEBUG_PORT)
webdriver.Chrome:
    """Si aggancia a una istanza di Chrome già aperta con --remote-
debugging-port."""
    opts = Options()
    # Aggancio alla finestra già aperta
    opts.debugger_address = f"127.0.0.1:{port}"

    # NIENTE detach, NIENTE DesiredCapabilities: con Selenium 4 bastano le
Options.
```

```

driver = webdriver.
Chrome(service=Service(ChromeDriverManager().install()), options=opts)
return driver

def load_host_ids(csv_path: str) -> pd.Series:
    """
    Legge il CSV rilevando automaticamente il separatore.
    Trova la colonna host_id (o simile), estrae solo cifre, elimina vuoti
    e ID corti,
    restituisce lista ordinata e unica (stringhe).
    """
    if not os.path.exists(csv_path):
        raise FileNotFoundError(f"CSV non trovato: {csv_path}")

    # Rileva automaticamente il separatore (funziona con , o ;)
    df = pd.read_csv(csv_path, sep=None, engine="python", dtype=str)

    # 1) trova la colonna host_id
    host_col = None
    for c in df.columns:
        cl = c.lower()
        if "host" in cl and "id" in cl:
            host_col = c
            break
    if host_col is None:
        # fallback: prova colonne note
        for c in ["host_id", "Host_ID", "id_host", "ID"]:
            if c in df.columns:
                host_col = c
                break
    if host_col is None:
        raise ValueError(f"Non trovo una colonna host_id nel file {csv_path}. ")

    print(f"Colonne: {list(df.columns)}")

    # 2) pulizia: stringa, solo cifre
    s = df[host_col].astype(str)

    # Prendi la sequenza numerica più lunga per ogni cella (es.
    # '27355682' da '27355682,$277,01')
    s_num = s.str.extract(r'(\d+)', expand=False)

    # 3) filtra: almeno 6 cifre (gli host_id reali sono lunghi)
    s_num = s_num.dropna()
    s_num = s_num.loc[s_num.str.len() >= 6]

    # 4) unici e ordinati numericamente
    host_ids = sorted(s_num.unique(), key=lambda x: int(x))

    # (opzionale) stampa i primi ID per verifica
    print("Esempi di host_id letti:", host_ids[:10])

```

```

return pd.Series(host_ids, dtype=str)

def read_progress(path: str) -> int:
    """Ritorna l'indice da cui riprendere. Se il file non esiste, 0."""
    if not os.path.exists(path):
        return 0
    try:
        return int(Path(path).read_text().strip())
    except Exception:
        return 0

def write_progress(path: str, index: int) -> None:
    """Scrive l'indice corrente nel file di progresso."""
    Path(path).write_text(str(index))

def page_has_error_banner(driver: webdriver.Chrome) -> bool:
    """
    Prova a rilevare il banner 'There was an error processing this image.'
    (quello di Chrome in caso di immagini non processabili).
    """
    try:
        # Match su testo in inglese (come nello screenshot)
        err = driver.find_elements(
            By.XPATH,
            "//*[contains(text(), 'There was an error processing this
image')]"
        )
        return len(err) > 0
    except Exception:
        return False

def wait_minimal_ready(driver: webdriver.Chrome, timeout: int =
WAIT_PAGE_READY) -> None:
    """Aspetta che il <body> esista; poi un piccolo delay per
stabilizzarsi."""
    | WebDriverWait(driver, timeout).until(
        EC.presence_of_element_located((By.TAG_NAME, "body"))
    )
    time.sleep(0.5)

def main():
    print(f"Carico host_id da '{CSV_PATH}' ...")
    host_ids = load_host_ids(CSV_PATH)
    print(f"Trovati {len(host_ids)} host unici (ordinati).")

    start_index = read_progress(PROGRESS_PATH)
    if start_index >= len(host_ids):
        print("Hai già finito tutti gli host")
        return

    print(f"Riprendo da indice {start_index} su {len(host_ids)-1}.")

```

```

# Aggancio a Chrome già aperto in debug
driver = attach_to_debug_chrome(REMOTE_DEBUG_PORT)

try:
    for i in range(start_index, len(host_ids)):
        host_id = host_ids.iloc[i]
        url = f"https://www.airbnb.it/users/show/{host_id}"

        print(f"\n[{i+1}/{len(host_ids)}] Apro: {url}")
        driver.get(url)

        try:
            wait_minimal_ready(driver, WAIT_PAGE_READY)
        except Exception:
            print("La pagina non sembra caricarsi; continuo
comunque.")

            # anche se non è pronta, salvo il progresso e vado oltre

            # Se c'è il banner di errore immagine, salto subito
            if page_has_error_banner(driver):
                print("Immagine non processabile (banner rilevato) → passo
al prossimo.")
                write_progress(PROGRESS_PATH, i + 1)
                continue

            # Attendi il tempo per valutare manualmente la foto
            for sec in range(TIME_PER_PROFILE, 0, -1):
                print(f"    Valutazione manuale... {sec}s      ", end="\r",
flush=True)
                time.sleep(1)
            print(" " * 40, end="\r") # pulizia riga

            # Salva avanzamento
            write_progress(PROGRESS_PATH, i + 1)

except KeyboardInterrupt:
    print("\n🛑 Interrotto dall'utente. Progresso salvato.")
finally:
    # Non chiudo il browser: stai usando la sessione 'umana' in debug
    print(f"Progresso salvato in '{PROGRESS_PATH}'.")
    print("Puoi rilanciare lo script quando vuoi: riparte da dove hai
lasciato.")

if __name__ == "__main__":
    main()

```

Figura 12 – Script Python per analizzare le foto profilo degli host manualmente

### 4.3 - Misure di performance

Per studiare la possibile presenza di differenze di performance viene effettuata un'analisi delle prestazioni degli host presenti sulla piattaforma.

A questo scopo vengono considerati alcuni indicatori di performance, considerati particolarmente significativi nel settore:

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"* di K. L. Xie e Z. Mao del 2017. (57)

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"* di K. L. Xie e Z. Mao del 2017. (57)

3. **ADR (Average Daily Rate):** rappresenta il prezzo medio pagato per notte prenotata, ed è uno degli indicatori più utilizzati nel comparto dell'ospitalità.

Utilizzato nel paper: *"Gender earning gap on digital platforms: The Airbnb case"* di Ert, Fleischer e Kopolovich. (1)

Si calcola tramite la formula:

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

Dove:

- Revenue indica i ricavi complessivi generati dall'host;
- ReservationDays corrisponde al numero totale di notti effettivamente occupate.

- 4. OCC (Occupancy Rate):** misura la percentuale di occupazione dell'alloggio confrontando le notti prenotate con quelle disponibili in un determinato periodo. Valori più alti indicano una maggiore capacità di sfruttare le opportunità di guadagno. Utilizzato nel paper: *"Gender earning gap on digital platforms: The Airbnb case"* di Ert, Fleischer e Kopolovich. (1)

La formula di calcolo è:

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

dove:

- ReservationDays rappresenta le notti riservate dagli ospiti;
  - AvailableDays sono le notti in cui l'alloggio è disponibile per la prenotazione.
- 5.** Il RevPAN (Revenue per Available Night) deriva dal rapporto tra i ricavi totali e il numero complessivo di notti disponibili, senza considerare se siano state effettivamente prenotate oppure no. È un indicatore utile perché permette di capire se la gestione dell'offerta è davvero efficiente. Quando il valore del RevPAN è alto, vuol dire che l'attività riesce a sfruttare bene le proprie risorse, trovando un buon equilibrio tra prezzo, domanda e disponibilità. Al contrario, un RevPAN basso può far pensare a un utilizzo poco efficace delle notti disponibili. Magari ci sono stanze o appartamenti vuoti troppo spesso, oppure i prezzi sono troppo bassi rispetto al mercato. Utilizzato nel paper: *"How to react to a shock? Effects of Airbnb hosts' choices and market segmentation at the time of Covid-19"* di Buzzacchi et al. (58)

La formula è la seguente:

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

dove:

- ADR è la tariffa media per notte;
- OCC è la percentuale di occupazione.

### 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 venivano modificati file originali, bastava un clic per ricaricare tutto e avere le tabelle aggiornate e coerenti tra di loro.

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 istogrammi, 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.

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 descrittive**

L'analisi descrittiva elaborata in questo sottocapitolo realizzata con l'utilizzo di Excel è utile per capire la distribuzioni delle variabili del dataset in modo da comprendere meglio le caratteristiche degli host e come si comportano su Airbnb.



Vengono considerate delle variabili di tipo demografico (come il genere e la nazionalità) e quelle che descrivono il comportamento e il guadagno degli host come il tasso di risposta, il tasso di occupazione, l'ADR, il RevPAN e delle valutazioni medie.

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 genere e nazionalità

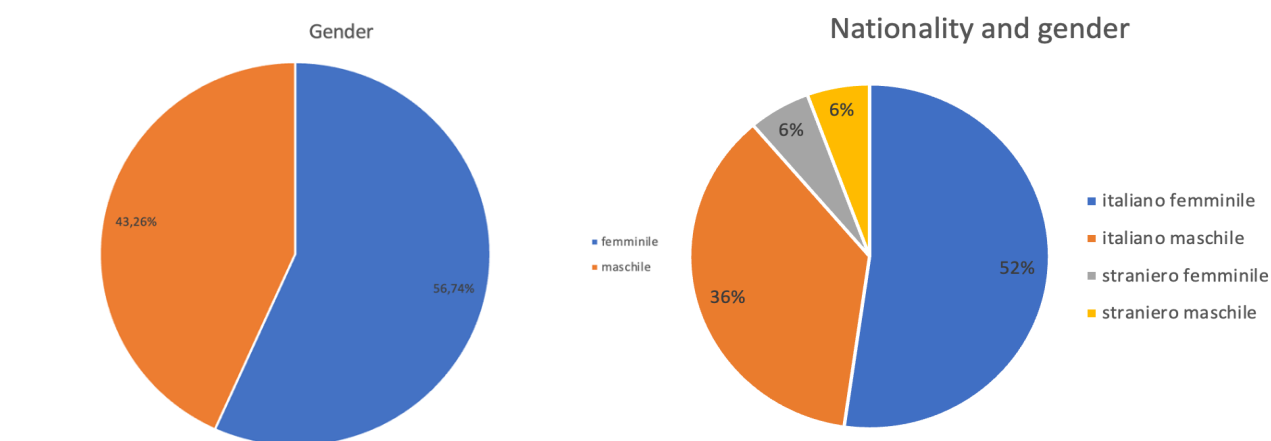


Figura 13 - Distribuzione annunci per genere e nazionalità

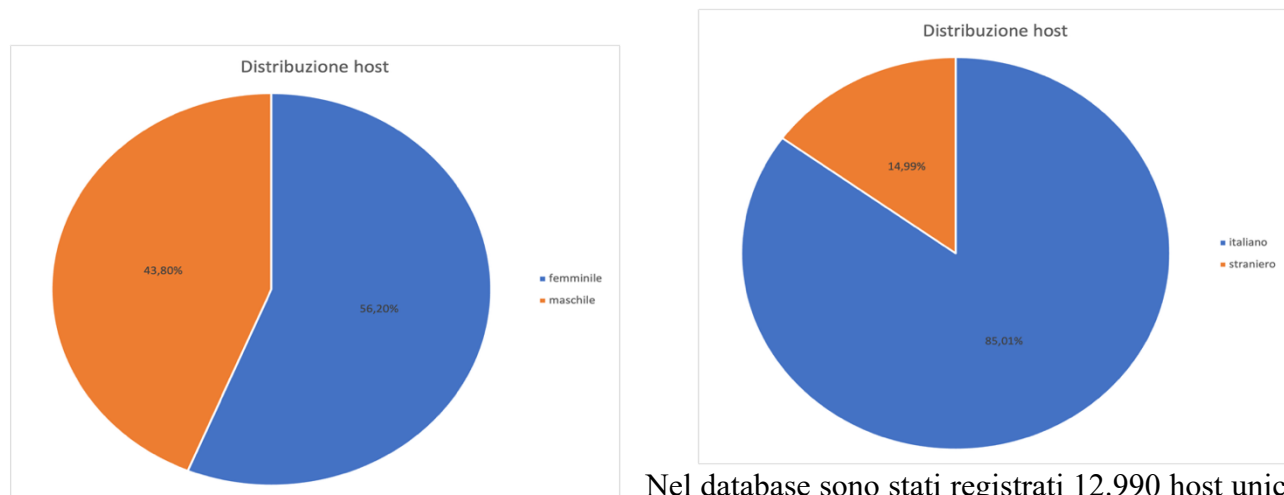
Il dataset analizzato contiene 530.960 annunci Airbnb, ognuno associato a un singolo host.

Dando un'occhiata al grafico della Figura 13 si nota subito che le donne sono più presenti degli uomini. Circa il 56,7% degli annunci è infatti gestito da host femminili, mentre il 43,3% da uomini. Questo fa pensare che le donne siano più attive sulla piattaforma, forse perché più inclini a occuparsi di attività legate all'accoglienza o alla cura dell'esperienza degli ospiti alcuni degli aspetti centrali nel mondo dell'ospitalità.

Quando poi si incrociano genere e nazionalità (come nel Grafico 2 della stessa figura) si nota che più della metà degli annunci (52%) arriva da donne italiane, seguite dagli uomini italiani con circa il 36%. Gli host stranieri, invece, rappresentano una parte più piccola del totale: solo il 6% di donne e un altro 6% di uomini.

Nel complesso, quindi, Airbnb in Italia sembra essere dominato da host donne italiane, con una presenza maschile comunque rilevante ma meno marcata. La componente straniera, per ora, rimane più marginale, anche se ben equilibrata tra uomini e donne.

#### 4.5.2 - Distribuzione degli host per genere e nazionalità



Nel database sono stati registrati 12.990 host unici

attivi su Airbnb.

*Figura 14 - Distribuzione host per genere e nazionalità*

Analizzando la distribuzione per genere (Grafico 1 - Figura 14) si può notare un leggero e costante vantaggio per le donne: 56,2% contro il 43,8% degli uomini quindi più della metà degli host è di sesso femminile. Questo potrebbe spiegarsi considerando che le donne mostrano una naturale predisposizione per lavori dove servono attenzione, empatia e cura dei dettagli. Qualità che, in un ambito come quello dell'accoglienza, fanno la differenza: bastano piccoli gesti per trasformare un soggiorno qualsiasi in un'esperienza che resta impressa.

Passando invece alla nazionalità (Grafico 2 – Figura 14) si nota che circa l'85% degli host attivi su Airbnb risulta italiano, mentre il restante 15% proviene da altri Paesi.

Questa distribuzione rispecchia bene il funzionamento del mercato degli affitti brevi in Italia, dove la gestione diretta da parte dei residenti è ancora la regola. Gli host stranieri ci sono ma restano una minoranza. Quindi si può dire che la scena italiana di Airbnb è fatta soprattutto di persone del posto che decidono di aprire le porte delle proprie case ai viaggiatori.

In sintesi, quindi, il profilo più comune dell'host è quello di una donna italiana. Un dato che è curioso e anche utile per capire meglio le differenze di performance che verranno analizzate più avanti come ad esempio il prezzo medio per notte (ADR), il ricavo per host (RevPAN) o i punteggi di valutazione lasciati dagli ospiti.

4.5.3 - Ricavi medi per genere

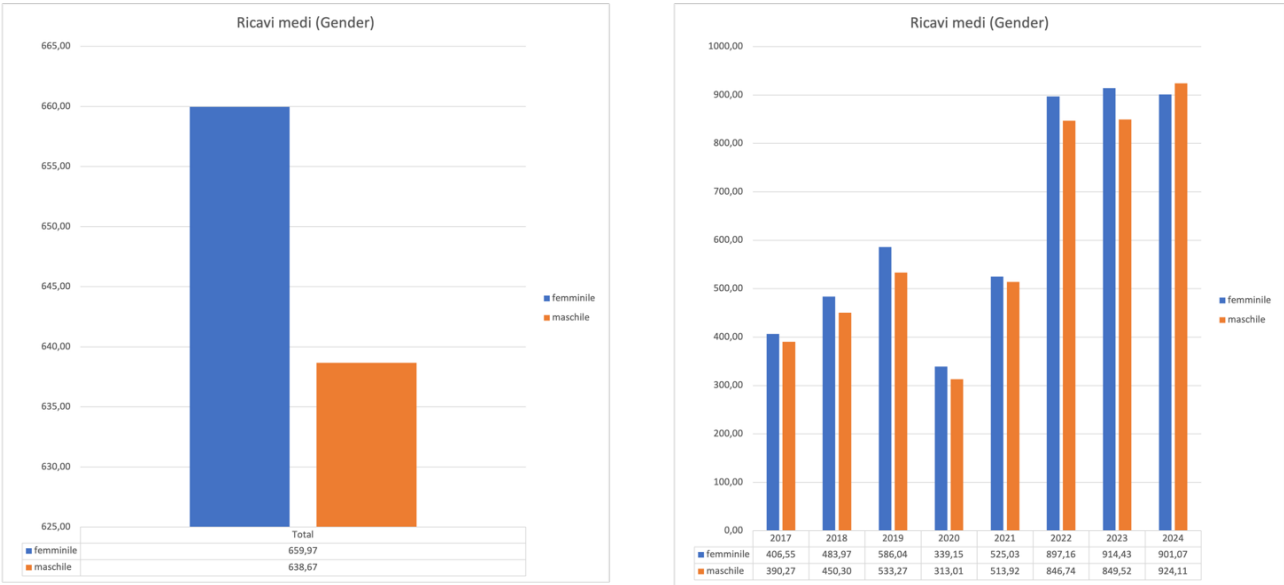


Figura 15 - Ricavi medi per genere

Dai dati sui ricavi medi annui (Tabella 1 – Figura 15) si nota che, in generale, le donne guadagnano un po’ di più degli uomini: circa 659,97 USD contro 638,67. Non è una grande differenza, ma è costante per gran parte del periodo analizzato, dal 2017 al 2024, anche se con qualche alti e bassi.

Negli anni dal 2017 al 2019, le donne registrano ricavi più elevati magari perché tendevano a curare meglio i propri annunci ad esempio con foto più accattivanti, descrizioni più precise, risposte più rapide agli ospiti, quindi un’attenzione in più che si è tradotta in guadagni leggermente superiori.

Nei due anni della pandemia (2020-2021) la differenza si è ridotta, successivamente nel 2022 però i valori diventano quasi più stabili infatti i redditi medi oggi sono molti simili tra donne e uomini.

Il t-test per campioni indipendenti conferma la significatività della differenza:

$t = 6.412, \quad p < 0.001$ , con una differenza media di 21,3 USD a favore delle donne.

Si rifiuta dunque l’ipotesi nulla di uguaglianza delle medie, concludendo che le host femminili percepiscono, nel complesso, ricavi significativamente più alti rispetto agli uomini.

4.5.4 - Rating medio per genere

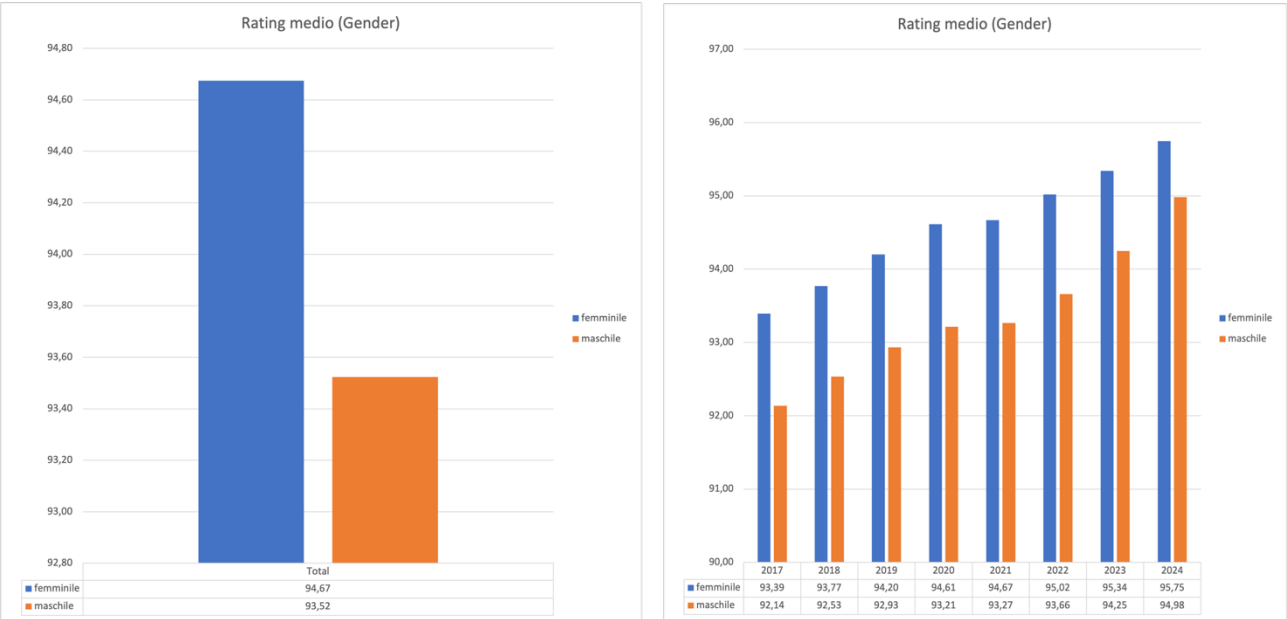


Figura 16 - Rating medio per genere

Le valutazioni degli ospiti (Figura 16 – Grafico 1) mostrano punteggi più alti nella media per le donne rispetto agli uomini: 94,67 contro 93,52. Sebbene la distanza sia minima, essa rimane stabile lungo tutto l’arco temporale 2017–2024 (Figura 16 – Grafico 2), suggerendo che le host donna mantengono costantemente una qualità del servizio superiore.

Dal 2019 in avanti si nota un miglioramento generale dei punteggi questo fa capire che la piattaforma è diventata più attenta ai bisogni e quindi magari host più attenti ai dettagli e alla soddisfazione del cliente.

Nonostante questo miglioramento collettivo, il piccolo vantaggio femminile non scompare, rimane stabile come se le donne avessero un punto in più nel comprendere i bisogni dei viaggiatori e nel creare un ambiente accogliente, su misura, dove le persone si sentono davvero a casa.

Il test t conferma la significatività statistica della differenza:

$t = 41.0796$ ,  $p < 0.001$ , con una differenza media di 1,15 punti a favore delle donne.

Pertanto, si conclude che le host femminili ottengono valutazioni significativamente più alte rispetto agli uomini.

4.5.5 - Numero medio di prenotazioni per genere

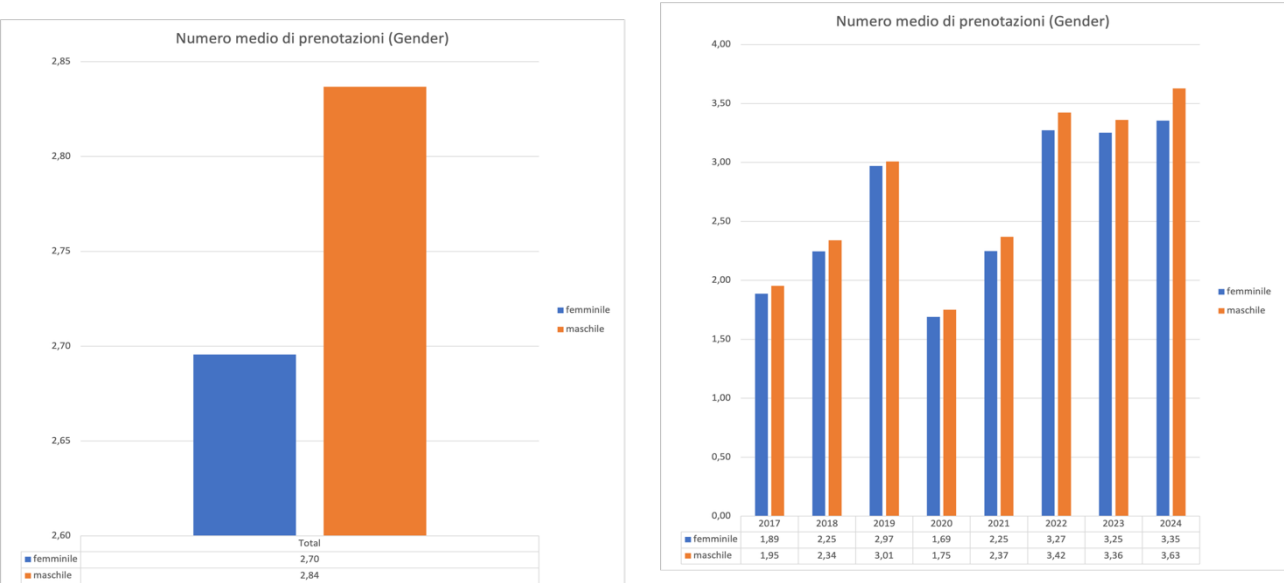


Figura 17 - Numero medio di prenotazioni per genere

Come si vede dalla Figura 17 gli host uomini ricevono in media più prenotazioni delle donne (2,84 contro 2,70). Questa tendenza è costante lungo tutto il periodo 2017–2024 e risulta molto evidente negli anni 2023–2024.

Questo risultato potrebbe riflettere differenze nella tipologia degli alloggi o nella localizzazione con una maggiore concentrazione degli host uomini in immobili più grandi o in zone a domanda più elevata.

Dall’analisi statistica si conferma che la differenza è reale:

il t-test ha restituito un valore di  $t = -12.9495$  con  $p < 0.001$ , quindi il risultato è statisticamente significativo.

Di conseguenza la media di  $-0,14$  prenotazioni a favore degli uomini non è dovuta al caso, quindi il genere sembra effettivamente avere un piccolo impatto sul volume delle prenotazioni.

Questa differenza può far intendere forse che le donne puntano di più sulla qualità e sull’esperienza del soggiorno, mentre gli uomini gestiscono strutture più orientate alla quantità o al turismo di passaggio. In ogni caso, i dati suggeriscono che gli host uomini ottengono un numero leggermente

superiore di prenotazioni rispetto alle loro colleghe, probabilmente per motivi legati al tipo di offerta o alla posizione degli alloggi.

#### 4.5.6 - Tipologia di alloggio per genere

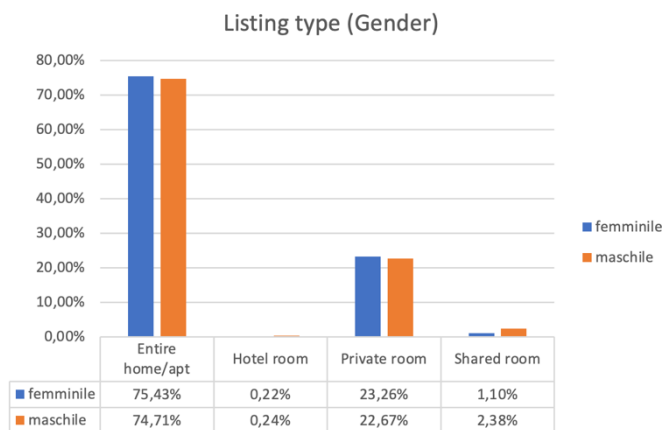


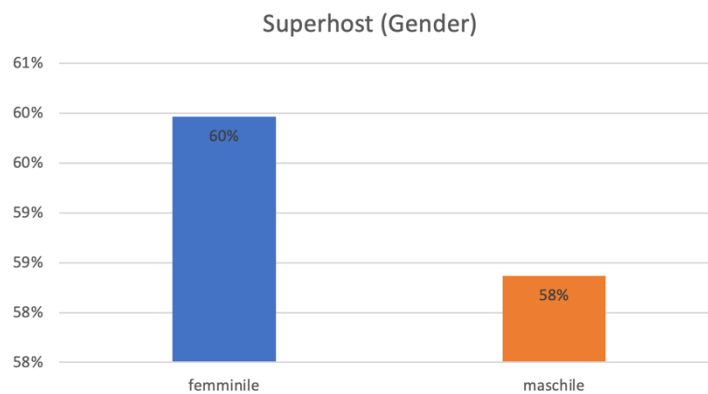
Figura 18 - tipologia di alloggio per genere

La distribuzione delle tipologie di alloggio (Figura 18) è sostanzialmente analoga per uomini e donne. Le case o gli appartamenti interi è caratterizzata da valori molto alti infatti rappresentano circa il 75% delle inserzioni (75,43% per le donne, 74,71% per gli uomini).

Dopo ci sono le stanze private (circa 23%) e in misura contenuta le stanze condivise (<2,5%) e le camere d'hotel (<0,3%).

La quasi perfetta sovrapposizione delle percentuali suggerisce che il genere non influisce sulla scelta della tipologia di alloggio. Entrambi i gruppi tendono a privilegiare soluzioni più redditizie e autonome, come gli appartamenti interi, rispetto alle camere.

#### 4.5.7 - Status di Superhost per genere



*Figura 19 – Status di superhost*

La quota di Superhost (Figura 19) mostra un lieve vantaggio per le donne: 60% contro 58% per gli uomini. Sebbene la differenza sia contenuta, indica una maggiore propensione femminile a mantenere elevati standard qualitativi e un'attenzione più marcata all'esperienza del cliente.

Il test t mostra che la differenza è molto significativa e cioè  $t = 19.8417$  con  $p < 0.001$  quindi in media, le donne hanno uno scarto positivo di circa 0,025 punti percentuali. Quindi diciamo che si rifiuta l'ipotesi nulla, il genere influenza la probabilità di ottenere lo stato di Superhost anche se in modo moderato.

Il risultato è in linea a quanto spiegato precedentemente nei capitoli di letteratura e cioè che la maggiore empatia e la maggiore preoccupazione per la soddisfazione dell'ospite sono dei tratti tipici di host donne.

#### 4.5.8 - Numero medio di foto per genere

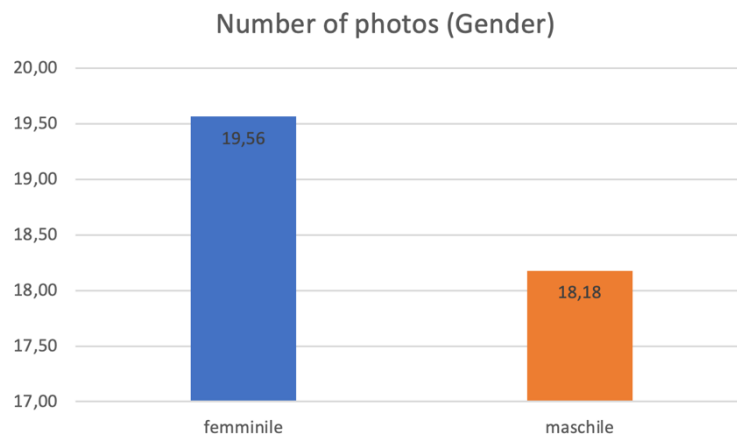


Figura 20 – Numero medio di foto per genere

Le host femminili pubblicano in media più immagini per annuncio rispetto agli uomini (19,56 contro 18,18, Figura 20).

Con il t-test otteniamo i seguenti risultati:  $t = 32.7638$  con  $p < 0.001$ , con una differenza statisticamente significativa con una media di 1,39 foto a favore delle host donne.

Questo mostra che le donne tendono a curare di più l'aspetto visivo dell'alloggio, un elemento chiave quando si tratta di catturare l'interesse dei potenziali ospiti. Questo comportamento delle host può essere pensato come una strategia per trasmettere più fiducia, più professionalità e più qualità.

#### 4.5.9 - Politica di cancellazione per genere

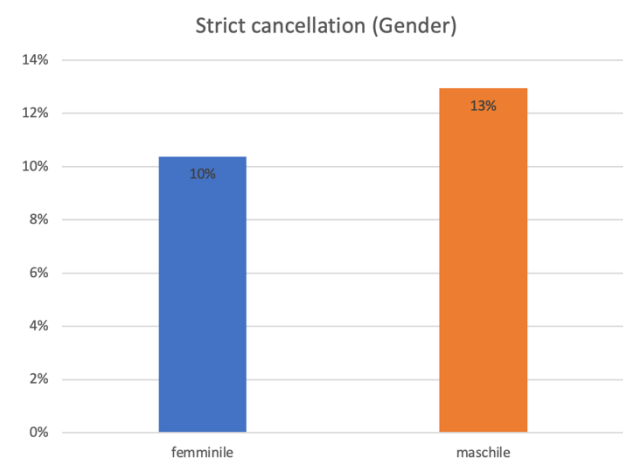


Figura 21 – Politica di cancellazione



Gli host maschili applicano più frequentemente una politica di cancellazione rigida rispetto alle donne (13% contro 10%, Figura 21). Questo comportamento riflette una maggiore tendenza maschile a preferire strategie conservative, mirate a ridurre il rischio di cancellazioni e a garantire stabilità dei ricavi.

Viene effettuato il test t che mostra quanto la differenza tra uomini e donne è statisticamente significativa con i seguenti risultati:  $t = -25.7671$ ,  $p < 0.001$ . In media, gli uomini hanno un valore leggermente più alto, circa  $-0,0257$  punti, e questo viene considerato come una maggiore tendenza a scegliere politiche di cancellazione più rigide.

Quindi le donne preferiscono degli approcci più flessibili e orientati alla relazione con l'ospite mentre gli uomini mostrano una gestione più orientata al controllo.

### 4.5.10 - ADR (Average Daily Rate) per genere

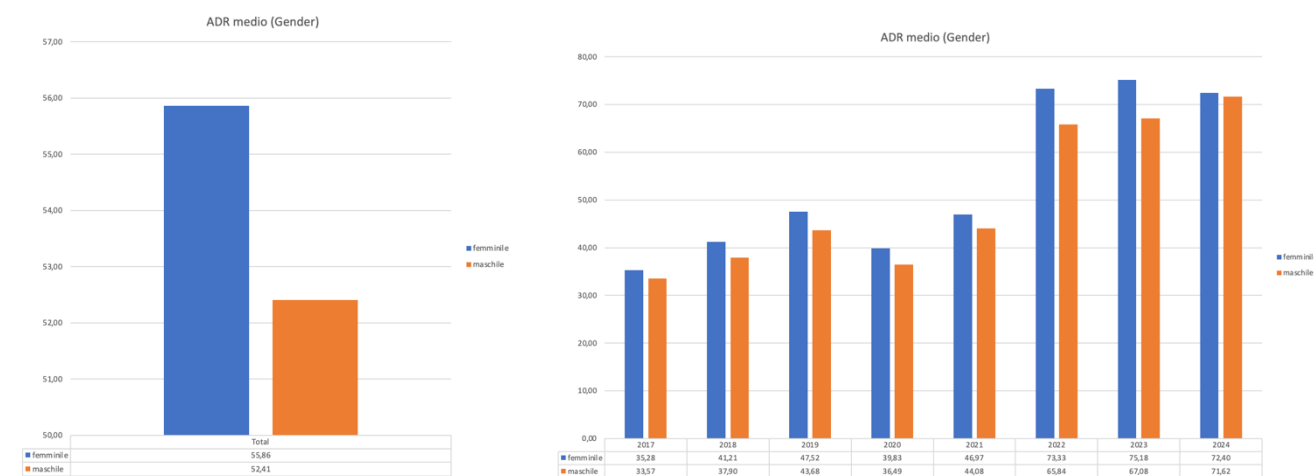


Figura 22 – ADR per genere

Dalla Figura 22 emerge una differenza significativa nel prezzo medio per notte: le donne applicano in media un ADR di 55,86, contro 52,41 per gli uomini.

Dal t-test troviamo i seguenti risultati:

- $t = 15.3727$  con  $p < 0,001$ ;
- Una distanza media di 3,45 a favore delle host donne.

Rifiutiamo quindi l’ipotesi di uguaglianza e di conseguenza quindi le donne fissano dei prezzi più alti nella media.

E’ possibile dare un’idea per questi risultati e cioè che le donne tendono a valorizzare maggiormente l’esperienza offerta, mentre gli uomini adottano strategie di pricing più competitive per massimizzare l’occupazione.

4.5.11 - Occupancy Rate (OCC) per genere

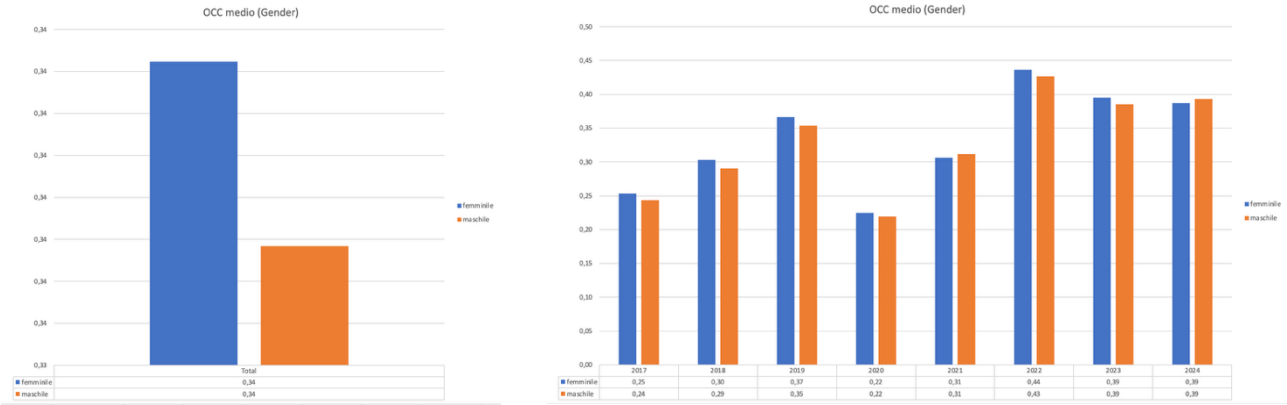


Figura 23 – OCC per genere

L’analisi dell’Occupancy Rate mostra che le donne presentano un tasso medio leggermente più alto (0,3412) rispetto agli uomini (0,3368, Figura 23 - Grafico 1). La differenza di 0,0044 punti, è statisticamente significativa ( $t = 3.8445$ ,  $p < 0.001$ ). Le host donne riescono a mantenere un tasso di occupazione più alto in media, questo grazie magari a una gestione più curata nella comunicazione e nella costruzione della fiducia tramite le recensioni. L’andamento nel tempo (Figura 23 – Grafico 2) mostra valori leggermente superiori per le donne in quasi tutti gli anni analizzati.

## 4.5.12 - RevPAN per genere

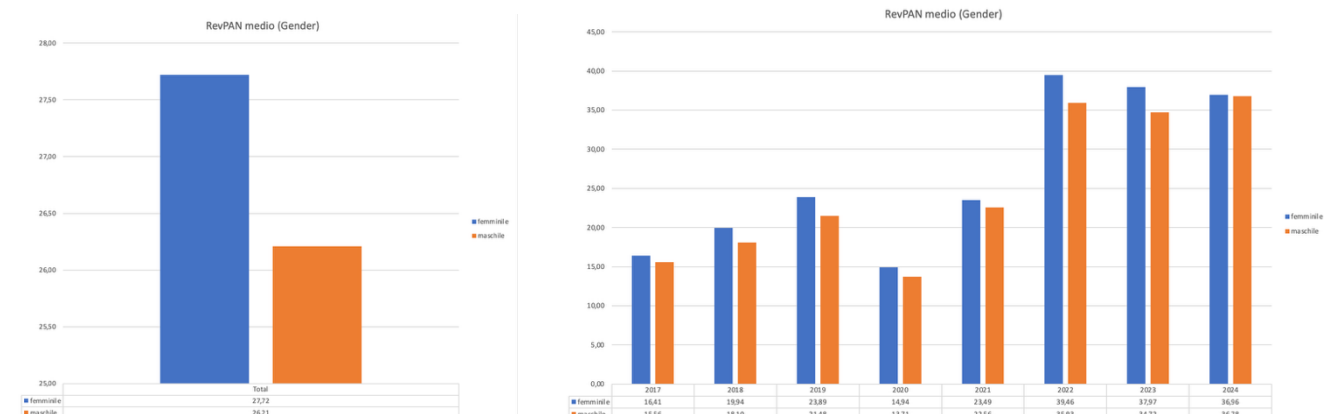


Figura 24 – RevPAN per genere

Infine, il ricavo medio per annuncio (RevPAN) evidenzia un vantaggio per le donne: 27,72 rispetto ai 26,21 degli uomini (Figura 24 – Grafico 1).

Per valutare se questa minima differenza è significativa viene effettuato il test t che rileva i seguenti risultati:

- $t = 10.5693$
- Distanza media di 1,51

Di conseguenza la differenza è significativa quindi si rifiuta l'ipotesi nulla di uguaglianza delle medie.

Dalla Figura 24 – Grafico 2 si nota invece più l'andamento nel tempo che mostra come il divario rimane costante e cioè che le donne ottengono un RevPAN più alto. Questo risultato può essere interpretato come la combinazione di prezzi più alti (ADR) e buoni livelli di occupazione (OCC), che si traducono in ricavi complessivamente superiori.

### 4.5.13 – Distribuzione delle caratteristiche visive

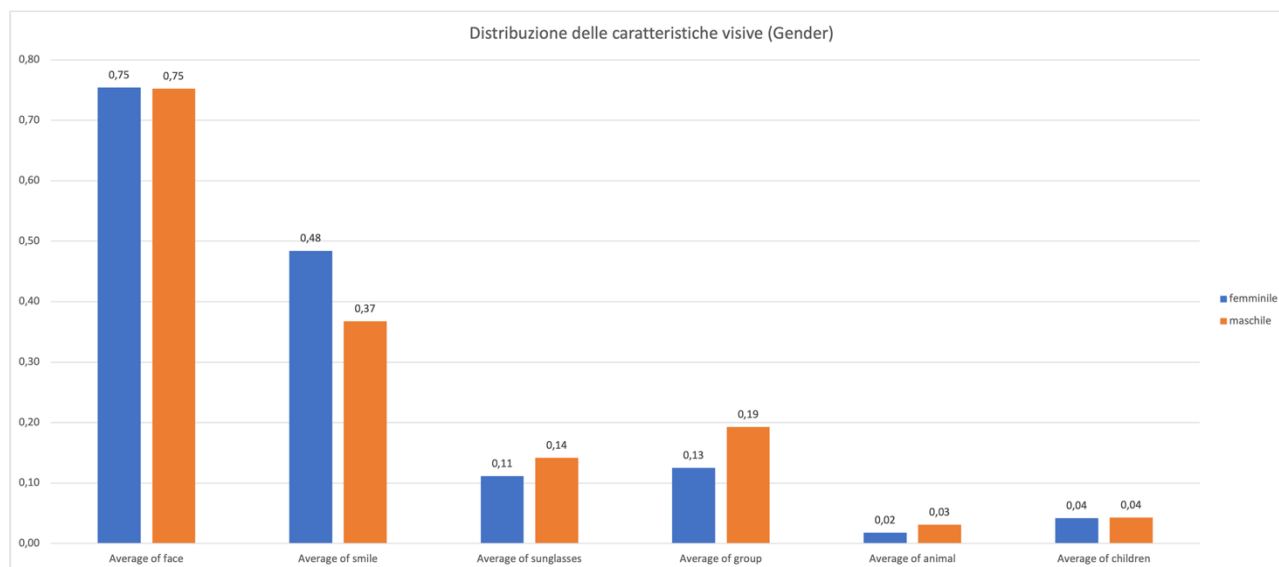


Figura 25 – Distribuzione delle caratteristiche visive

L'analisi delle caratteristiche visive dei profili Airbnb, suddivise per genere, rivela alcune differenze interessanti nel modo in cui uomini e donne scelgono di presentarsi sulla piattaforma.

Prima di tutto, la presenza del volto nelle foto è praticamente identica: sia uomini che donne mostrano il viso in circa il 75% dei casi (0,75). Questo fa pensare che entrambi i gruppi vogliano trasmettere fiducia e autenticità, qualità fondamentali in un contesto come quello dell'ospitalità online, dove la prima impressione spesso conta più di ogni descrizione.

Le differenze emergono invece quando si guarda alla frequenza del sorriso. Le host donne sorridono più spesso, con un valore medio di 0,48, contro 0,37 degli uomini. È un piccolo scarto ma comunque significativo infatti il sorriso sembra essere un modo per comunicare calore, gentilezza e disponibilità, tutti tratti che aiutano a far sentire l'ospite il benvenuto ancora prima della prenotazione. Questo risultato è coerente con altri studi che mostrano come, nei contesti digitali di ospitalità, le donne tendano a usare il linguaggio visivo in modo più accogliente e relazionale. (9)

Gli uomini, invece, si distinguono per alcuni dettagli più difensivi o distaccati: indossano più spesso occhiali da sole (0,14 contro 0,11) e compaiono più di frequente in foto di gruppo (0,19 contro 0,13). Queste scelte possono indicare una preferenza per immagini meno personali, magari più informali o orientate a una rappresentazione collettiva, piuttosto che intima.

Per il resto, la presenza di animali (0,02 donne, 0,03 uomini) e bambini (0,04 per entrambi) è rara e sostanzialmente simile, quindi non incide in modo significativo sulla percezione complessiva.

In sintesi, entrambi i generi scelgono di mostrarsi apertamente, ma con stili leggermente diversi: le donne puntano su espressività e accoglienza, mentre gli uomini tendono a rappresentarsi in modo più sobrio e sociale.

### **4.5.14 – Andamento delle metriche chiave**

La Tabella 26 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.

	Year	Mean	Std. Dev.	Min	Max
<b>Number of Reservations</b>					
	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
<b>Revenues per Active Night (RevPAN)</b>					
	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
<b>Overall Rating</b>					
	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

Figura 26 – Statistiche descrittive per anno

## CAPITOLO 5 - RISULTATI

In questo capitolo vengono presentati i risultati delle analisi econometriche, utili per capire quali elementi influenzano di più le performance economiche e la reputazione degli host su Airbnb. E' stato scelto di usare modelli di regressione logaritmica invece di quelli lineari, perché consentono di interpretare i coefficienti in termini percentuali. È un approccio più intuitivo e, allo stesso tempo, riduce l'impatto dei valori estremi — quelli troppo alti o troppo bassi — che potrebbero falsare i risultati.

### 5.1 Analisi dei modelli logaritmici – Diretti per genere e nazionalità.

Tabella 4 – Modelli logaritmici – Diretti per genere e nazionalità

	(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 <sup>2</sup>	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 4 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 ( $\ln(\text{Prenotazioni})$ ), tariffa media giornaliera ( $\ln(\text{ADR})$ ), tasso di occupazione (OCC), ricavo medio per annuncio ( $\ln(\text{RevPAN})$ ) e valutazione media ( $\ln(\text{Rating})$ ).

Le variabili che vengono considerate per lo studio vanno da caratteristiche di tipo individuale dell’host come ad esempio genere e nazionalità a quelle strutturali dell’alloggio quali per esempio il numero di stanze e la capienza. Ci sono anche variabili legate alla reputazione e alla fiducia (es. stato di Superhost, attivazione della prenotazione immediata, qualità dei servizi e delle immagini) e variabili di controllo che tengono conto della flessibilità dell’annuncio (es. politica di cancellazione, soggiorno minimo).

In tutti i modelli sono stati inseriti effetti fissi per anno e mese, così da tenere sotto controllo le influenze legate alle stagioni o ai cambiamenti nel tempo. Per avere delle stime più affidabili e robuste si sono corretti gli errori standard con eteroschedasticità.

I valori di  $R^2$  variano tra 0.064 e 0.383, una situazione del tutto comune.

### **Effetti delle caratteristiche demografiche dell’host**

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.

Per quanto riguarda il genere, la variabile “Uomo” mostra un andamento quasi inverso rispetto a quello delle host donne. Gli host maschili, infatti, ricevono in media più prenotazioni ( $+0.042^*$ ), ma tendono a fissare prezzi un po’ più bassi ( $-0.028$ ). Nonostante questo, riescono comunque a mantenere un tasso di occupazione leggermente positivo ( $+0.003$ ).



Guardando però ai ricavi complessivi l'effetto è negativo ( $-0.030$ ) e anche sul punteggio medio ottenuto dagli ospiti si nota una piccola diminuzione ( $-0.011$ )\*\*.

Questi risultati possono suggerire differenze nel modo in cui uomini e donne gestiscono l'attività o nel modo in cui i clienti percepiscono il servizio. Le host donne, forse, riescono a costruire relazioni più solide con gli ospiti o a curare meglio i dettagli, e questo si riflette in una reputazione un po' più alta nel tempo.

### **Effetti delle caratteristiche reputazionali e di fiducia**

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.

Essere Superhost è, di fatto, un segnale di reputazione molto potente. I dati mostrano chiaramente che questo status ha un effetto positivo e rilevante su diversi aspetti: più prenotazioni ( $+0.192$ \*), maggiore tasso di occupazione ( $+0.155$ \*), ricavi più alti ( $+0.120$ \*\*\*) e valutazioni migliori ( $+0.034$ \*\*\*).

L'unico aspetto un po' contrario riguarda le tariffe che tendono a essere leggermente più basse ( $-0.079$ \*)\*\* e ciò può essere inteso in modo che chi è Superhost sembra puntare più sulla fiducia e sulla soddisfazione degli ospiti che sul guadagno per singola notte. In pratica, preferiscono offrire un'esperienza positiva e costruire una reputazione solida, piuttosto che alzare troppo i prezzi.

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.

### **Effetti delle caratteristiche strutturali dell'alloggio**

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.

### **Effetti delle caratteristiche qualitative e visive**

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 (cioè i comfort di buona qualità ma essenziali) sono legate a migliori performance: +0.035\*\*\* sull'ADR, +0.034\*\*\* sull'OCC e +0.096\*\*\* sul RevPAN, quindi i servizi che rendono il soggiorno più comodo e pratico contano davvero per la soddisfazione degli ospiti.

Al contrario, le luxury amenities, quelle più costose o decorative, sembrano avere un effetto opposto. I dati mostrano valori negativi su tutte le metriche (−0.048\*\*\* per ADR, −0.026\*\*\* per OCC e −0.125\*\*\* per RevPAN). Questo suggerisce che il lusso, da solo, non basta per migliorare le performance e può persino ridurre l'efficienza economica 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.

### **Discussione complessiva**

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.

## 5.2 Analisi dei modelli logaritmici – Diretti e di Interazione per genere e nazionalità

Tabella 5 - Modelli logaritmici - Diretti e interaction per genere e nazionalità

	(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 <sup>2</sup>	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 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 (M1), che riflette quanto spesso un alloggio viene scelto;
- il prezzo medio giornaliero (ADR) o (M2), che rappresenta quanto si paga in media per notte;
- il tasso di occupazione (OCC) (M3), cioè la percentuale di notti effettivamente prenotate;
- il ricavo per annuncio (RevPAN) (M4), che misura il guadagno medio per inserzione;
- e la valutazione media (Rating) (M5), ovvero il punteggio medio assegnato dagli ospiti.

Ogni modello include errori standard robusti che vengono indicati tra parentesi e 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,  $p < 0.01$ ), ma positivo per l’ADR (+0.045,  $p < 0.01$ ), per il tasso di occupazione (+0.013,  $p < 0.01$ ), per il RevPAN (+0.094,  $p < 0.01$ ) e per il rating medio (+0.018,  $p < 0.01$ ). Gli host italiani quindi 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* mostra invece un effetto negativo e molto significativo su quasi tutte le metriche di performance. Gli host uomini hanno più prenotazioni (+0.034,  $p < 0.01$ ), ma allo stesso tempo prezzi medi più bassi (-0.036 sull’ADR,  $p < 0.01$ ), minore occupazione (-0.012,  $p < 0.01$ ), ricavi inferiori (-0.029,  $p < 0.01$ ) e valutazioni peggiori (-0.017,  $p < 0.01$ ). 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  $\times$  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 e reputazione**

Tra le variabili di controllo, la funzione *Instant Book* 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 ( $p < 0.01$ ).

Da questo risultato si può capire che la possibilità di prenotare subito senza attendere la conferma dell'host rende l'annuncio più attraente per gli utenti e va a favorire sia la domanda che i ricavi complessivi. Allo stesso tempo però si può notare una leggera diminuzione della soddisfazione media.

Probabilmente ciò accade perché, con l'Instant Book 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 anche importante infatti 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,  $p < 0.01$ ). Quindi essere Superhost migliora la fiducia e la domanda, ma spinge a una politica di prezzo più competitiva.

### **Struttura dell'alloggio e qualità percepita**

Le caratteristiche strutturali mostrano effetti misti ma coerenti.

Il *numero di bagni* mostra un'associazione negativa con le prenotazioni (-0.043,  $p < 0.01$ ), 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,  $p < 0.01$ ), mentre l'ADR cresce di +0.063 ( $p < 0.01$ ). 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,  $p < 0.01$ ), 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,  $p < 0.01$ ) e sul RevPAN (-0.125,  $p < 0.01$ ). 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. Per quanto riguarda invece gli amenities di qualità (ovvero quelli legati alla praticità e al comfort) mostrano effetti positivi e significativi: +0.035 sull'ADR e +0.096 sul RevPAN (entrambi  $p < 0.01$ ). Questo indica che gli ospiti apprezzano di più un ambiente accogliente, ben curato e pensato per farli sentire a proprio agio, piuttosto che dettagli lussuosi ma poco utili.

### **Tipologia di host e modalità di offerta**

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. Questo porta al fatto 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 magari per minore attenzione individuale.

## CAPITOLO 5 - RISULTATI

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 e significatività globale

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).

Dall'evidenza empirica quindi possiamo notare un divario di genere a sfavore degli uomini in termini di prezzo, ricavi e valutazioni, ma con effetti moderati per gli host italiani dove il legame tra nazionalità e performance sembra 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 nella pagina seguente 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).

```
capture drop uomo_italiano
gen byte uomo_italiano = uomo*italiano
label var uomo_italiano "Uomo × Italiano"
eststo clear

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)
estadd local YearFE "YES"
estadd local MonthFE "YES"

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)
estadd local YearFE "YES"
estadd local MonthFE "YES"

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)
estadd local YearFE "YES"
estadd local MonthFE "YES"

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)
estadd local YearFE "YES"
estadd local MonthFE "YES"

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)
estadd local YearFE "YES"
estadd local MonthFE "YES"

esttab, b(%9.3f) se(%9.3f) brackets star(* 0.10 ** 0.05 *** 0.01) label
compress nogaps 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) 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) coelabels(italiano "Italiano") uomo
"Uomo " uomo_italiano "Uomo × Italiano") mtitles("M1: ln(Prenotazioni)"
"M2: ln(ADR)" "M3: OCC" "M4: ln(RevPAN)" "M5: ln(Rating)") stats(YearFE
MonthFE r2 N, fmt(%9s %9s 3 0) labels("Year FE" "Month FE" "R^2" "N"))
```

### 5.3 Analisi dei modelli logaritmici – Diretti per i segnali visivi delle foto profilo degli host.

Tabella 6 - modelli logaritmici – Diretti per i segnali visivi delle foto profilo degli 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 <sup>2</sup>	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 6 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 sono 5 come nei modelli precedenti: 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 hanno un impatto negativo sui risultati e in particolare -0.221 sulle prenotazioni, -0.074 sul tasso di occupazione e -0.098 sul RevPAN. Quando nelle immagini compaiono più persone per l'utente diventa meno chiaro chi sia davvero l'host e questo può influenzare negativamente la sua intenzione di 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.

### **Caratteristiche dell'annuncio e reputazione dell'host**

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 di Superhost mostra un legame con delle performance più alte in tutte le principali metriche: +0.402 sulle prenotazioni e +0.301 sui ricavi. Questo esprime quanto la fiducia sia importante nel contesto della sharing economy. Gli ospiti, infatti, tendono a scegliere chi offre garanzie di affidabilità e qualità costante, e il badge di Superhost diventa un segnale forte in questa direzione. Il response rate, invece, ha un effetto positivo ma piuttosto limitato quindi rispondere velocemente ai messaggi aiuta ma non troppo.

La variabile Multiproperty invece ha effetti negativi sul tasso di occupazione (-0.039) e sul rating (-0.024) e ciò suggerisce che la gestione di più proprietà può diminuire 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 e robustezza dei modelli**

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.

## 5.4 Analisi dei modelli logaritmici – Diretti e di Interazione per i segnali visivi delle foto profilo degli host con la variabile Uomo.

Tabella 7 – Modelli logaritmici Diretti e di interazione per i segnali visivi nelle foto profilo degli host con Uomo

	(1) M1: ln(Prenotazioni)	(2) M2: ln(ADR)	(3) M3: OCC	(4) M4: ln(RevPAN)	(5) M5: ln(Rating)
Uomo	0.179*** [0.031]	-0.238*** [0.024]	0.083*** [0.014]	-0.057 [0.044]	-0.012*** [0.004]
Face	-0.040 [0.027]	-0.082*** [0.020]	-0.007 [0.012]	-0.048 [0.039]	0.008*** [0.003]
Smile	0.069*** [0.022]	-0.048*** [0.014]	0.067*** [0.010]	0.142*** [0.028]	0.005*** [0.002]
Sunglasses	0.095*** [0.032]	0.092*** [0.021]	0.024* [0.014]	0.120*** [0.038]	0.016*** [0.002]
Group	0.138*** [0.032]	-0.070*** [0.020]	0.094*** [0.016]	0.068* [0.039]	0.033*** [0.002]
Children	-0.149*** [0.049]	0.023 [0.033]	-0.079*** [0.024]	-0.066 [0.066]	-0.016*** [0.003]
Animal	0.253*** [0.069]	-0.248*** [0.052]	0.088*** [0.032]	-0.138 [0.093]	-0.033*** [0.006]
Face × Uomo	-0.206*** [0.042]	0.276*** [0.032]	-0.071*** [0.018]	0.175*** [0.059]	0.013*** [0.005]
Smile × Uomo	0.164*** [0.037]	0.019 [0.025]	0.006 [0.016]	-0.039 [0.046]	-0.004 [0.004]
Sunglasses × Uomo	-0.123** [0.049]	-0.137*** [0.033]	-0.037* [0.020]	-0.248*** [0.062]	0.013*** [0.004]
Group × Uomo	-0.695*** [0.046]	0.326*** [0.034]	-0.326*** [0.021]	-0.455*** [0.065]	-0.061*** [0.005]
Children × Uomo	0.933*** [0.083]	-0.054 [0.052]	0.415*** [0.038]	0.899*** [0.107]	0.061*** [0.008]
Animal × Uomo	-0.381*** [0.088]	0.338*** [0.077]	-0.136*** [0.042]	0.059 [0.140]	0.023 [0.015]
Instantbook	0.138*** [0.019]	0.042*** [0.012]	0.062*** [0.008]	0.290*** [0.024]	-0.003 [0.002]
Response rate	0.005*** [0.000]	0.000 [0.000]	0.002*** [0.000]	0.005*** [0.001]	0.000*** [0.000]
Superhost	0.363*** [0.021]	-0.012 [0.012]	0.168*** [0.009]	0.230*** [0.023]	0.041*** [0.001]
Multiproperty host	-0.055*** [0.015]	0.080*** [0.011]	-0.027*** [0.007]	0.008 [0.020]	-0.023*** [0.001]
Bathrooms	-0.185*** [0.017]	0.217*** [0.017]	-0.054*** [0.008]	0.137*** [0.030]	0.021*** [0.002]
Bedrooms	-0.160*** [0.016]	0.032*** [0.012]	-0.056*** [0.007]	-0.055** [0.023]	0.010*** [0.002]
Number of photos	0.011*** [0.001]	0.002*** [0.000]	0.003*** [0.000]	0.004*** [0.001]	-0.000*** [0.000]
Accommodates	0.076*** [0.006]	0.076*** [0.005]	0.020*** [0.003]	0.109*** [0.009]	-0.003*** [0.001]
Luxury amenities	-0.019 [0.020]	0.042*** [0.013]	-0.006 [0.009]	0.035 [0.026]	0.012*** [0.002]
Quality amenities	0.066** [0.032]	0.293*** [0.020]	0.027** [0.013]	0.316*** [0.052]	0.015*** [0.005]
Strict cancellation	0.022 [0.023]	-0.070*** [0.017]	0.022** [0.010]	-0.017 [0.036]	-0.025*** [0.003]
Shared room	-0.400*** [0.047]	-0.749*** [0.044]	-0.182*** [0.019]	-1.135*** [0.119]	-0.078*** [0.014]
Constant	0.031 [0.048]	3.185*** [0.042]	-0.025 [0.020]	1.464*** [0.089]	4.476*** [0.007]
Year FE	YES	YES	YES	YES	YES
Month FE	YES	YES	YES	YES	YES
R <sup>2</sup>	0.230	0.412	0.196	0.283	0.214
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 7 riporta i risultati di cinque modelli di regressione log-lineare stimati con effetti fissi per anno e mese, che analizzano l'impatto delle caratteristiche visive del profilo host e delle variabili demografiche sulle performance economiche e reputazionali degli annunci Airbnb.

Le variabili dipendenti 5 come prima: *numero di prenotazioni (M1)*, *prezzo medio giornaliero (ADR) (M2)*, *tasso di occupazione (M3)*, *ricavo per notte disponibile (M4)* e *valutazione media (M5)*.

Tutti i modelli includono errori standard robusti, riportati tra parentesi, e controllano per variabili strutturali e di reputazione che possono influenzare le performance.

### **Effetto del genere e interazioni visive**

La variabile *Uomo* mostra una relazione positiva con il numero di prenotazioni ( $\beta = 0.179$ ,  $p < 0.01$ ). In pratica, se il resto delle condizioni rimangono uguali, gli host di sesso maschile ricevono mediamente più prenotazioni rispetto alle donne. D'altra parte, il quadro cambia quando si guarda ad altri aspetti. Lo stesso coefficiente diventa negativo, e anche significativo, nei modelli che riguardano il prezzo medio giornaliero ( $\beta = -0.238$ ,  $p < 0.01$ ) e il punteggio delle recensioni ( $\beta = -0.012$ ,  $p < 0.01$ ). Questo vuol dire che gli uomini, pur ricevendo più prenotazioni, tendono a proporre prezzi un po' più bassi. Inoltre, le loro valutazioni sono leggermente inferiori rispetto a quelle delle donne.

In sintesi quindi i profili maschili sembrano puntare più sulla quantità che sulla qualità o sul guadagno per singola prenotazione. Non è un grande risultato, ma abbastanza chiaro da mostrare una differenza di comportamento tra host uomini e donne secondo quanto descritto.

Dalle interazioni tra genere e caratteristiche visive vengono fuori delle dinamiche più complesse del previsto. L'effetto dell'interazione *Uomo*  $\times$  *Sunglasses* è negativo e molto significativo in quasi tutti i modelli (M1–M4). In parole semplici, quando un profilo maschile mostra una foto con gli occhiali da sole, tende a ottenere risultati peggiori. Probabilmente il motivo è che gli occhiali coprono parte del viso, rendendo più difficile percepire espressioni e quindi riducendo la fiducia che gli utenti provano verso l'host. Al contrario con l'interazione *Uomo*  $\times$  *Children* si ottiene un effetto positivo sia sul tasso di occupazione ( $\beta = 0.415$ ,  $p < 0.01$ ) che sul rating ( $\beta = 0.061$ ,  $p < 0.01$ ). Mostrarsi in un contesto familiare sembra ammorbidire la penalità legata al genere, trasmettendo un senso di calore, accoglienza e affidabilità. In pratica, gli uomini che appaiono in foto con bambini vengono percepiti



come più rassicuranti. Da discutere è anche l'interazione *Uomo*  $\times$  *Group* che risulta fortemente negativa ( $\beta = -0.695$ ,  $p < 0.01$  nel modello M1 e  $\beta = -0.455$ ,  $p < 0.01$  in M4). Le immagini di uomini in gruppo risultano meno efficaci, forse perché sembrano meno personali o meno professionali. In quelle foto, il volto dell'host si perde un po', e chi guarda può faticare a capire chi sia realmente la persona che offre l'alloggio.

I profili degli host uomini con volto visibile e sorridente hanno migliori performance mentre le immagini che riducono la capacità di riconoscere l'identità dell'host (gruppo o occhiali da sole) sono penalizzate.

### **Impatto delle caratteristiche visive indipendenti**

Oltre al genere, le caratteristiche visive della foto profilo influenzano in modo rilevante le performance economiche.

La presenza di occhiali da sole (*sunglasses*) ha un effetto positivo e significativo su diverse metriche: tasso di occupazione (+0.024,  $p < 0.01$ ) e *RevPAN* (+0.12,  $p < 0.01$ ), suggerendo che in generale l'immagine "professionale" o curata dell'host possa favorire la domanda, se non associata al genere maschile.

La presenza di un gruppo (*Group*) mostra un effetto negativo e significativo sull'ADR ( $\beta = -0.070$ ,  $p < 0.01$ ) e un impatto positivo sulla *RevPAN* ( $\beta = 0.068$ ,  $p < 0.05$ ), suggerendo che, se da un lato la presenza di più persone nella foto può ridurre la percezione di professionalità e quindi il prezzo medio, dall'altro può favorire un leggero incremento della performance complessiva in termini di ricavi per disponibilità.

Quando ci sono dei bambini nelle immagini e cioè quindi parliamo della variabile *Children* è legata a una diminuzione del tasso di occupazione (*OCC*,  $\beta = -0.079$ ,  $p < 0.01$ ) e a una diminuzione maggiore delle prenotazioni ( $\beta = -0.149$ ,  $p < 0.01$ ). Da questo si capisce che quando una foto mostra dei bambini l'annuncio potrebbe essere percepito come meno professionale. Questo tipo di immagine può trasmettere un'idea più familiare o privata, riducendo la fiducia di chi cerca un alloggio per motivi turistici o di lavoro. Di conseguenza, gli utenti sono meno propensi a prenotare, e le performance complessive dell'annuncio si abbassano.

La presenza di animali (Animal) risulta invece positivamente correlata all'occupazione (OCC,  $\beta = 0.088$ ,  $p < 0.01$ ), ma negativamente associata al rating ( $\beta = -0.033$ ,  $p < 0.01$ ). Questo risultato suggerisce che, se da un lato gli animali nelle immagini possono aumentare l'attrattività e la probabilità di prenotazione, dall'altro possono ridurre lievemente la valutazione complessiva, forse a causa di una percezione di minore professionalità o pulizia.

Nel complesso, le evidenze mostrano che le foto che trasmettono fiducia, autenticità e calore umano (volto visibile, sorriso, presenza di animali) migliorano la performance dell'annuncio, mentre le immagini impersonali o di gruppo la riducono.

### **Caratteristiche operative e reputazionali dell'annuncio**

Le variabili operative e quelle legate alla reputazione mostrano risultati in linea con quanto emerso negli studi precedenti.

Tra tutti i fattori analizzati, l'opzione Instant Book si distingue per l'impatto fortemente positivo. Quando gli ospiti possono prenotare subito, senza attendere conferme, le performance migliorano quasi ovunque: le prenotazioni aumentano di +0.138, il tasso di occupazione cresce di +0.062, e anche il ricavo per annuncio (RevPAN) sale di +0.290, tutti con valori statisticamente significativi. In sostanza, la possibilità di prenotare al volo riduce l'incertezza e spinge più persone a completare la prenotazione.

Anche altri aspetti legati alla reputazione e alla reattività degli host hanno un ruolo importante. Un tasso di risposta più alto e lo status di Superhost migliorano le prestazioni, soprattutto per quanto riguarda il tasso di occupazione (+0.168) e il punteggio medio delle recensioni (+0.041) riferito a Superhost. È una conferma di quanto la fiducia conta nel mondo degli affitti brevi: chi risponde in fretta e mantiene un buon livello di servizio viene premiato con più prenotazioni e valutazioni migliori.

### **Struttura dell'alloggio e qualità percepita**

Le caratteristiche strutturali dell'annuncio (come, ad esempio, il numero di bagni, camere da letto o anche la quantità di foto caricate) risultano strettamente legate sia al prezzo che ai ricavi. In sostanza

più un annuncio appare completo e dettagliato più cresce la disponibilità degli ospiti a pagare. Non si tratta solo di dimensioni o comfort, ma anche di quanto l'annuncio riesce a trasmettere chiarezza e professionalità.

La variabile *quality amenities* mostra un effetto positivo e molto significativo nel RevPAN con un aumento di +0,316 e rating di + 0.015. Questo mostra che gli host che offrono maggiore qualità (come ad esempio un alloggio ben curato) ottengono maggiori guadagni.

Per quanto riguarda invece le stanze condivise esse rappresentano una caratteristica non positiva. Gli effetti negativi si notano su tutte le metriche economiche: l'ADR cala di -0.749, mentre il RevPAN scende di -1.135, entrambi con elevata significatività statistica. Questo conferma che, sebbene le stanze condivise possano attrarre un certo tipo di viaggiatore, in termini di redditività penalizzano fortemente le performance complessive.

### **Capacità esplicativa e significatività complessiva**

I valori di  $R^2$  variano tra 0.196 e 0.412, indicando una buona capacità esplicativa dei modelli, considerando la natura cross-section dei dati. Il miglior fit si osserva nei modelli relativi al prezzo (M2) e al ricavo (M4), dove le caratteristiche strutturali e reputazionali degli annunci spiegano una quota rilevante della varianza. Tutti i modelli includono effetti fissi temporali (anno e mese) per controllare le variazioni stagionali della domanda turistica e della disponibilità degli host.

### **Sintesi interpretativa**

Nel complesso, i risultati confermano che le caratteristiche visive del profilo costituiscono un potente segnale di fiducia e influenzano significativamente la performance economica su Airbnb.

Il genere maschile mostra un vantaggio in termini di volume di prenotazioni ma uno svantaggio in termini di prezzo e valutazioni, delineando un gender gap complesso e multidimensionale.

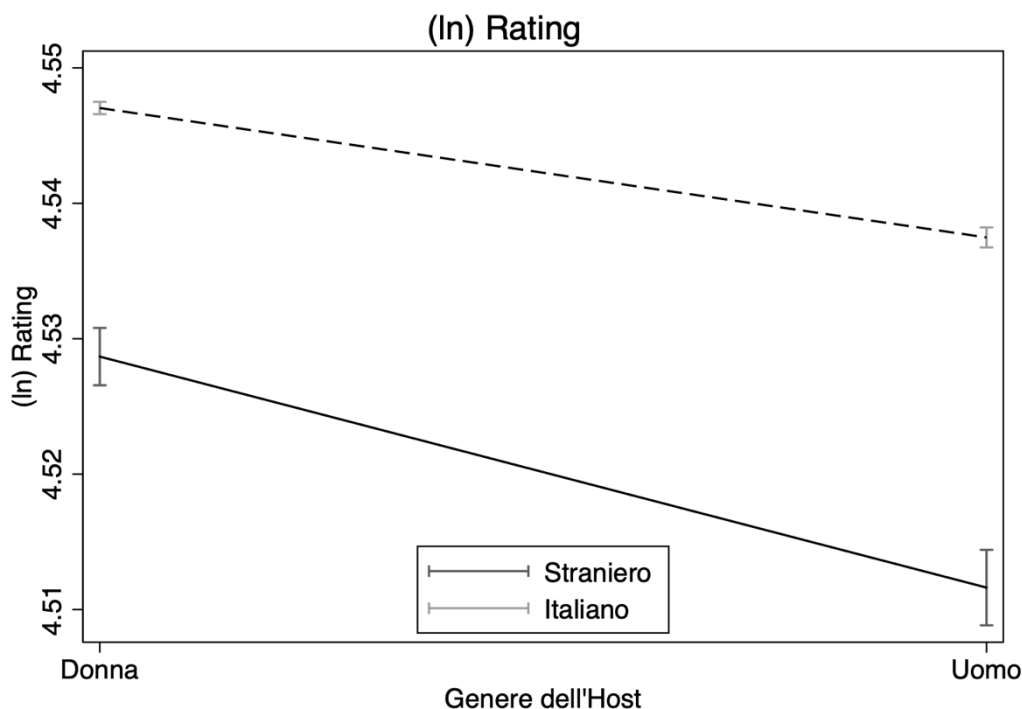
L'efficacia delle foto dipende fortemente dal contenuto e dal contesto visivo: le immagini autentiche e riconoscibili migliorano i risultati, mentre quelle impersonali (occhiali da sole, foto di gruppo) riducono la fiducia, soprattutto per gli host maschi.

Le variabili di reputazione (Superhost, Instant book, response rate) e la qualità percepita dell'alloggio (amenities, numero di foto, capacità ricettiva) si confermano determinanti chiave per la visibilità e la redditività, coerentemente con i principi della teoria del signaling e della fiducia nelle piattaforme peer-to-peer.

### **5.5 – 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 il genere dell'host influenzi diverse metriche di performance (Rating, RevPAN, Occupancy rate, Numero di Prenotazioni e ADR) in modo differenziato a seconda della nazionalità.

**RATING***Figura 27 - Interaction plot - 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 leggermente superiore a 4.55, mentre gli uomini registrano un rating di circa 4.53. 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 italiani ottengono in generale rating leggermente superiori rispetto agli stranieri, indipendentemente dal genere, ma il gender gap rimane comparabile tra le due categorie.

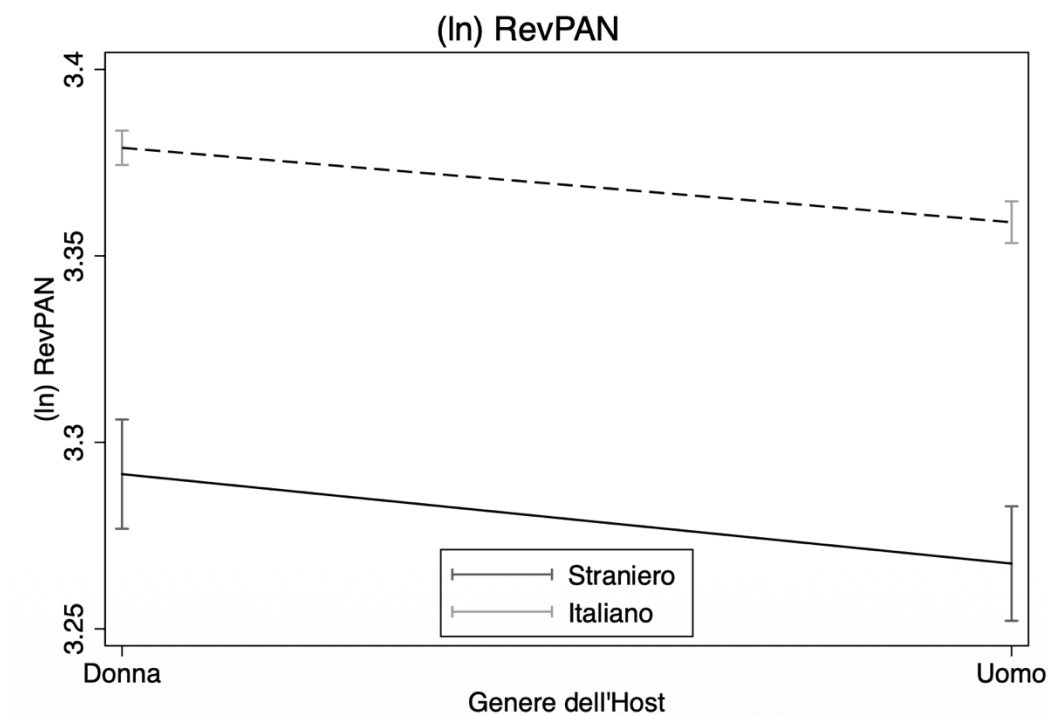
**REVPAN**

Figura 28 - Interaction plot - 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. E' possibile osservare che entrambe le linee mostrano una pendenza negativa, il che significa che gli host uomini, in media, ottengono un RevPAN più basso rispetto alle host donne, a prescindere dalla loro nazionalità. La distanza tra le due linee resta quasi identica lungo tutto l'asse del genere, segno che la nazionalità dell'host influisce sul livello complessivo delle performance, ma non modifica in modo rilevante il divario di genere nel RevPAN.

## OCCUPANCY RATE

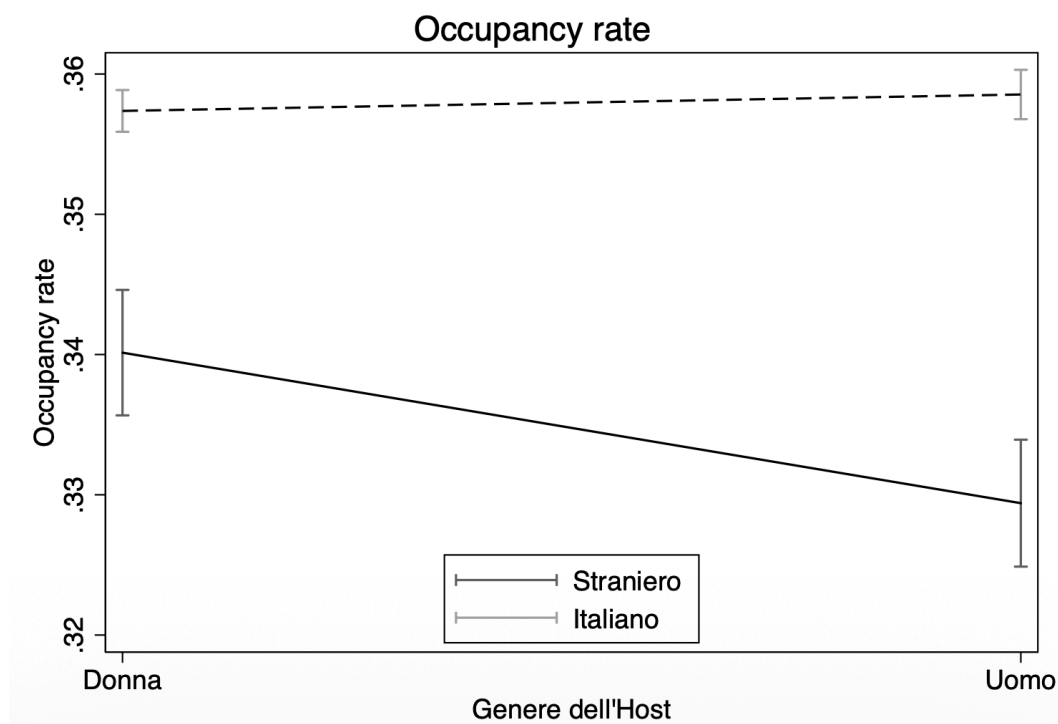


Figura 29 - Interaction plot - OCC

Il tasso di occupazione mostra dinamiche parzialmente diverse rispetto alle metriche precedenti. 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. Possiamo osservare due linee quasi parallele ma leggermente distaccate e ciò mostra quindi che gli host italiani mantengono un piccolo vantaggio intorno ai 2 punti percentuali sul tasso di occupazione. Questo suggerisce che la nazionalità può moderare il rapporto tra genere e occupancy rate. In particolare, tra gli host italiani il genere non sembra avere un impatto rilevante: uomini e donne ottengono risultati simili. Al contrario, tra gli host stranieri, essere donna rappresenta un piccolo vantaggio competitivo, con tassi di occupazione più alti, anche se partendo da valori medi generalmente inferiori rispetto agli italiani.

## NUMERO DI PRENOTAZIONI

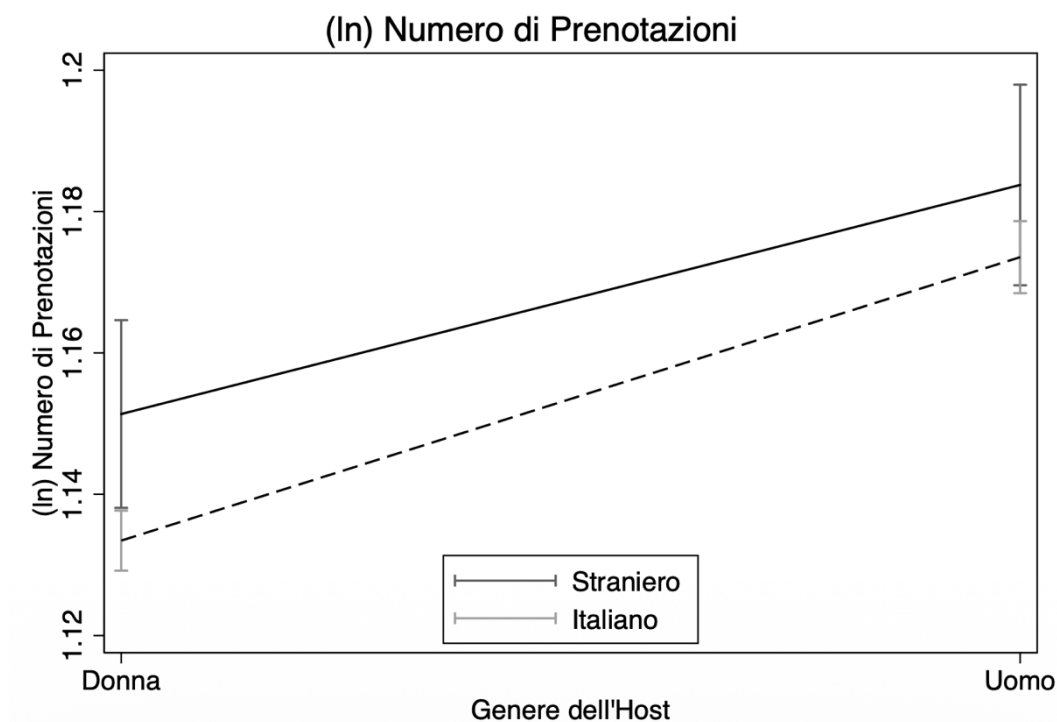
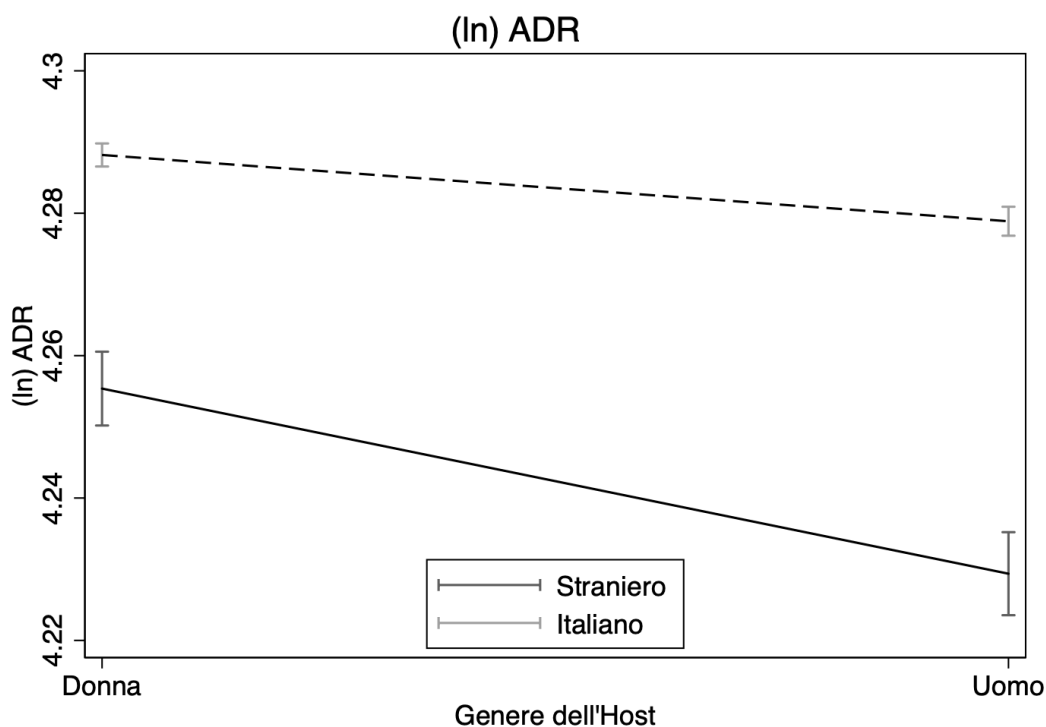


Figura 30 - Interaction plot - Numero di prenotazioni

L'interaction plot relativo al numero di prenotazioni presenta un pattern unico e particolarmente interessante, differenziandosi nettamente dalle altre metriche analizzate. Diversamente dai grafici precedenti, 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 (ln) 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. Possiamo vedere delle linee quasi parallele e ciò significa che il gender gap nel numero delle prenotazioni è simile per italiani e stranieri. Quindi essere uomo rappresenta un vantaggio in termini di volume di prenotazioni ricevute.



**ADR (AVERAGE DAILY RATE)***Figura 31 - Interaction plot - 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 (ln) 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) hanno dei valori inferiori in quanto le donne partono da circa 4.26 mentre gli uomini si attestano intorno a 4.23. La pendenza negativa è maggiore rispetto agli italiani suggerendo un gender gap più grande 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 donna.
- 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 italiani mantengono un prezzo più alto, indipendentemente dal genere.

## CAPITOLO 6 – CONCLUSIONI E IMPLICAZIONI FUTURE

L'ultimo capitolo viene scritto in ottica di definire 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.

Allo stesso tempo vengono definiti i limiti dello studio e delle possibili direzioni future, come nuovi approcci, approfondimenti o ambiti poco esplorati che potrebbero arricchire la tematica alla base.

### 6.1 Sintesi dei risultati e risposte alle domande di ricerca

La presente ricerca si è proposta di analizzare in che modo il genere e i segnali visivi (*visual cues*) delle foto profilo degli host su Airbnb influenzino la domanda di prenotazioni, il rating degli annunci e i ricavi medi per notte (RevPAN).

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 parte delle ipotesi formulate.

In particolare:

- **RQ1:** Il genere e i segnali visivi generano un effetto significativo sulla *domanda* di prenotazioni.  
→ I risultati mostrano che gli host percepiti come più affidabili, in particolare le donne e 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.
- **RQ2:** Il genere e i segnali visivi influenzano significativamente il *rating* medio.  
→ Anche in questo caso l'ipotesi è confermata e cioè che le host donne 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:** Il genere e i segnali visivi influenzano i ricavi per notte (*RevPAN*).

→ I risultati mostrano un effetto meno marcato: nonostante la maggiore fiducia e il rating più alto, i ricavi medi delle host donne non risultano superiori a quelli degli uomini. Ciò suggerisce che la relazione tra reputazione e performance economica sia mediata da altri fattori (posizione geografica, tipologia di alloggio, strategie di prezzo), portando a una parziale conferma di HP3.

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, anche senza parole, trasmettono fiducia o diffidenza. Questi dettagli, spesso percepiti in maniera automatica, 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 e di policy

Oltre ai contributi accademici, i risultati offrono spunti pratici per i gestori di piattaforme digitali, i professionisti del settore turistico e i decisori pubblici.

#### 1. Design inclusivo delle piattaforme:

Sarebbe utile nascondere alcune informazioni dei profili per evitare che si verifichino discriminazioni basate su pregiudizi che si basano sull'aspetto o sul nome. Anche lo sviluppo degli algoritmi dovrebbe andare nella direzione di garantire una maggiore equità tra gli utenti ed evitare discriminazioni.

#### 2. 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.

### 3. 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.

### 4. 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 parecchio.

In secondo luogo, i dati sono dati reali e non ottenuti da esperimenti controllati, quindi è possibile solo osservare delle associazioni statistiche, ma non si possono trarre conclusioni sulla effettiva presenza di relazioni di causa ed effetto.

## 6.4 Direzioni per ricerche future

Partendo dai limiti sopra indicati, si individuano alcune direzioni di approfondimento:

### 1. Analisi multi-contesto:

Estendere l'indagine ad altre città o piattaforme della sharing economy (come Booking o Uber) per confrontare la presenza e l'intensità del *gender gap* in ambienti diversi.

### **2. Integrazione di tecniche di computer vision:**

L'utilizzo di strumenti di riconoscimento delle immagini permetterebbe di analizzare in modo più preciso le caratteristiche visive dei profili, come il sorriso o la luminosità della foto, e di capire meglio come questi elementi influenzano i risultati economici.

### **6.5 Considerazioni conclusive**

In conclusione, questa tesi mette in evidenza come, anche nei sistemi digitali più automatizzati, le percezioni e le dinamiche sociali legate al genere continuino a lasciare un segno profondo. Influenzano la fiducia che le persone si accordano a vicenda, il modo in cui si costruisce la reputazione online e persino i risultati economici che ne derivano. Le piattaforme digitali, quindi, non sono semplici spazi neutri di scambio, ma veri e propri ambienti sociali complessi, dove algoritmi e immagini influenzano chi ha più o meno possibilità di successo.

L'obiettivo, in fondo, è costruire piattaforme che non limitino, ma che diano libertà. Spazi digitali dove la tecnologia diventi un mezzo per ridurre le disuguaglianze e favorire la parità, rendendo la nuova economia della condivisione un terreno più aperto e umano per tutti.

## BIBLIOGRAFIA

- 1. Ert, E., Fleischer, A. e Kopolovich, D. Gender earning gap on digital platforms: The Airbnb case. *Tourism Economics*. 2025.
- 2. *Image features and demand in the sharing economy: A study of Airbnb listings*. He, J., Li, B. e Wang, X. 2020, International Journal of Research in Marketing.
- 3. *Learning To Rank Diversely At Airbnb*. Halder, M., et al. 2023.
- 4. *Reputation and Feedback Systems in Online Platform Markets*. Tadelis, Steven. 2016, Annual Review of economics.
- 5. *The gender earnings gap in sharing economy services: The role of price, number of stays, and guests accommodated on Airbnb*. Davidson, A. e Gleim, M.R. 2022, Journal of Marketing Theory and Practice.
- 6. *Does gender congruence matter? The impact of gender matching on trust in the sharing economy*. Su, L. e Mattila, A.S. 2020, International Journal of Hospitality Management.
- 7. *"Doing gender" by sharing: Examining the gender gap in the European sharing economy*. Eichhorn, Thomas, Hoffmann, Christian e Heger, Katharina. 2022, Internet Policy Review.
- 8. *Women who host: An intersectional critique of rentier capitalism on AirBnB*. Maier, C. e Gilchrist, K. 2022, Gender, Work & Organization published by John Wiley & Sons Ltd.
- 9. *Put on your sunglasses and smile: The secret of Airbnb hosts' profile photos?* Banerjee, S., Lens, M. e Pal, A. 2022, International Journal of Hospitality Management.
- 10. Becker, G. S. The Economics of Discrimination. Chicago, IL : University of Chicago Press, 1957.

- 11. Arrow, Kenneth J. The theory of discrimination. *Discrimination in Labor Markets*. Princeton, NJ : Princeton University Press, 1973.
- 12. *The statistical theory of racism and sexism*. Phelps, Edmund S. 1972, The American Economic Review.
- 13. *Wage discrimination: Reduced form and structural estimates*. Blinder, A.S. 1973, Journal of Human Resources.
- 14. *Male–female wage differentials in urban labor markets*. Oaxaca, R. 1973, International Economic Review.
- 15. *The Gender Wage Gap: Extent, Trends, and Explanations*. Blau, Francine D. e Kahn, Lawrence M. 2017, Journal of economic literature.
- 16. *A meta-analysis of the international gender wage gap*. Weichselbaumer, D. e Winter-Ebmer, R. 2005, JOURNAL OF ECONOMIC SURVEYS.
- 17. *Gender wage gap in European emerging markets: a meta-analytic perspective*. Iwasaki, I. e Satogami, A. 2023, Journal for Labour Market Research.
- 18. *The Rise of the Platform Economy*. Kenney, M. e Zysman, J. 2016, Issues in Science and Technology.
- 19. *Airbnb and the sharing economy*. Hall, C.M., et al. 2022, Current Issues in Tourism.
- 20. *Privacy in the sharing economy*. Teubner, T. e Flath, C.M. 2019, Journal of the Association for Information Systems.
- 21. *Pros vs Joes: Agent pricing behavior in the sharing economy*. Li, J. 2016.
- 22. *The Rise of the Sharing Economy: Estimating the Impact of Airbnb on the Hotel Industry*. Zervas, G., Proserpio, D. e Byers, J.W. 2017, Journal of Marketing Research.

- 23. *Who Benefits from the “Sharing” Economy of Airbnb?* Quattrone, G., et al. 2016, Proceedings of the 25th international conference on world wide web.
- 24. *Digital Discrimination: The Case of Airbnb.com.* Edelman, B. e Luca, M. 2014, Harvard Business School.
- 25. *Offline biases in online platforms: a study of diversity and homophily in Airbnb.* Koh, V., et al. 2019, EPJDataScience.
- 26. *Occupational gender segregation in an equilibrium search model.* Usui, Emiko. 2015, Usui IZA Journal of Labor Economics .
- 27. *Labor market sorting and the gender pay gap revisited.* Strittmatter, A. e Wunsch, C. 2025, Journal of Population Economics.
- 28. OECD. Gender wage gap (indicator) . *OECD Data* . [Online] 2023 .
- 29. Ridgeway, C. L. *Framed by gender: How gender inequality persists in the modern world.* 2011.
- 30. OECD, Publishing. *Reporting Gender Pay Gaps in OECD Countries.* Paris : s.n., 2023.
- 31. OECD. Social Institutions and Gender Index 2022: Global report. [Online] 2022.
- 32. —. Social Institutions & Gender Index Dashboard. <https://www.oecd.org/en/data/dashboards/social-institutions-gender-index.html>. [Online]
- 33. Nations, United. *Transforming our World: The 2030 Agenda for Sustainable Development* . New York : s.n., 2015.
- 34. *Platform competition in two-sided markets.* Rochet, J.C. e Tirole, J. 2003, Journal of the European Economic Association.



- 35. *Matchmakers: The new economics of multisided platforms*. Evans, G.G. e Schmalensee, R. 2016, Harvard Business Review Press.
- 36. Parker, G.G., Van Alstyne, M.W. e Choudary, S.P. *Platform revolution: How networked markets are transforming the economy—and how to make them work for you*. 2016.
- 37. *Two-Sided Network Effects: A Theory of Information Product Design*. Parker, G. e Van Alstyne, M. 2005, Management Science.
- 38. *The Rise of the Platform Enterprise. A Global Survey*. Evans, P.C. e Gawer, A. 2016, The Emerging Platform Economic Series.
- 39. Valant, J. *A European Agenda for the Collaborative Economy*. s.l. : Members' Research Service Briefing. Brussels: European Parliamentary Research Service, 2016.
- 40. *Chicken & egg: Competition among intermediation service providers*. Caillaud, B. e Jullien, B. 2003, Journal of Economics.
- 41. *When does the winner take all in two-sided markets?* . Sun, M. e Tse, E. 2007, Review of Network Economics.
- 42. Botsman, R. e Rogers, R. *What's mine is yours: The rise of collaborative consumption*. 2010.
- 43. *From Zipcar to the sharing economy*. Sundararajan, A. 2013, Harvard Business Review.
- 44. *The sharing economy globalization phenomenon: A research agenda*. Parente, R.C., Geleilate, J.M.G. e Rong, K. 2018, Journal of International Management.
- 45. *Oxford English Dictionary – Sharing economy*. [Online] [https://www.oed.com/dictionary/sharing-economy\\_n?tab=factsheet#1335003350100](https://www.oed.com/dictionary/sharing-economy_n?tab=factsheet#1335003350100) .
- 46. *You are what you can access: Sharing and collaborative consumption online*. Belk, R. 2014, Journal of Business Research.

- 47. *Sharing economy: A review and agenda for future research*. Cheng, M. 2016, International Journal of Hospitality Management.
- 48. *Sharing economy: A comprehensive literature review*. Hossain, M. 2020, International Journal of Hospitality Management.
- 49. *Airbnb: the future of networked hospitality businesses*. Oskam, J. e Boswijk, A. 2016, Journal of Tourism Futures.
- 50. Airbnb: case vacanze, chalet, alloggi ed esperienze. [Online] 2025. <https://www.airbnb.it/>.
- 51. wikipedia. *Airbnb*. [Online] 2025. <https://it.wikipedia.org/wiki/Airbnb>.
- 52. *Monetizing network hospitality: Hospitality and sociability in the context of Airbnb*. Ikkala, T. e Lampinen, A. 2015.
- 53. *Is home sharing driving up rents? Evidence from Airbnb in Boston*. Horn, K. e Merante, M. 2017, Journal of Housing Economics.
- 54. *Airbnb and the rent gap: Gentrification through the sharing economy*. Wachsmuth, D. e Weisler, A. 2018, Environment and Planning A: Economy and Space.
- 55. Raboni, Andrea. Cos'è la gentrificazione, quando interi quartieri si modernizzano e "cambiano" cittadini. [Online] 2023. <https://www.geopop.it/cose-la-gentrificazione-quando-interi-quartieri-si-modernizzano-e-cambiano-cittadini>.
- 56. Airdna - About. [Online] <https://www.airdna.co/about>.
- 57. «*The Impacts of Quality and Quantity Attributes of Airbnb Hosts on Listing Performance*». Xie, K.L. e Mao, Z. 2017, International Journal of Contemporary Hospitality Management.
- 58. «*How to react to a shock? Effects of Airbnb hosts' choices and market segmentation at the time of Covid-19*». Buzzacchi, L., et al. 2023, Information & Management.

- 59. [www.istat.it](http://www.istat.it). Earnings differences by gender, generation, educational attainment and type of contract. [www.istat.it](http://www.istat.it). [Online]
- 60. World Economic Forum. *Global Gender Gap Report 2023*. [Online] 2023.
- 61. *Airbnb phenomenon: a review of literature and future research directions*. Negi, A. e Tripathi, V. 2022, Journal of Hospitality and Tourism Insights.
- 62. *The sharing economy and digital platforms: A review and research agenda*. Sutherland, W. e Jarrahi, M.H. 2018, International Journal of Information Management.
- 63. *Toward a geographic understanding of the sharing economy: Systemic biases in UberX and TaskRabbit*. Thebault-Spieker, J., Terveen, L. e Hecht, B. 2017, ACM Transactions on Computer-Human Interaction.
- 64. *Does gender bias exist? The impact of gender congruity on consumer's Airbnb booking intention and the mediating role of trust*. Na Su, A. S. e Mattila, Anna S. 2020, International Journal of Hospitality Management.
- 65. *The sharing economy: A critical review and research agenda*. Rojanakit, P., de Oliveira, R. T. e Dulleck, U. 2022, Journal of Business Research.
- 66. *Sharing economy business models for sustainability*. Curtis, SK e Mont, O. 2020, Journal of Cleaner Production.
- 67. *A qualitative investigation of microentrepreneurship in the sharing economy*. Zhang, T., Byfquin, D. e Lu, C. 2019, International Journal of Hospitality Management.
- 68. *Gender Pay Gap in the Gig Economy*. Dong, Z, Peng, J. e Liu, S. 2024, China Finance and Economic Review.
- 69. *Platform inequality: Gender in the gig-economy*. Barzilay, A. R. e Ben-David, A. 2016, Seton Hall Law Review.

- 70. *What Airbnb does to the housing market*. Benítez-Aurioles, Beatriz. 2020, *Annals of Tourism Research*.
- 71. *Is sharing the solution? Exploring public acceptability of the sharing economy*. Cherry, CE e Pidgeon, NF. 2018, *Journal of Cleaner Production*.
- 72. *How sustainable is the sharing economy? On the sustainability connotations of sharing economy platforms*. Geissinger, A., et al. 2019, *Journal of Cleaner Production*.
- 73. *The gender earnings gap in the gig economy: Evidence from over a million rideshare drivers*. Cook, C., et al. 2021, *The Review of Economic Studies*.
- 74. *Airbnb research: an analysis in tourism and hospitality journals*. Andreu, L., et al. 2020, *International Journal of culture, tourism and hospitality research*.

## ELENCO DELLE FIGURE

Figura 1 – Gender Wage Gap – Dati del 2023 – Fonte:OECD.....	6
Figura 2 – Grafico che mostra la discriminazione della famiglia – Fonte: OECD.....	12
Figura 3 – Grafico che mostra l'integrità fisica per le donne - Fonte: OECD.....	13
Figura 4 – Mostra l'accesso alle risorse produttive – Fonte: OECD .....	13
Figura 5 – Mostra le libertà civili per le donne – Fonte: OECD.....	14
Figura 6 – Mostra il caso Italiano confrontato internazionalmente – Fonte: OECD .....	14
Figura 7 - Tipologie di piattaforme digitali - (38) .....	18
Figura 8 - Visualizzazione delle relazioni in una piattaforma a due parti Fonte: A European Agenda for the Collaborative Economy (39) .....	19
Figura 9 - Codice Python per lo scraping dei nomi degli host dal sito Airbnb.....	35
Figura 10 - Codice Python per la pulizia del database.....	38
Figura 11 - Codice Python per la richiesta Namsor per osservare il genere associato al nome degli host.....	40
Figura 12 – Script Python per analizzare le foto profilo degli host manualmente.....	46
Figura 13 - Distribuzione annunci per genere e nazionalità .....	51
Figura 14 - Distribuzione host per genere e nazionalità .....	52
Figura 15 - Ricavi medi per genere.....	53
Figura 16 - Rating medio per genere .....	54
Figura 17 - Numero medio di prenotazioni per genere.....	55
Figura 18 - tipologia di alloggio per genere.....	56
Figura 19 – Status di superhost.....	57
Figura 20 – Numero medio di foto per genere.....	58
Figura 21 – Politica di cancellazione .....	58
Figura 22 – ADR per genere .....	59
Figura 23 – OCC per genere .....	60
Figura 24 – RevPAN per genere .....	61
Figura 25 – Distribuzione delle caratteristiche visive.....	62
Figura 26 – Statistiche descrittive per anno .....	64
Figura 27 - Interaction plot - Rating .....	87
Figura 28 - Interaction plot - RevPAN.....	88
Figura 29 - Interaction plot - OCC.....	89

Figura 30 - Interaction plot - Numero di prenotazioni.....	90
Figura 31 - Interaction plot - ADR.....	91

## ELENCO DELLE TABELLE

Tabella 1 - Tabella sinottica della letteratura sul gender gap in Airbnb .....	4
Tabella 2 - Dataset con host-id e host-name .....	33
Tabella 3 - Output restituito da NamSor .....	38
Tabella 4 – Modelli logaritmici – Diretti per genere e nazionalità .....	65
Tabella 5 - Modelli logaritmici - Diretti e interaction per genere e nazionalità .....	70
Tabella 6 - modelli logaritmici – Diretti per i segnali visivi delle foto profilo degli host.....	77
Tabella 7 – Modelli logaritmici Diretti e di interazione per i segnali visivi nelle foto profilo degli host con Uomo .....	81