

# DATA ANALYSIS SUL DATASET AIRBNB

## MATERIA : DATA ANALYSIS

# **DOCENTE : PISANO GIANMARCO GIUSEPPE**

# GLI ARTISTI

**'FAVA EMANUELE', 'ZAMMITO MIRIANA', 'COSTANZO FEDERICO'**



# INTRODUZIONE



**FONTE DEI DATI:** ANALISI CONDOTTA UTILIZZANDO I DATI DI AIRBNB TRAMITE INSIDE AIRBNB (LISTINGS.CSV, AGGIORNATO IL 29/06/2024).

**OBIETTIVO DELL'ANALISI:** IDENTIFICARE TREND RELATIVI AI POSTI LETTO E ALTRE CORRELAZIONI.

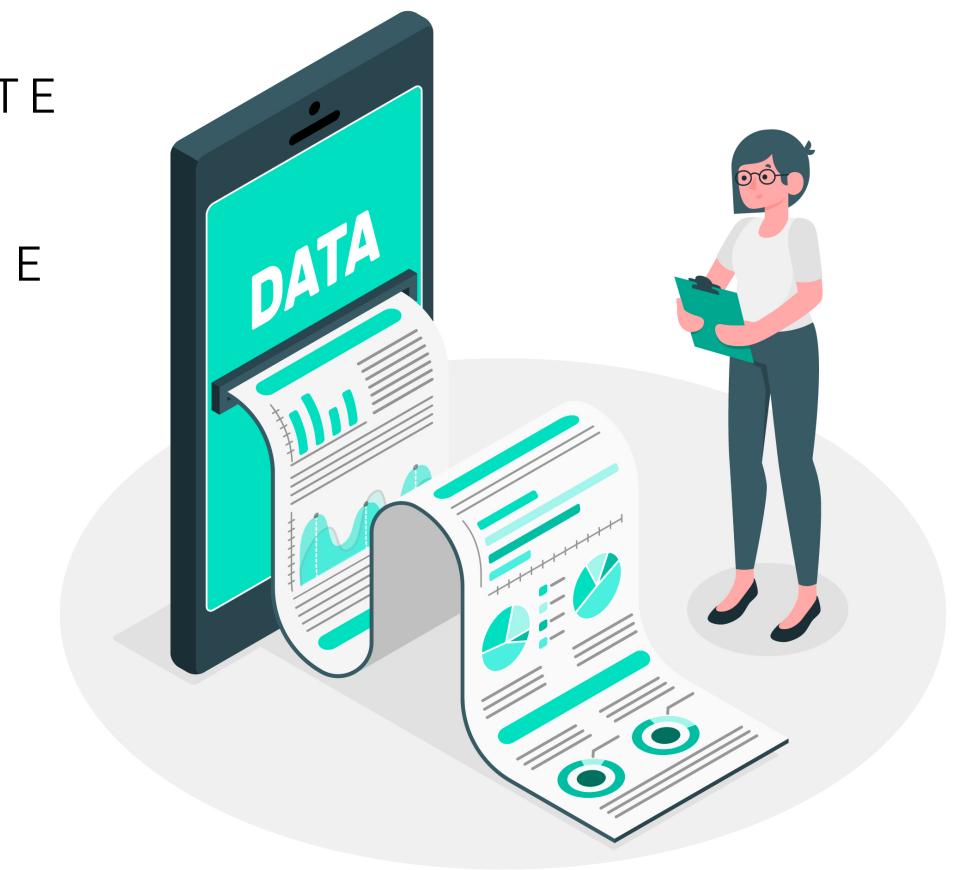
**FOCALIZZAZIONE INIZIALE:** ESCLUSIONE DELLE CITTÀ SICILIANE PER CONCENTRARSI SU PALERMO, DOPO UN'ANALISI PRELIMINARE DELLE CITTÀ CON MAGGIORI ANNUNCI IN SICILIA.

**PRE-PROCESSING DEI DATI:** SELEZIONE DELLE COLONNE, GESTIONE DEI VALORI DUPLICATI E NAN. FILTRAGGIO DELLE COLONNE PRINCIPALI OGGETTO DI STUDIO.

**VISUALIZZAZIONE INIZIALE:** CREAZIONE DI GRAFICI (PAIRPLOT E ISTOGRAMMI) PER EVIDENZIARE ELEMENTI SIGNIFICATIVI.

**ANALISI GEOLOCALE:** IDENTIFICAZIONE DELLE CORRELAZIONI E STANDARDIZZAZIONE DEL MODELLO RELATIVO A PALERMO.

**CONFRONTO CON ALTRE LOCALITÀ:** ANALISI INIZIALE CON LOCALITÀ OCCIDENTALI E SUCCESSIVAMENTE CON QUELLE ORIENTALI PER VERIFICARE LA PRESENZA DI TREND SIMILI.



# PRE PROCESSING



**FASE INIZIALE:** PULIZIA E LETTURA DEI DATI PER COMPRENDERE IL DATAFRAME E LA TIPOLOGIA DI DATI PRESENTI.

**OBIETTIVO DELLA PULIZIA:** IDENTIFICARE ELEMENTI PARTICOLARI E PRIME CORRELAZIONI VISIBILI.

**SORTING ESEGUITI:** UTILIZZO DI VALORI COME 'ACCOMMODATES', 'NUMBER OF REVIEWS', 'NAME'.  
ORDINAMENTI CASUALI TRAMITE LA FUNZIONE SAMPLE().  
ANALISI DEI PRIMI 5 E ULTIMI 5 ELEMENTI DEL DATAFRAME.

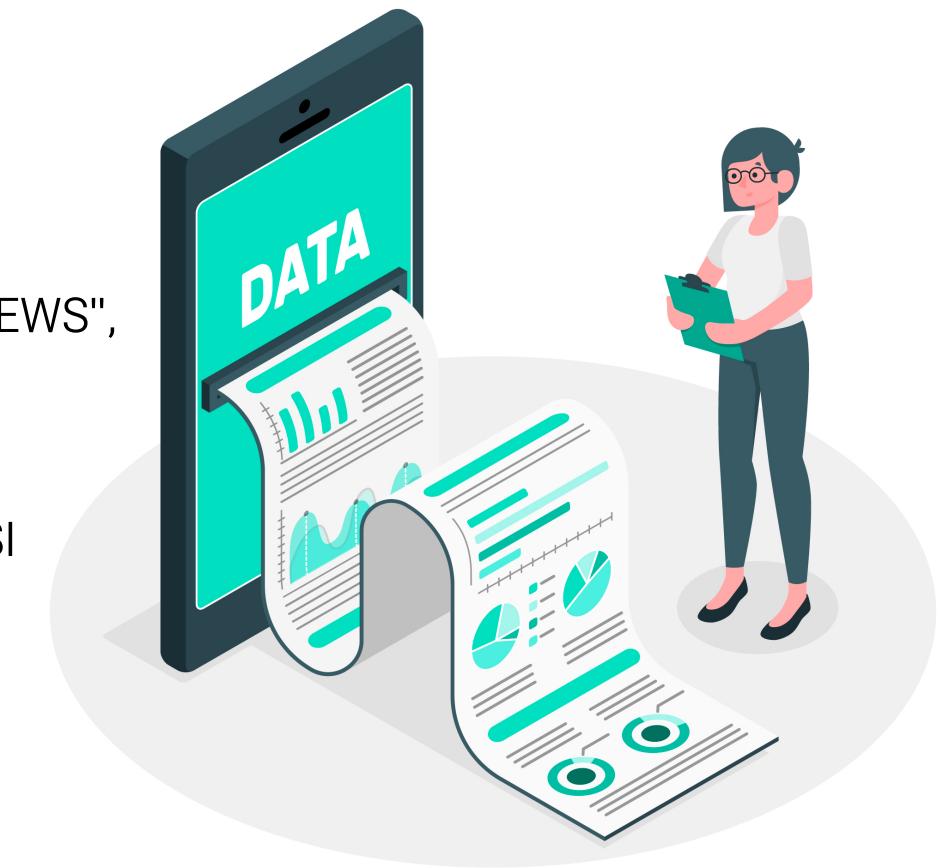
#### **STRUTTURA DEL DATASET:**

TOTALE COLONNE: 75

TOTALE RIGHE: 59,887

**COLONNE PRINCIPALI:** ID, 'ACCOMMODATES',  
'NUMBER\_OF\_REVIEWS', 'PRICE', 'MAXIMUM NIGHT', ECC.  
COLONNE SELEZIONATE PER APPROFONDIMENTO:  
["NAME", "DESCRIPTION", "NEIGHBOURHOOD\_CLEANSED",  
"ROOM\_TYPE", "ACCOMMODATES", "PRICE", "NUMBER\_OF\_REVIEWS",  
"FIRST REVIEW", "LAST REVIEW", "REVIEW\_SCORES\_RATING",  
"REVIEWS\_PER\_MONTH"]

**POSSIBILITÀ DI ULTERIORI SCREMATURE:** IN BASE ALL'ANALISI SPECIFICA E ALLE CORRELAZIONI DA STUDIARE.



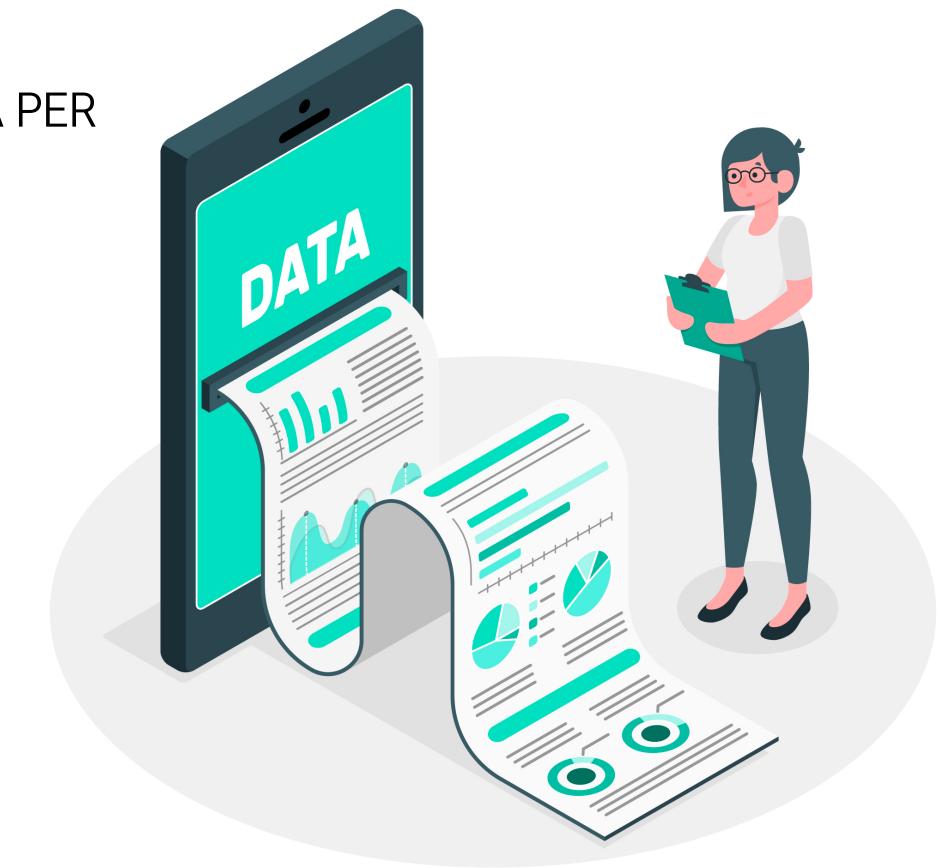
### **VALORI MANCANTI E DUPLICATI:**

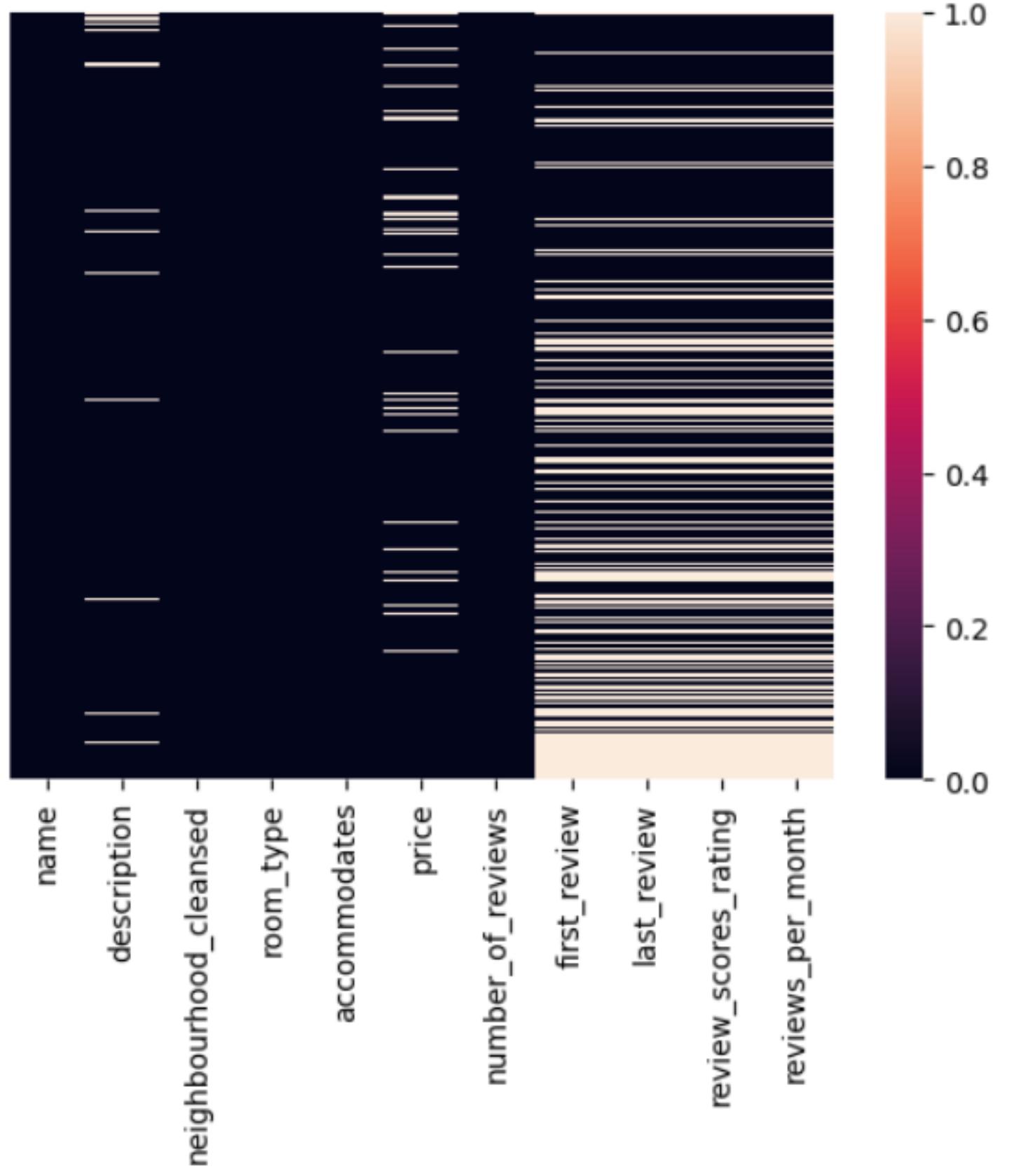
MINIMA PRESENZA DI VALORI DUPLICATI (50 NEL DATASET COMPLETO DELLA SICILIA).

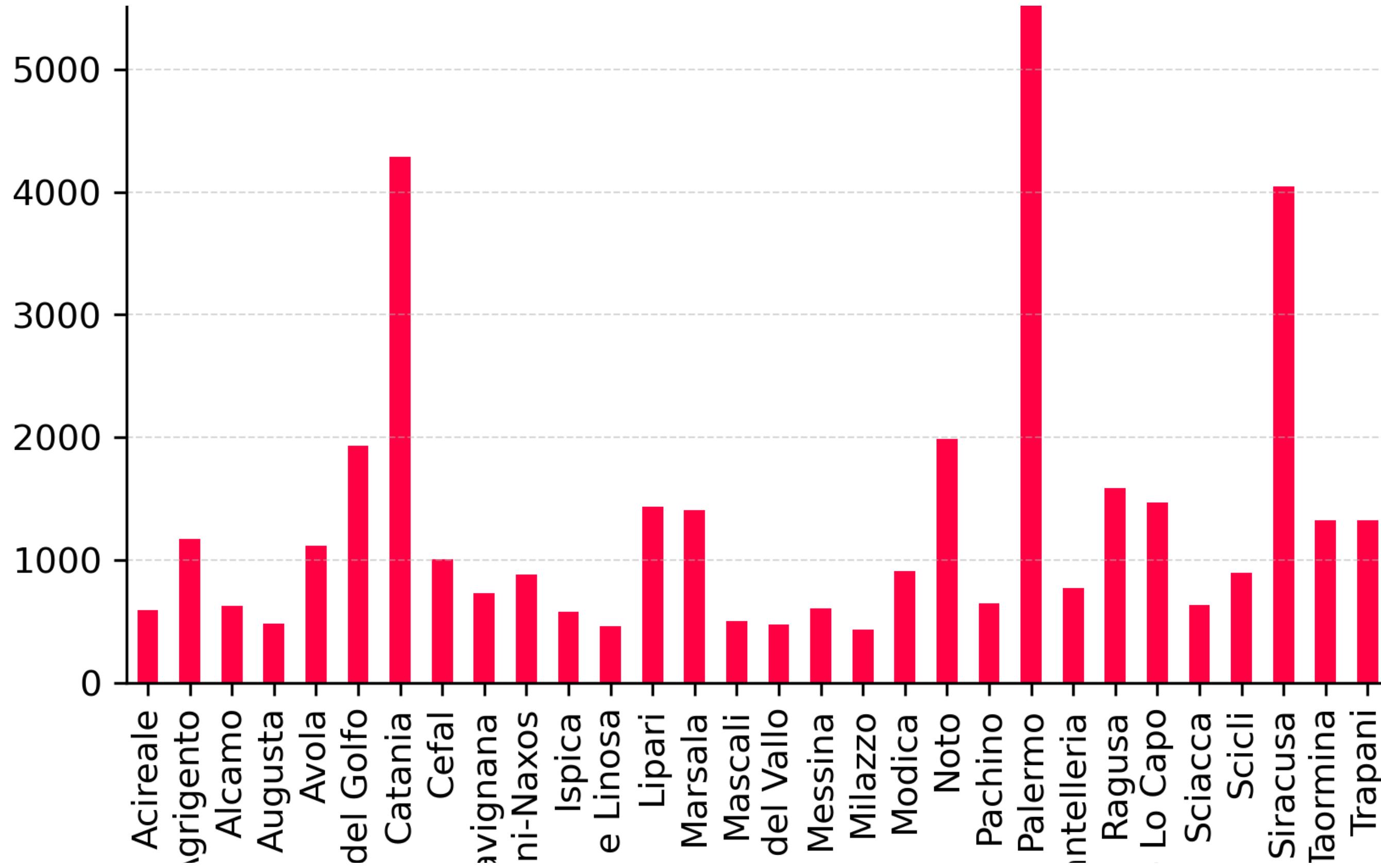
**ANALISI DEI VALORI NULLI:** UTILIZZO DEL METODO ISNULL() E DELLA LIBRERIA SEABORN PER VISUALIZZARE I VALORI NULLI TRAMITE UNA HEATMAP.

### **DECISIONI SUI VALORI NULLI:**

- **ESCLUSIONE** DI ALCUNE COLONNE CON ALTA CONCENTRAZIONE DI VALORI NULLI.
- **MANTENIMENTO** DEI VALORI NULLI PER PRESERVARE I DATI NECESSARI PER UN'ANALISI PIÙ ATTENDIBILE.
- **DECISIONE** DI NON RIEMPIRE I VALORI NULLI CON LA MEDIA PER EVITARE DI INFLUENZARE I DATI IMPORTANTI.







# PRIME OSSERVAZIONI



**FASE DI ANALISI:** DOPO LA PULIZIA DEI DATI E L'IDENTIFICAZIONE DEI VALORI NAN, SI È PROCEDUTO A DEFINIRE GLI OBIETTIVI DELL'ANALISI.

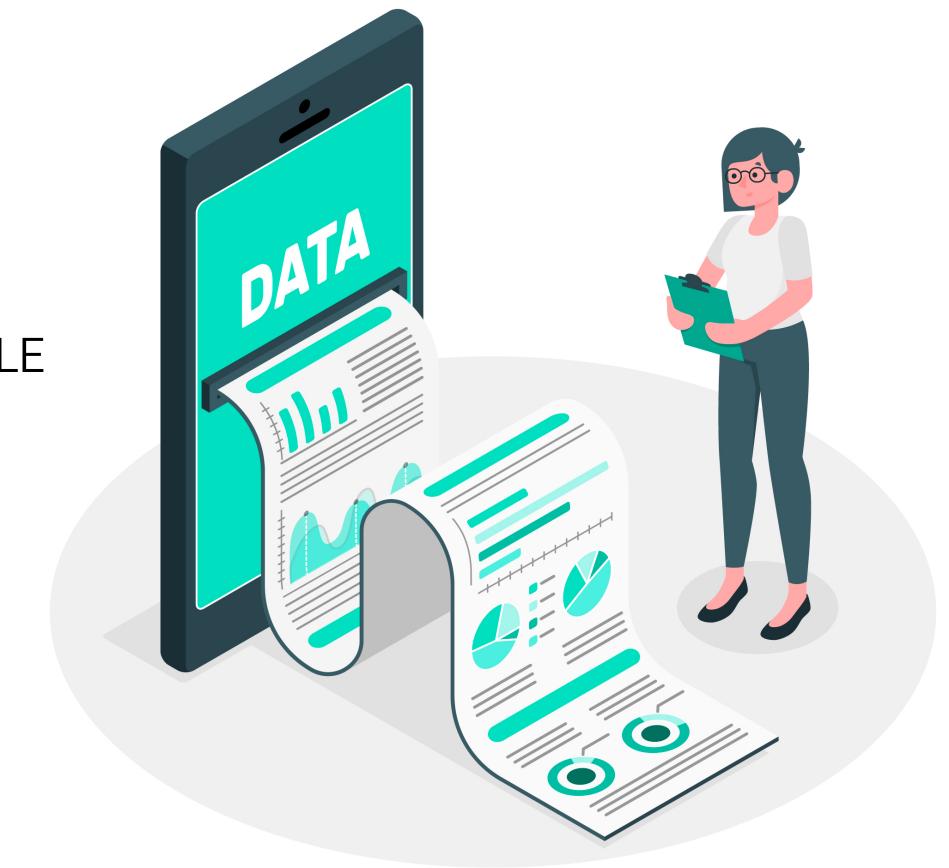
**OBIETTIVI IDENTIFICATI:**

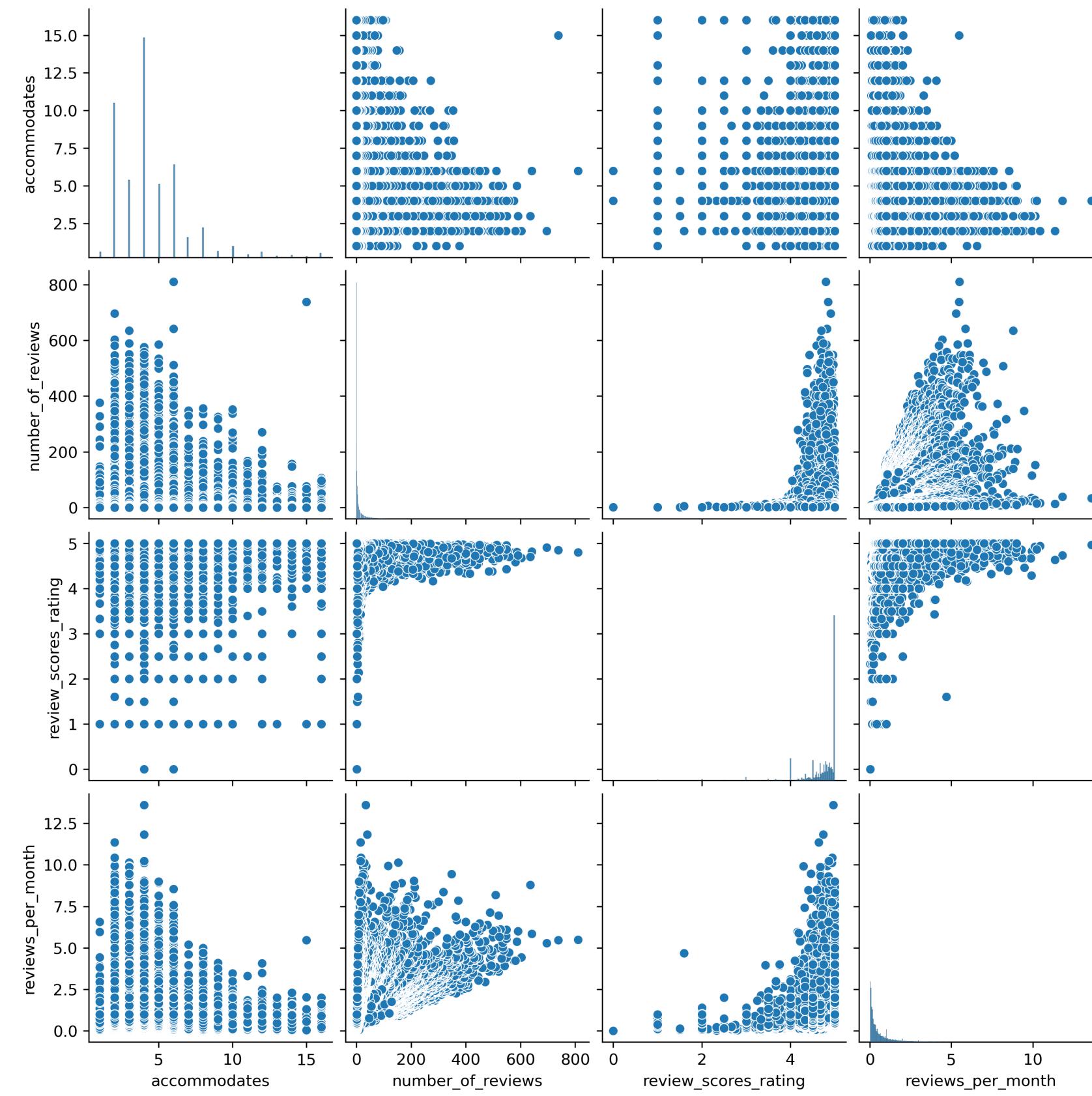
- VERIFICARE SE LA TIPOLOGIA DI CAMERA INFLUENZI IL PREZZO.
- ANALIZZARE SE LA TIPOLOGIA DI ANNUNCIO CAMBI IN BASE AL NUMERO DI POSTI LETTO.
- ESPLORARE LA CORRELAZIONE TRA IL NUMERO DI RECENSIONI E IL NUMERO DI POSTI LETTO/OSPITI.
- IDENTIFICARE QUALI POSTI HANNO RECENSIONI PIÙ ALTE.
- DETERMINARE IN QUALI ZONE CI SIANO PIÙ AFFITTI.
- VISUALIZZARE TREND IN DIVERSI CONTESTI CULTURALI

**STRUMENTI UTILIZZATI:** CREAZIONE DI UN PAIRPLOT E UN ISTOGRAMMA PER FORNIRE SPUNTI SUGLI OBIETTIVI DA PERSEGUIRE.

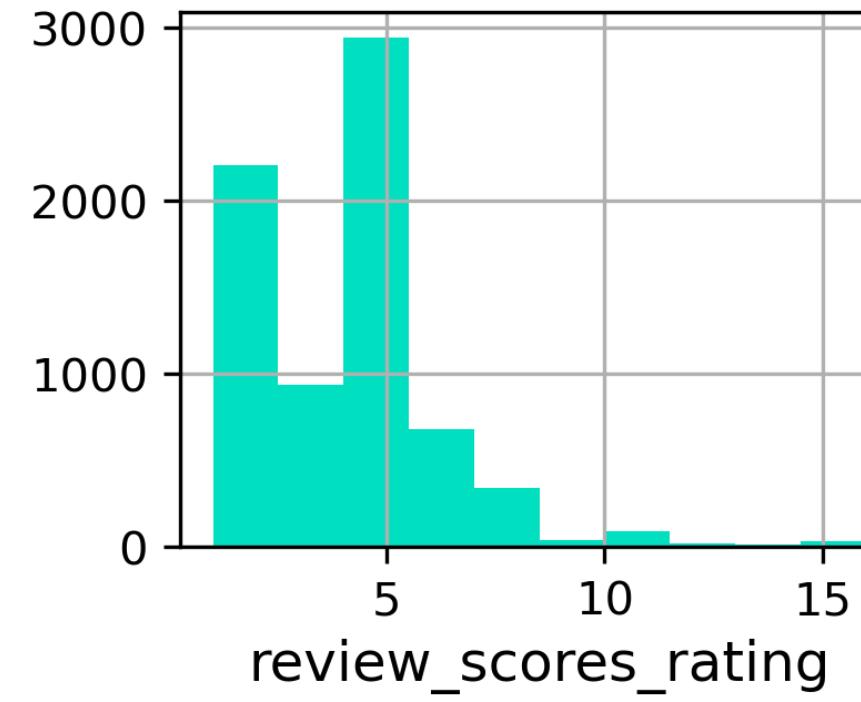
**OSSERVAZIONI INIZIALI:**

IL PAIRPLOT HA RIVELATO CONCENTRAZIONI PARTICOLARI NELLE AREE DEI GRAFICI, IN PARTICOLARE TRA 'ACCOMMODATES' (POSTI LETTO) E 'REVIEWS' (RECENSIONI).

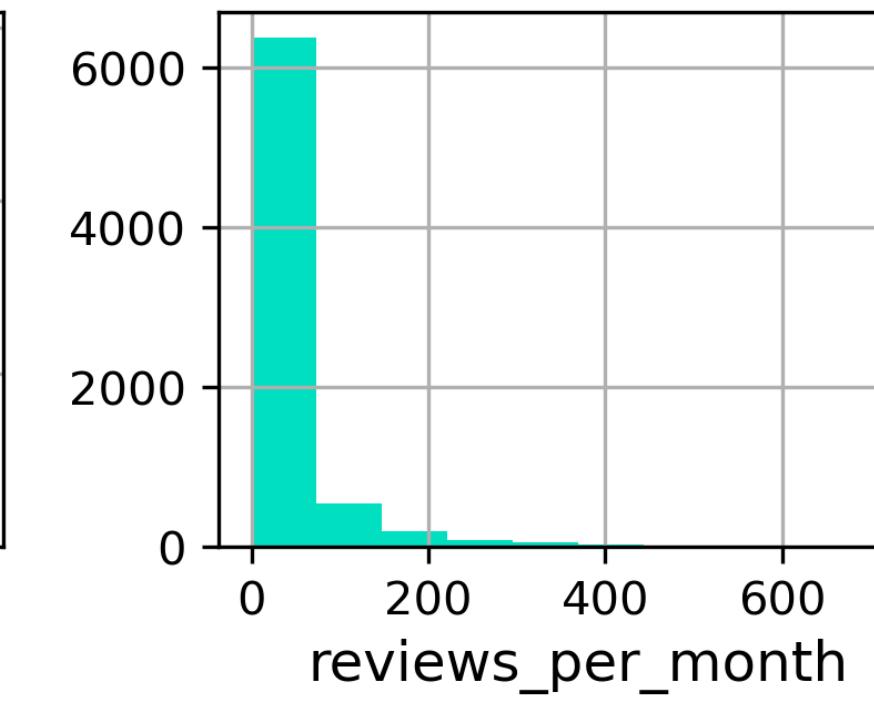




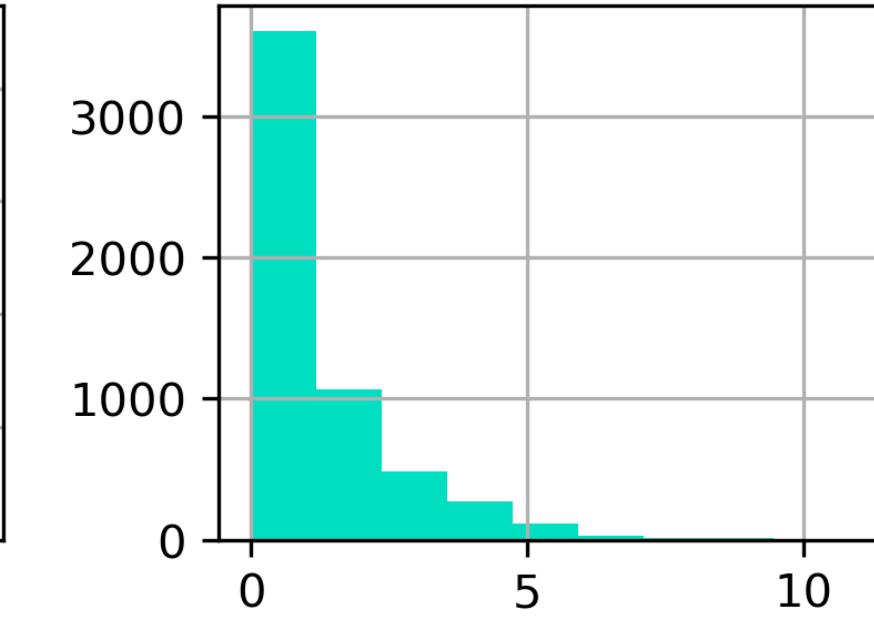
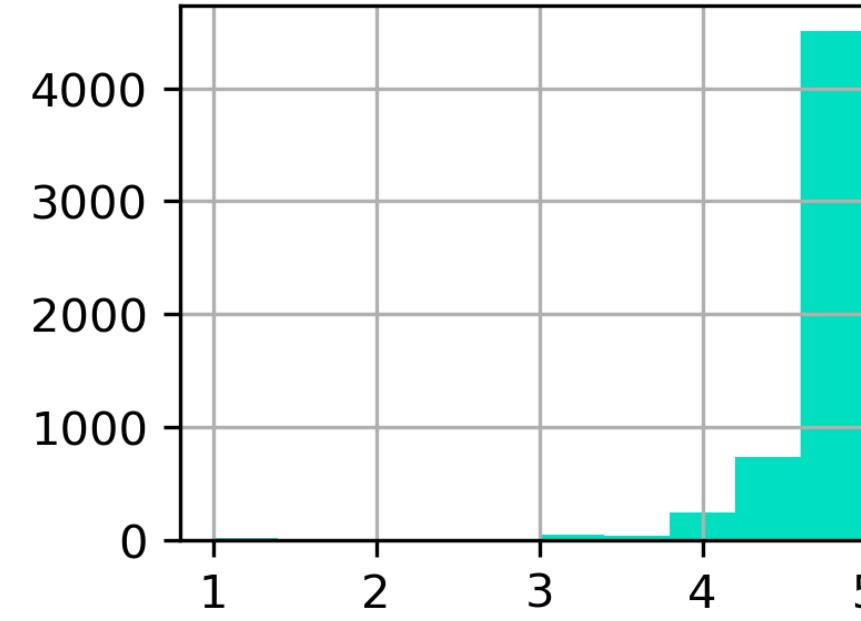
accommodates



number\_of\_reviews



review\_scores\_rating



# ANALISI E CORRELAZIONI

## RECENSIONI X POSTO LETTO



**OBIETTIVO DELL'ANALISI:** ESAMINARE LA CORRELAZIONE TRA POSTI LETTO E RECENSIONI PER IDENTIFICARE GLI ANNUNCI CON IL MAGGIOR NUMERO DI RECENSIONI.

**RISULTATI PRINCIPALI:**

*PICCHI SIGNIFICATIVI* DI RECENSIONI PER ANNUNCI CON DUE E QUATTRO POSTI LETTO.

*NOTEVOLE CALO* DELLE RECENSIONI PER GLI ANNUNCI CON TRE POSTI LETTO, SIA IN DISPONIBILITÀ CHE IN RECENSIONI.

**DATI SPECIFICI:**

NUMERO DI RECENSIONI PER:

**DUE POSTI LETTO:** 64,834

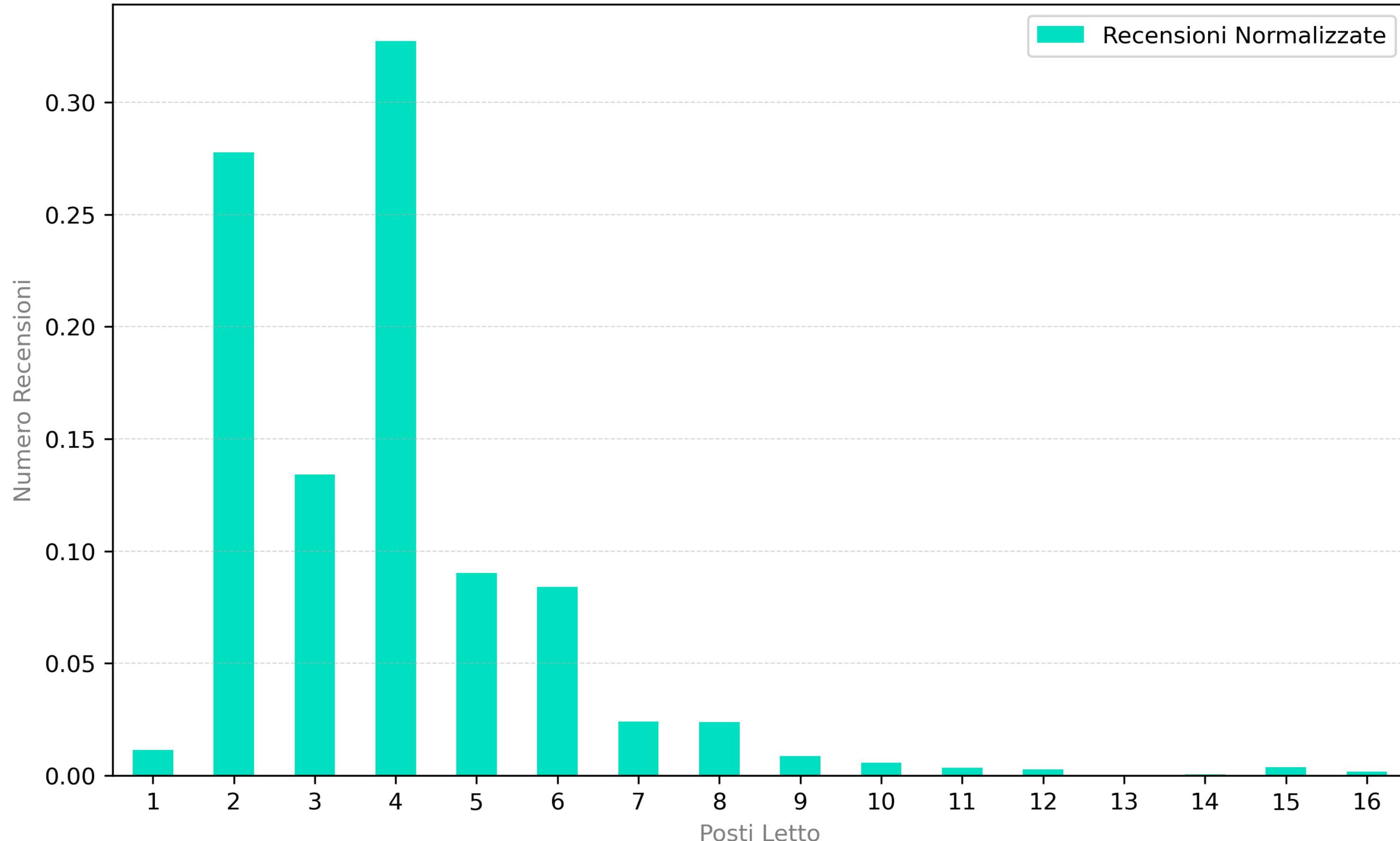
**QUATTRO POSTI LETTO:** 76,423

**TREDICI POSTI LETTO:** 90 (MINIME RECENSIONI).

**TREND OSSERVATO:** CORRELAZIONE RISCONTRATA ANCHE IN ALTRE ZONE DELLA SICILIA E IN DIVERSE CITTÀ OCCIDENTALI.



## Rapporto numero recensioni in Percentile e posti letto - Palermo



# ANALISI E CORRELAZIONI

## RECENSIONI X TIPO DI STANZA



**NUOVA VARIABILE AGGIUNTA:** ANALISI DELLA CORRELAZIONE TRA IL NUMERO DI RECENSIONI E LA TIPOLOGIA DI STANZA.

**TIPOLOGIE CON MAGGIOR NUMERO DI RECENSIONI:**

**INTERE CASE/APPARTAMENTI:** 199,693 RECENSIONI.

**STANZE PRIVATE:** 32,570 RECENSIONI.

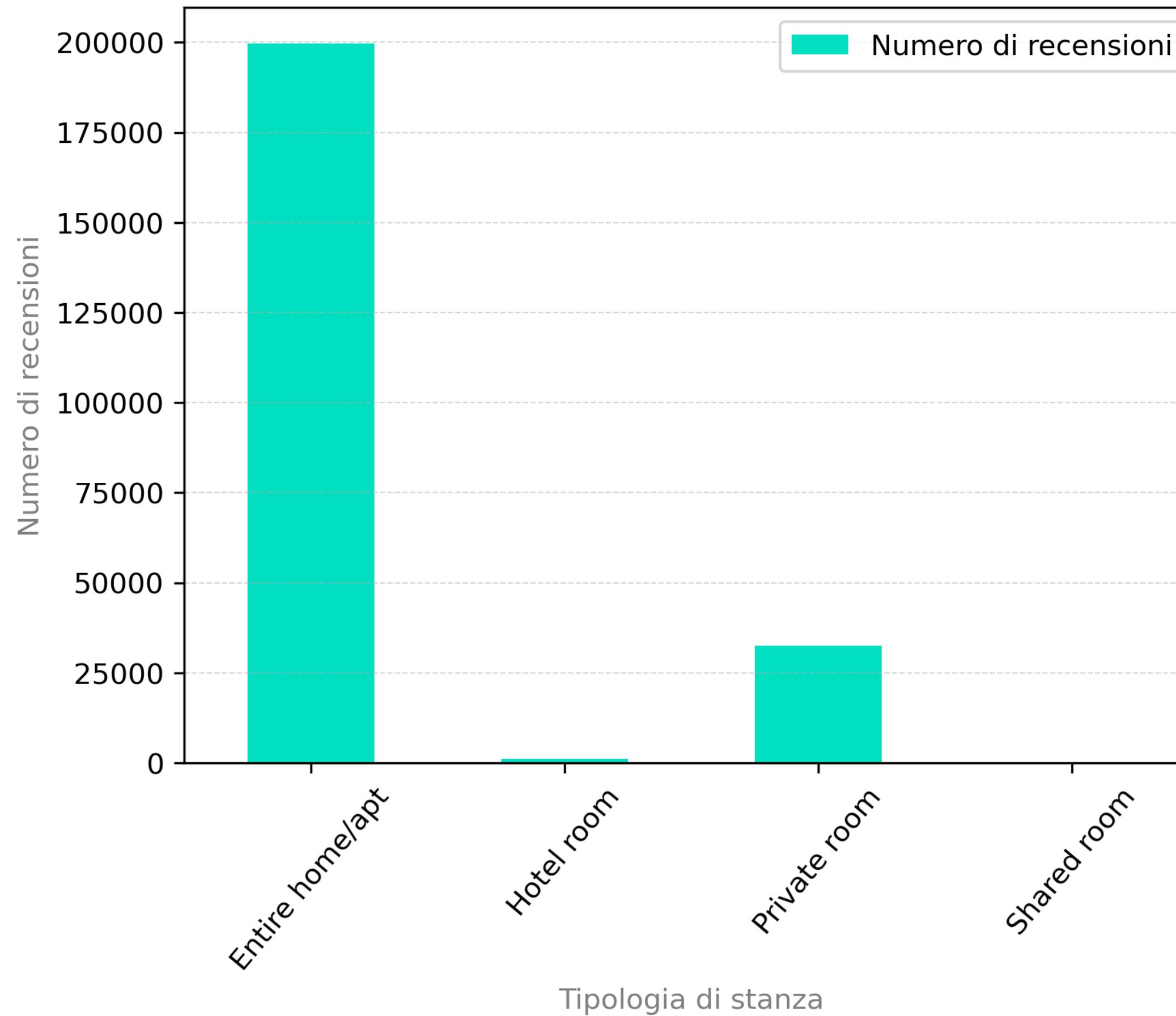
**TIPOLOGIE CON MINORE NUMERO DI RECENSIONI:**

**STANZE CONDIVISE (SHARED ROOM):** SOLO 215 RECENSIONI.

**OSSERVAZIONI CHIAVE:** LE TIPOLOGIE DI LOCALI CON MAGGIOR NUMERO DI RECENSIONI EVIDENZIANO UNA PREFERENZA PER INTERE CASE/APPARTAMENTI RISPETTO AD ALTRE OPZIONI.



## Numero di recensioni per tipologia di stanza - Palermo



## N° ANNUNCI PER TIPO DI STANZA E POSTI LETTO

**NUOVA ANALISI:** VISUALIZZAZIONE DEL NUMERO DI ANNUNCI PER POSTI LETTO IN RELAZIONE ALLA TIPOLOGIA DI STANZA.

**RAGGRUPPAMENTO DEI DATI:** CREAZIONE DI UNA TABELLA A DOPPIA ENTRATA PER CONTEGGIARE GLI ANNUNCI PER CIASCUNA COMBINAZIONE DI POSTI LETTO E TIPO DI STANZA.

### RISULTATI PRINCIPALI:

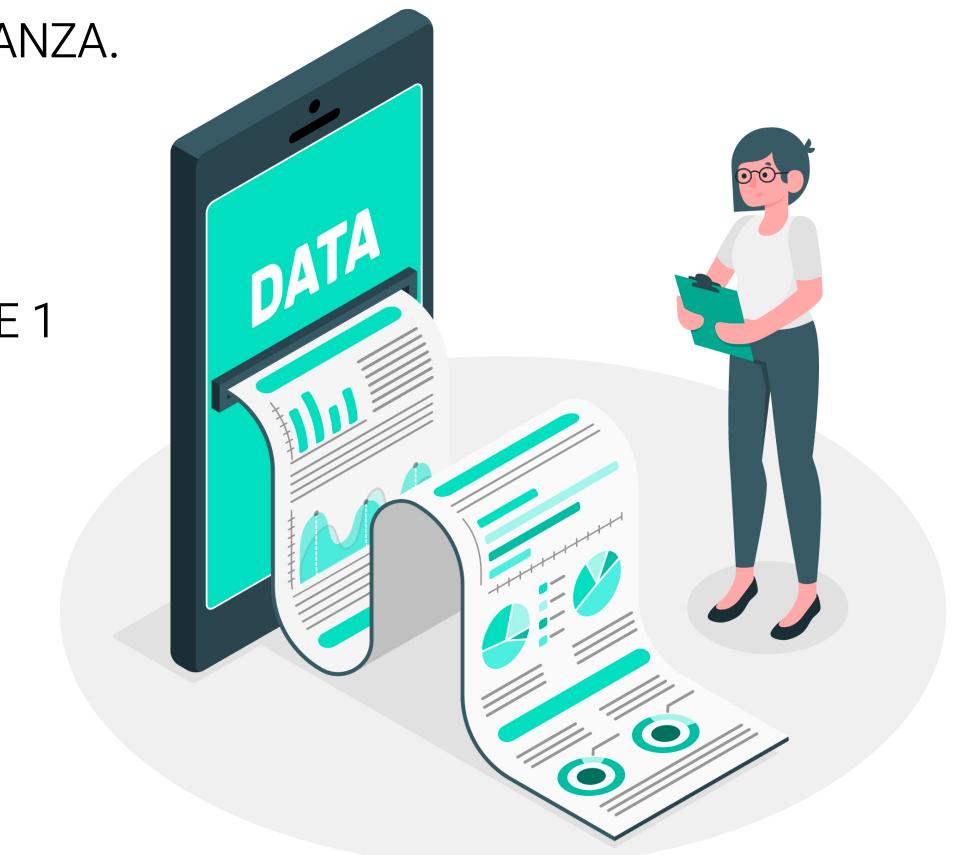
**LOCALI CON DUE POSTI LETTO:** DISTRIBUZIONE EQUA TRA INTERE CASE/APPARTAMENTI E STANZE PRIVATE.

**LOCALI CON QUATTRO POSTI LETTO:** PREDOMINANZA DELLE INTERE CASE/APPARTAMENTI, CON STANZE PRIVATE IN MINORANZA.

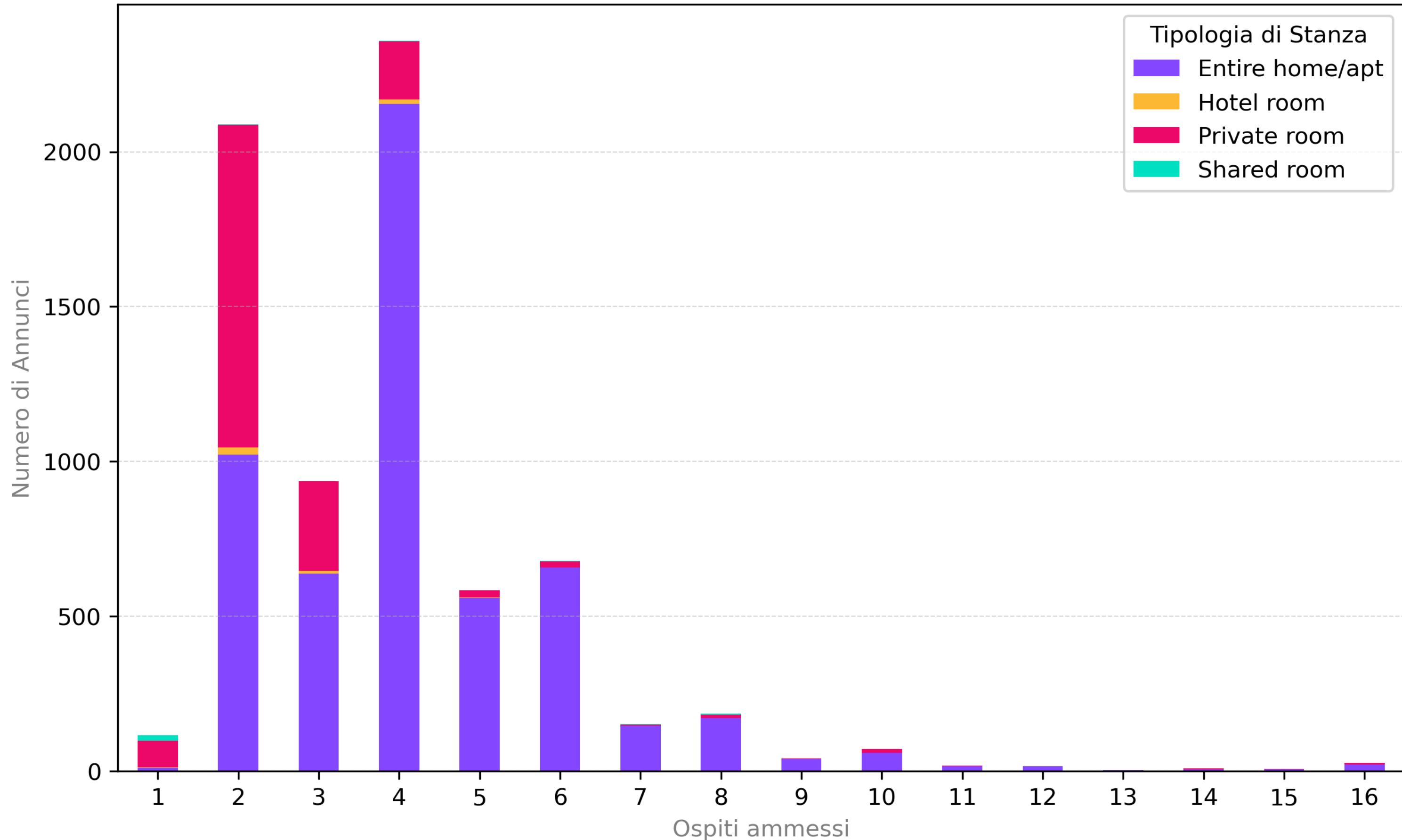
### TIPOLOGIE CON MINIME QUANTITÀ:

**LOCALI CON UN POSTO LETTO:** 24 ANNUNCI DI HOTEL E 1 ANNUNCIO DI STANZE CONDIVISE.

**LOCALI CON QUATTRO POSTI LETTO:** 14 ANNUNCI DI HOTEL E 1 ANNUNCIO DI STANZE CONDIVISE.



## Annunci per tipologia di stanza e posti letto - Palermo



## N° MASSIMO DI RECENSIONI PER POSTO LETTO

**OBIETTIVO DELL'ANALISI:** IDENTIFICARE L'ANNUNCIO CON IL MAGGIOR NUMERO DI RECENSIONI IN BASE AL NUMERO DI POSTI LETTO.

**RISULTATI PRINCIPALI:**

**MASSIMO NUMERO DI RECENSIONI PER:**

**DUE POSTI LETTO:** 548 RECENSIONI.

**QUATTRO POSTI LETTO:** 576 RECENSIONI.

**CINQUE POSTI LETTO:** 586 RECENSIONI.

**QUINDICI POSTI LETTO:** 738 RECENSIONI (PICCO MASSIMO).

**MINIME RECENSIONI:**

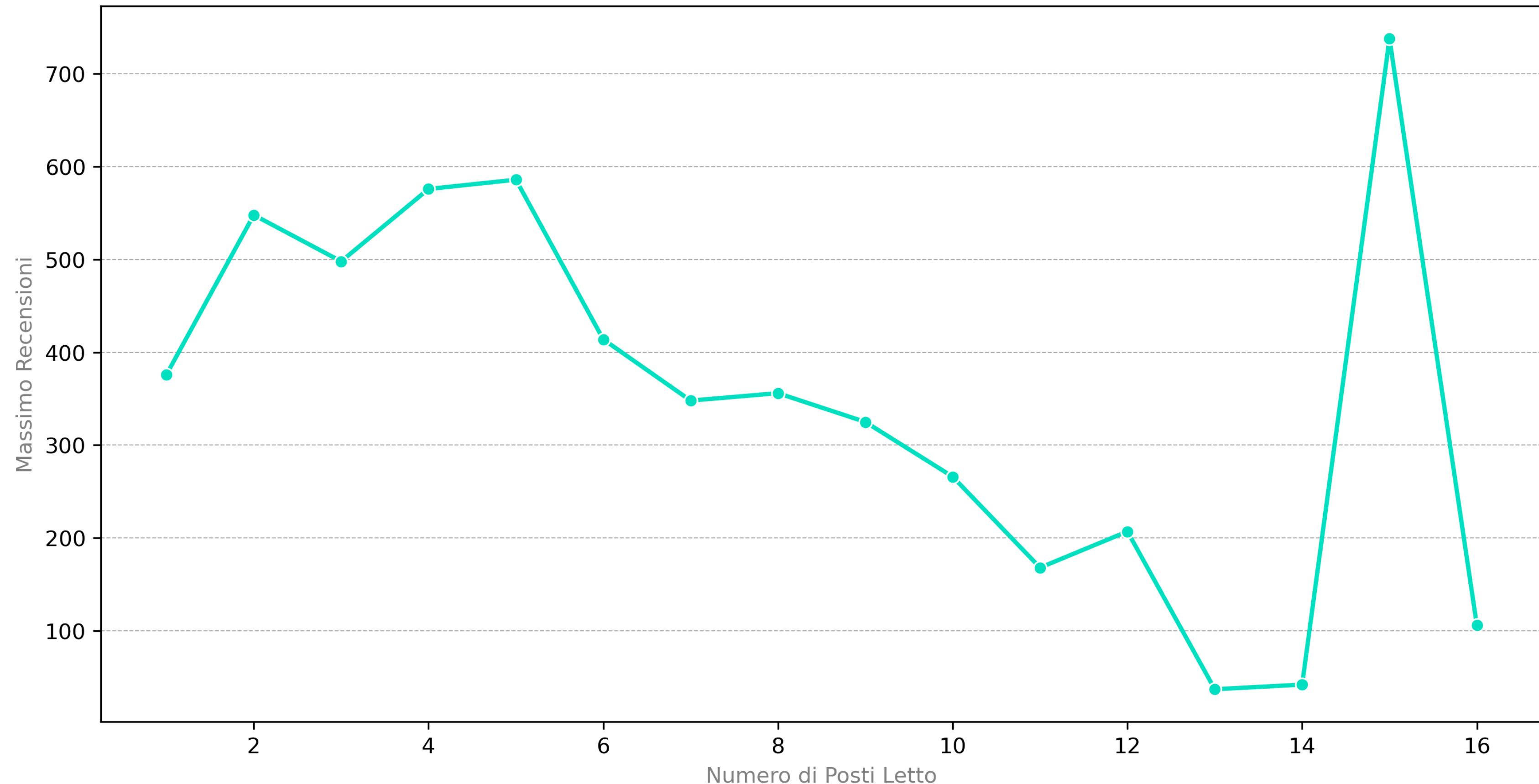
**TREDICI POSTI LETTO:** 37 RECENSIONI.

**QUATTORDICI POSTI LETTO:** 42 RECENSIONI.

**TREND OSSERVATO:** I PICCHI DI RECENSIONI PER SINGOLO ANNUNCIO NON SEGUONO NECESSARIAMENTE IL TREND GENERALE DEI POSTI LETTO.



## Massimo di Recensioni in Base ai Posti Letto - Palermo



## PREZZO MEDIO X POSTI LETTO

**DOMANDA DI RICERCA:** PERCHÉ C'È UN AFFLUSSO NEI LOCALI CON DUE E QUATTRO POSTI LETTO?

### POSSIBILI SPIEGAZIONI:

**PREFERENZE DI VIAGGIO:** MAGGIORE COMODITÀ NEL VIAGGIARE IN DUE O IN QUATTRO; VIAGGIARE IN TRE È MENO COMUNE.

**PREZZO DEI LOCALI:** I LOCALI CON DUE O QUATTRO POSTI LETTO POTREBBERO ESSERE PIÙ CONVENIENTI RISPETTO AD ALTRE TIPOLOGIE.

### ANALISI DEL PREZZO:

CREAZIONE DI UN GRAFICO SULLA STIMA MEDIA DEI PREZZI, RAGGRUPPANDO PER POSTI LETTO.

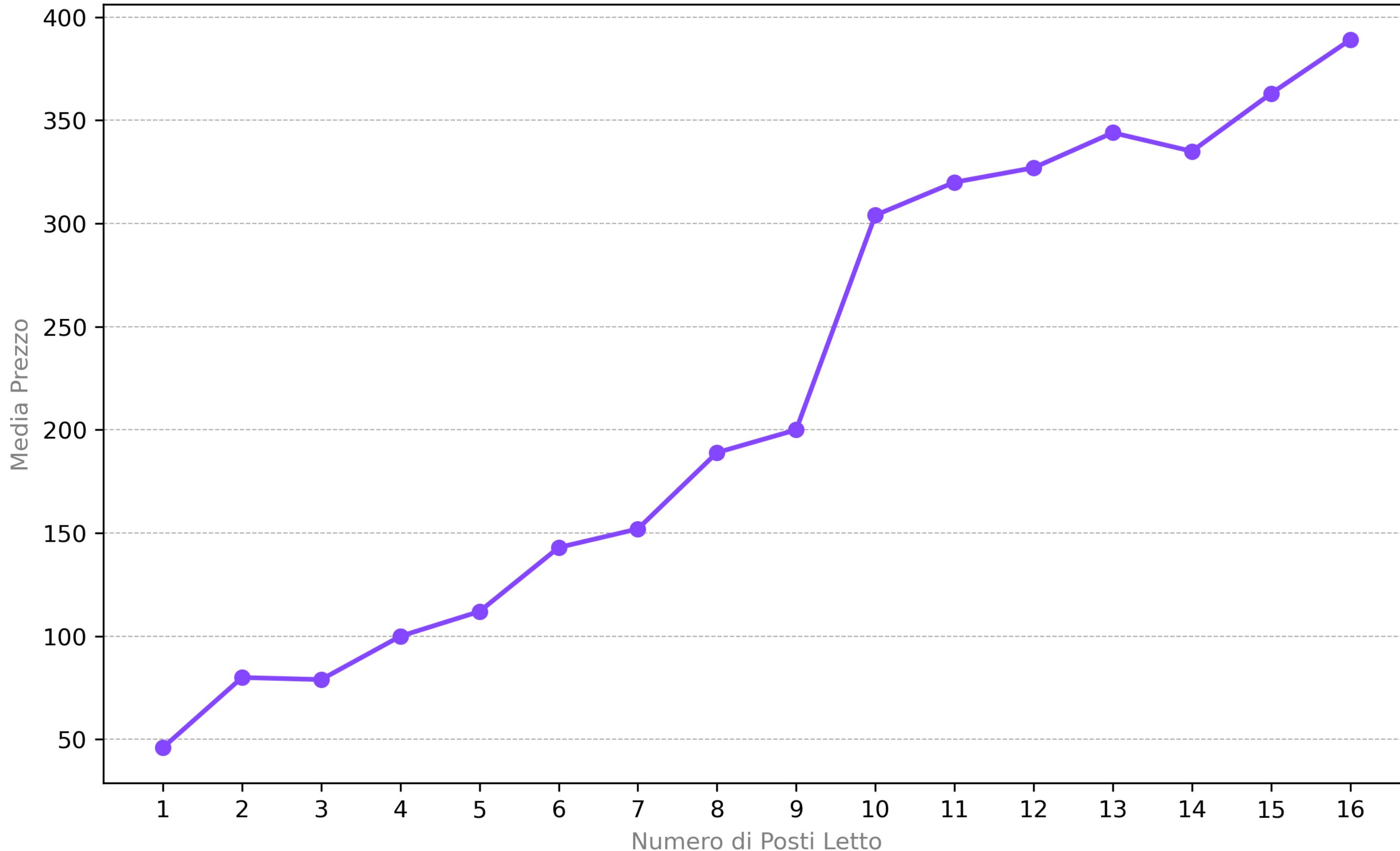
### OSSERVAZIONI DAI DATI:

NOTATA UNA CRESCITA SIGNIFICATIVA DEI PREZZI NELLA FASCIA TRA 9 E 10 POSTI LETTO, CON UN INCREMENTO DELLO 0,50.

**IMPLICAZIONI DELL'ANALISI:** COMPRENDERE LE PREFERENZE DI VIAGGIO E I PREZZI AIUTA A SPIEGARE IL TREND DEGLI AFFITTI NEI LOCALI CON DUE E QUATTRO POSTI LETTO.



## Media Prezzo per Posto Letto



# RAPPORTO NUMERO ANNUNCI PER POSTO LETTO

**DOMANDA DI RICERCA:** ESISTE UNA CORRELAZIONE TRA IL NUMERO DI RECENSIONI E LA DISPONIBILITÀ DI LOCALI CON DUE E QUATTRO POSTI LETTO?

## ANALISI CONDOTTA:

FOCUS SUL NUMERO DI ANNUNCI PER POSTO LETTO A PALERMO. RISULTATI INDICANO CHE LA MAGGIORE DISPONIBILITÀ DI LOCALI CON DUE E QUATTRO POSTI LETTO CONTRIBUISCE AI PICCHI DI RECENSIONI.

## DATI RILEVANTI:

**LOCALI CON DUE POSTI LETTO:** 2,088 ANNUNCI.

**LOCALI CON QUATTRO POSTI LETTO:** 2,358 ANNUNCI.

**LOCALI CON TRE POSTI LETTO:** 936 ANNUNCI.

## QUASI TOTALE MANCANZA DI ANNUNCI PER:

**13 POSTI LETTO:** 4 ANNUNCI.

**14 POSTI LETTO:** 9 ANNUNCI.

**15 POSTI LETTO:** 7 ANNUNCI.

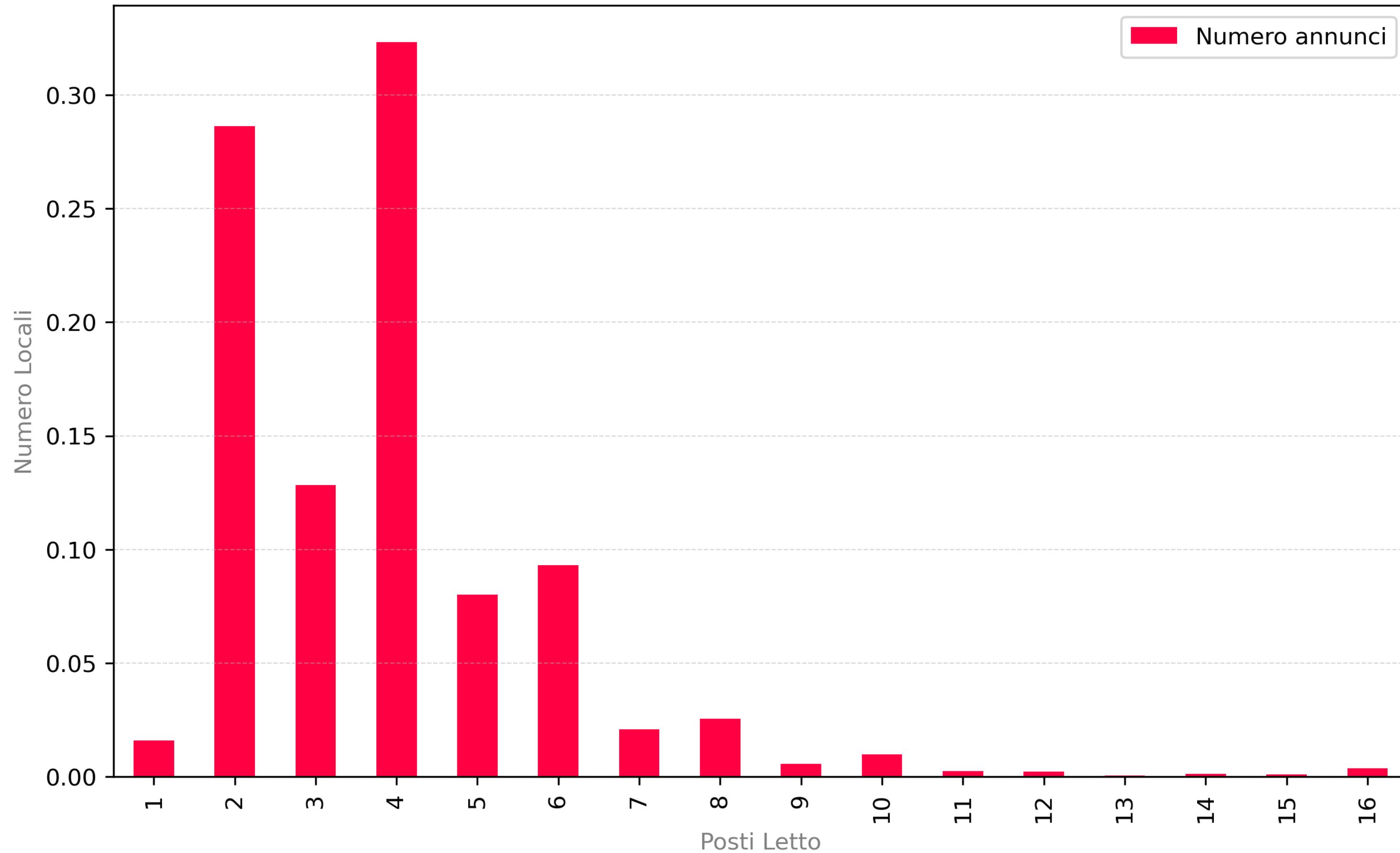
## OSSERVAZIONI DAI GRAFICI:

QUASI SOVRAPPOSIZIONE TRA IL GRAFICO DEL NUMERO DI RECENSIONI E QUELLO DEL NUMERO DI ANNUNCI PER POSTO LETTO.

**IMPLICAZIONI DELL'ANALISI:** LA DISPONIBILITÀ DI LOCALI CON DUE E QUATTRO POSTI LETTO È UN FATTORE CHIAVE NEL DETERMINARE IL NUMERO ELEVATO DI RECENSIONI RICEVUTE.



## Rapporto Numero locali e posti letto - Palermo



# GEOLOCALIZZAZIONE



## ZONE CON POSTI LETTO 2 E 4

**OBIETTIVO DELL'ANALISI:** IDENTIFICARE LE ZONE DI PALERMO CON LA MAGGIOR PARTE DEGLI ANNUNCI PER POSTI LETTO.

**UTILIZZO DI DATI GEOGRAFICI:** INCLUSIONE DI LONGITUDINE E LATITUDINE NELL'ANALISI PER VISUALIZZARE LA DISTRIBUZIONE DEGLI ANNUNCI.

### RISULTATI PRINCIPALI:

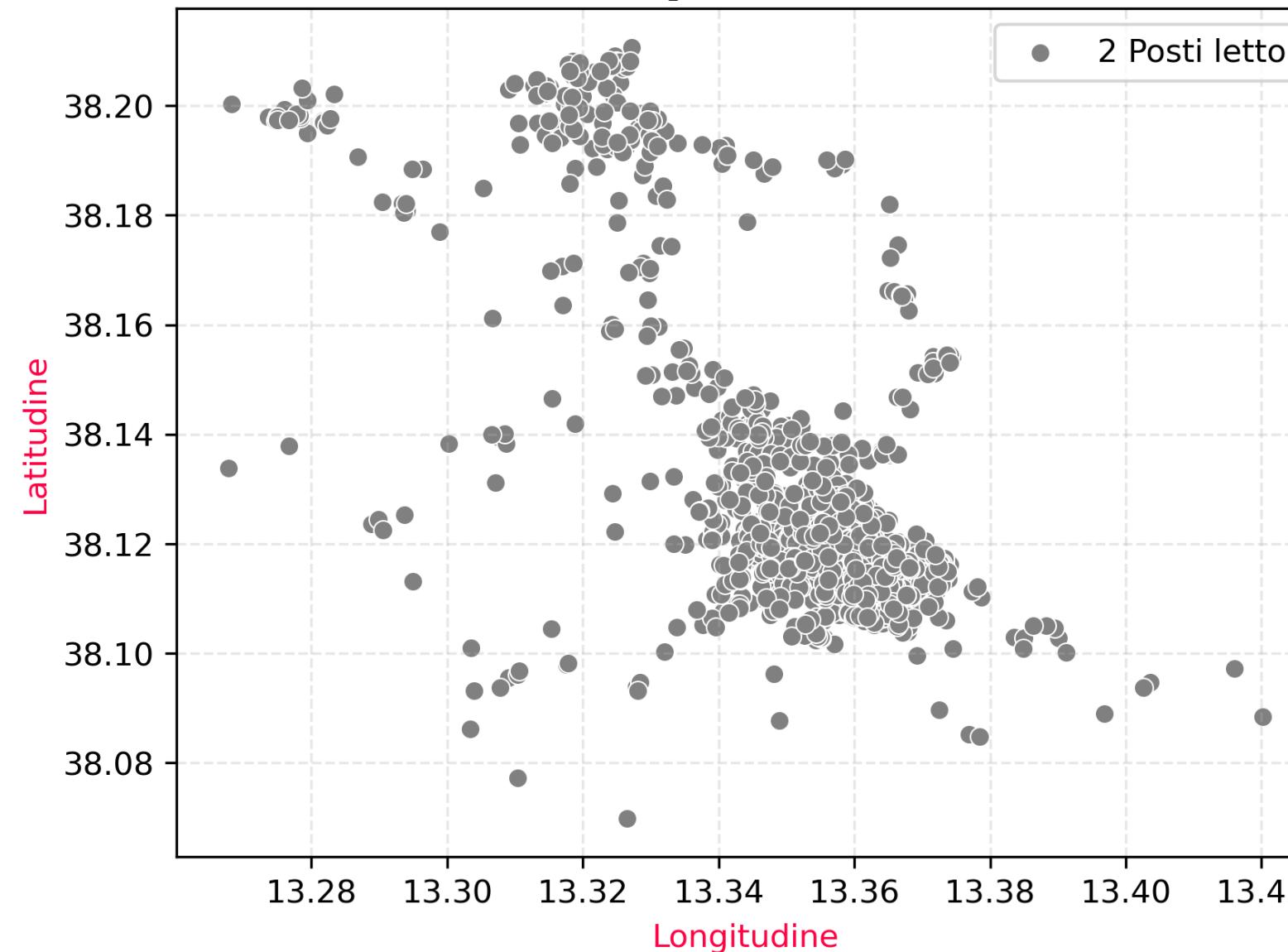
VISUALIZZAZIONE DI CLUSTER NELLE AREE CON UN'ALTA CONCENTRAZIONE DI LOCALI CON DUE E QUATTRO POSTI LETTO.

**CLUSTER PRINCIPALE:** SITUATO A SUD NELLA ZONA CENTRALE DI PALERMO, VERSO VIA VITTORIO EMANUELE E PIAZZA PRETORIA.

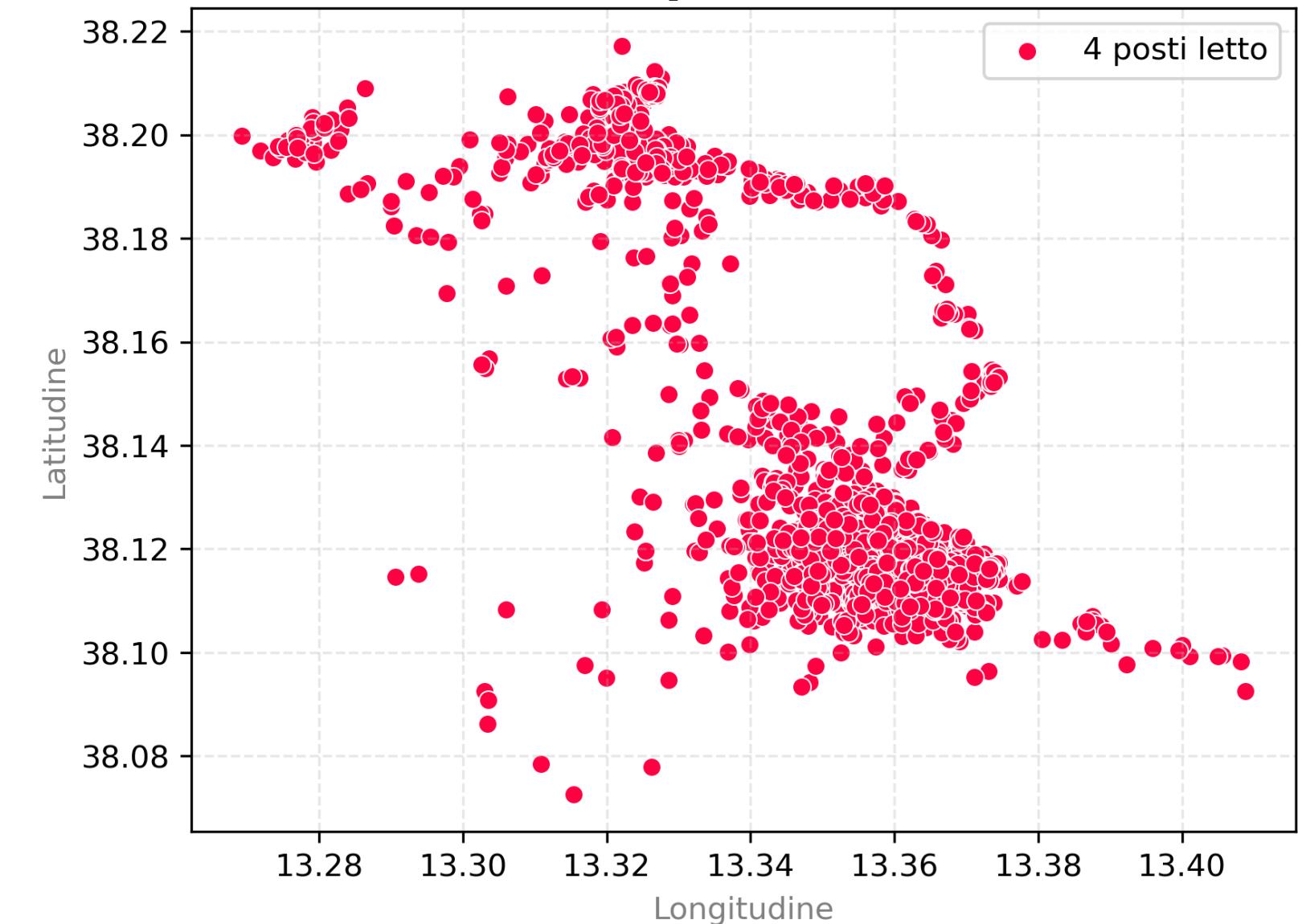
**CLUSTER MINORE:** PRESENTE A NORD NELLA ZONA CENTRALE DI PALERMO, CON UNA DISTRIBUZIONE PIÙ DISPERSA DI LOCALI.



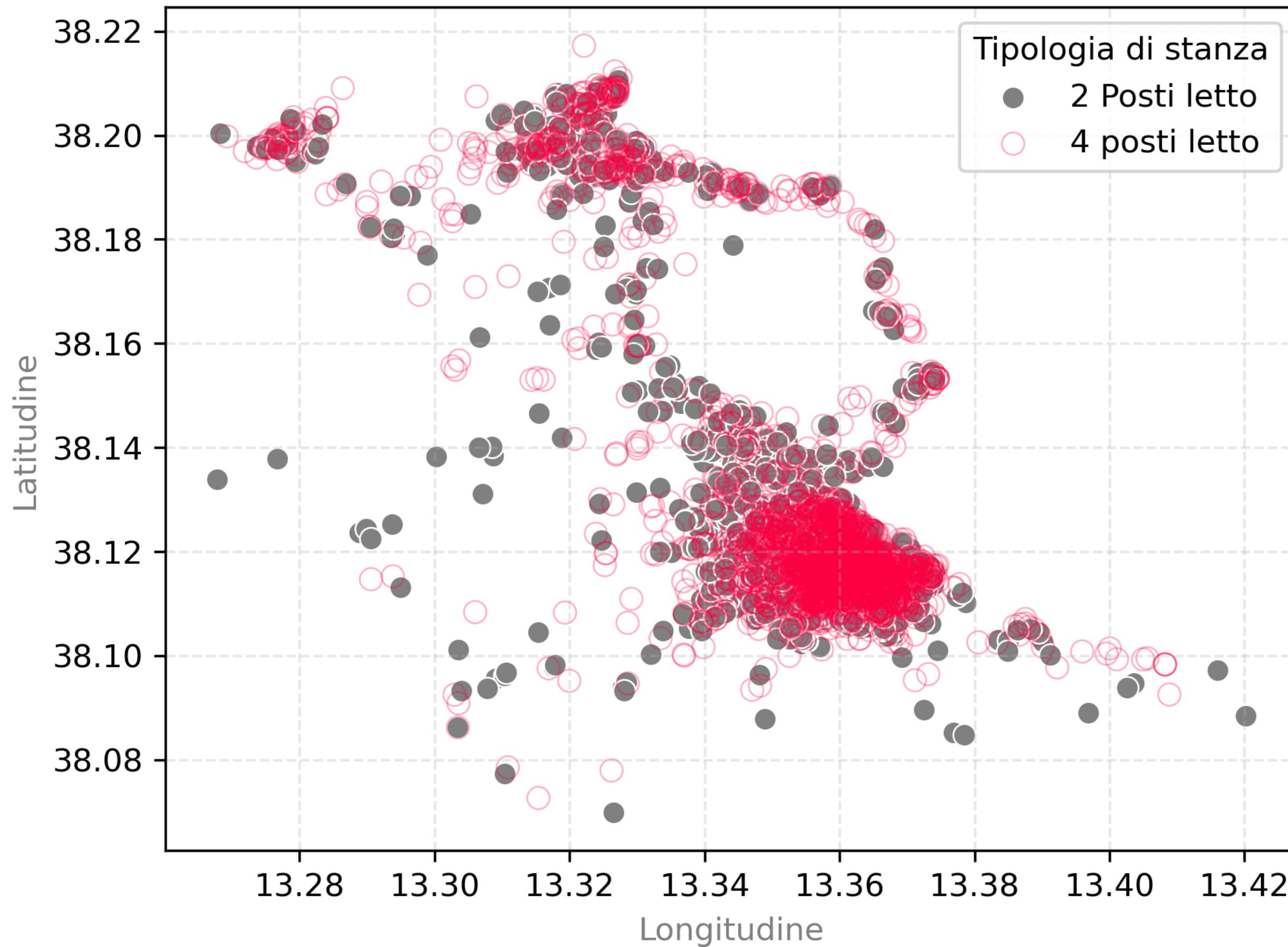
### **Locazione di 2 posti letto - Palermo**



### **Locazione di 4 posti letto - Palermo**



# Locazione di 2 e 4 posti letto - Palermo



# OCCIDENTE E ORIENTE

## TREND A CONFRONTO



# COMPARAZIONE DEI TREND AIR BNB

**OBIETTIVO DELL'ANALISI:** COMPARARE IL NUMERO DI RECENSIONI OTTENUTE IN DIVERSE LOCALITÀ PER IDENTIFICARE TREND RELATIVI AGLI ALLOGGI.

**FOCALIZZAZIONE INIZIALE:** ANALISI DELLE CITTÀ SICILIANE, IN PARTICOLARE CATANIA E PALERMO, EVIDENZIANDO UN AUMENTO DI RECENSIONI E DISPONIBILITÀ PER LOCALI CON DUE E QUATTRO POSTI LETTO.

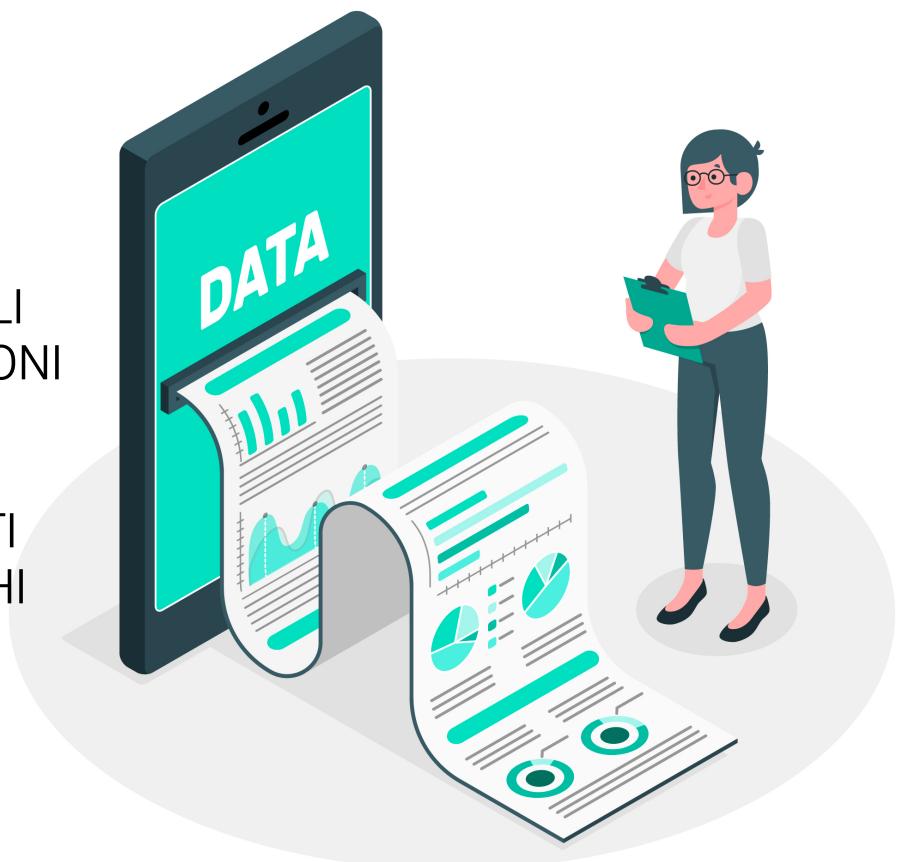
**ESPANSIONE DELL'ANALISI:** VERIFICA DEL TREND IN ALTRE LOCALITÀ OCCIDENTALI, COME LONDRA E NEW YORK CITY.

**RISULTATI INIZIALI:** RISCONTRATA UNA PREFERENZA PER GLI ALLOGGI CON DUE POSTI LETTO, CON UNA RIDUZIONE SIGNIFICATIVA PER QUELLI CON QUATTRO POSTI LETTO.

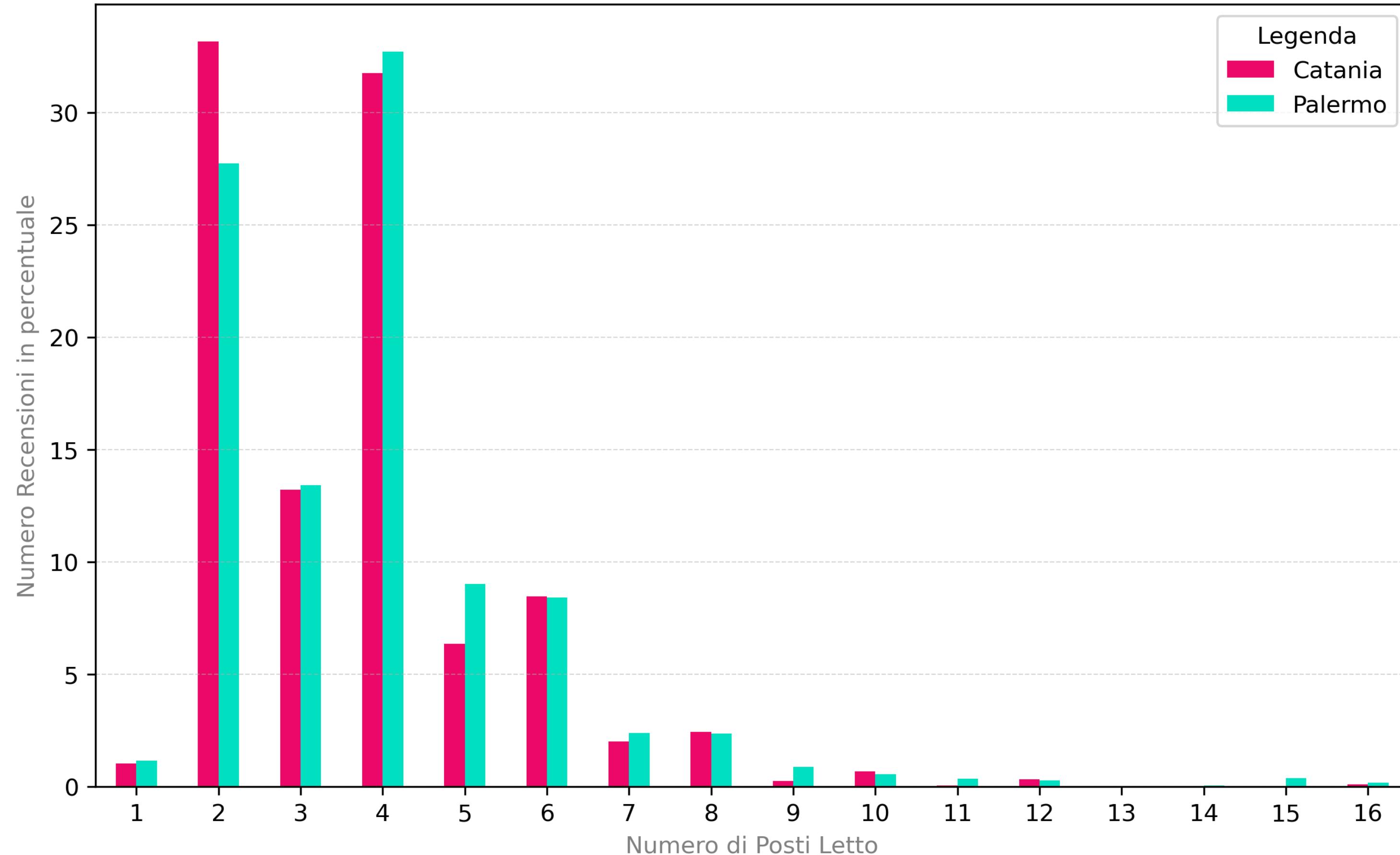
## CONFRONTO ORIENTALE:

STUDIO DEL TREND IN CONTESTI ORIENTALI, COME TOKYO E BANGKOK. SI È NOTATA UNA MAGGIORE PREFERENZA PER I LOCALI CON DUE POSTI LETTO E DISTRIBUZIONE OMOGENEA DI RECENSIONI PER ALTRE TIPOLOGIE.

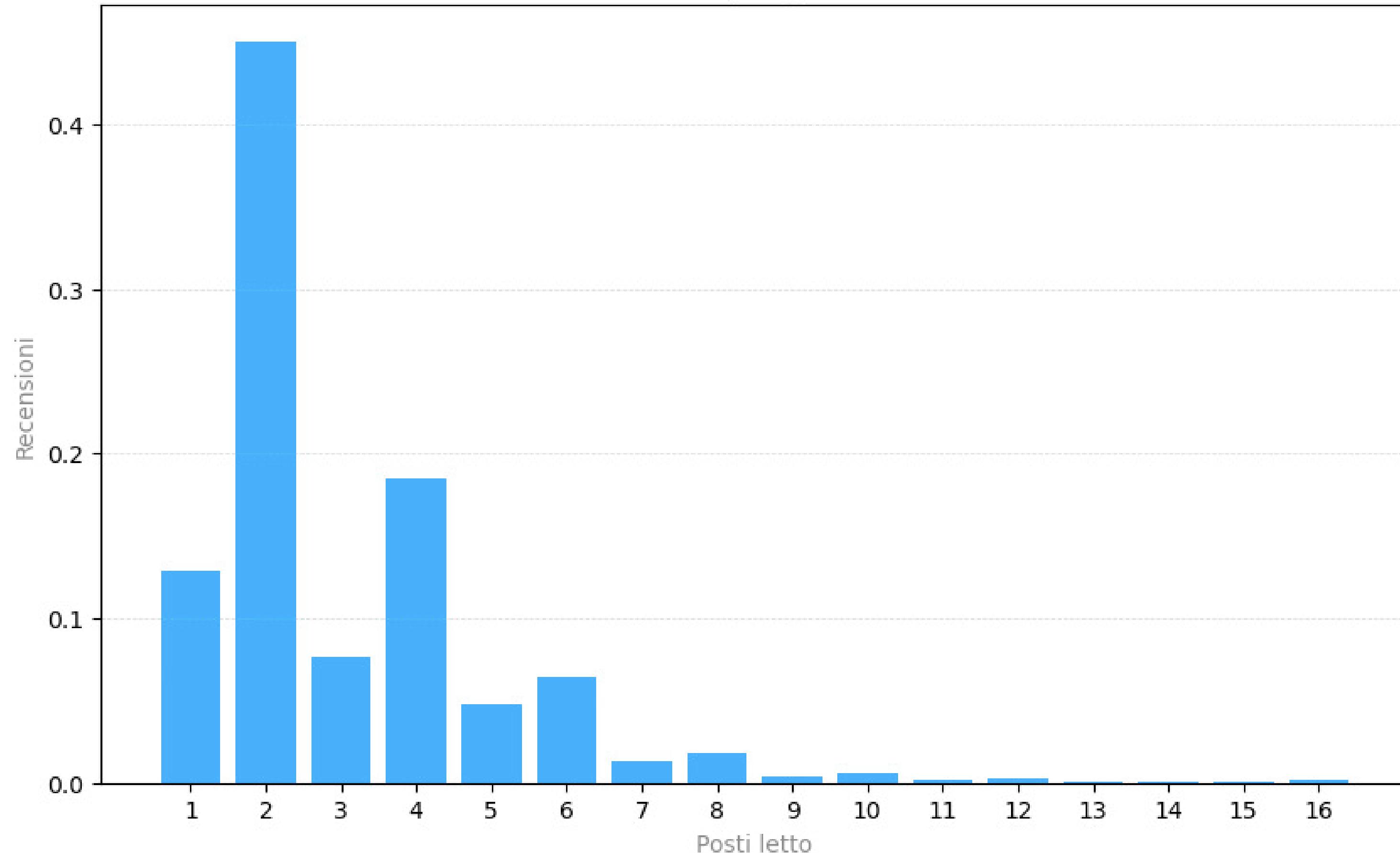
**OSSERVAZIONI FINALI:** CRESCITA NEI LOCALI CON CINQUE POSTI LETTO IN GIAPPONE, INDICANDO GRUPPI PIÙ NUMEROSI IN LUOGHI D'INTERESSE.



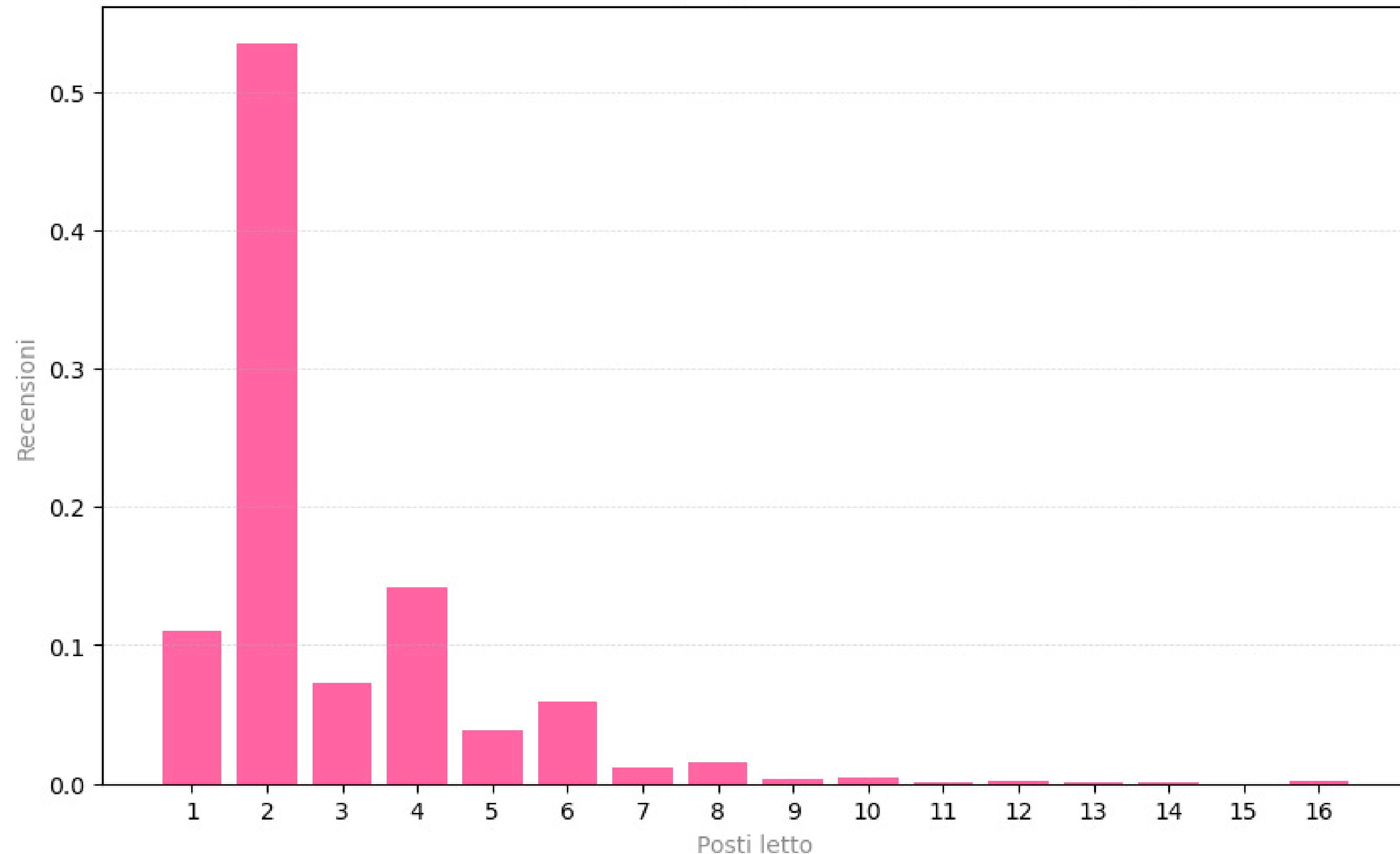
## Rapporto Numero Recensioni e posti letto in percentuale - Palermo - Catania



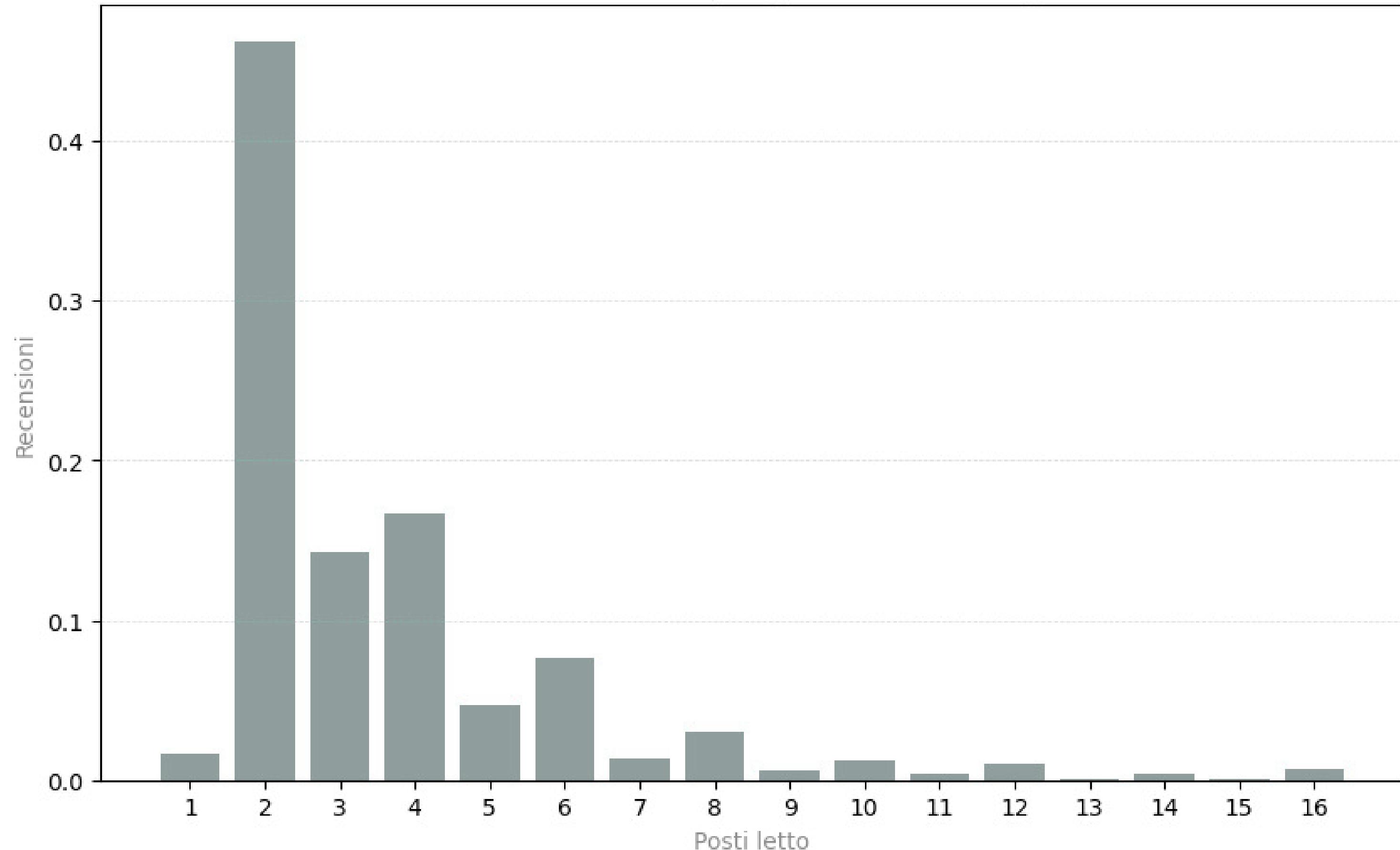
## **Numero recensioni per posti letto - Londra**



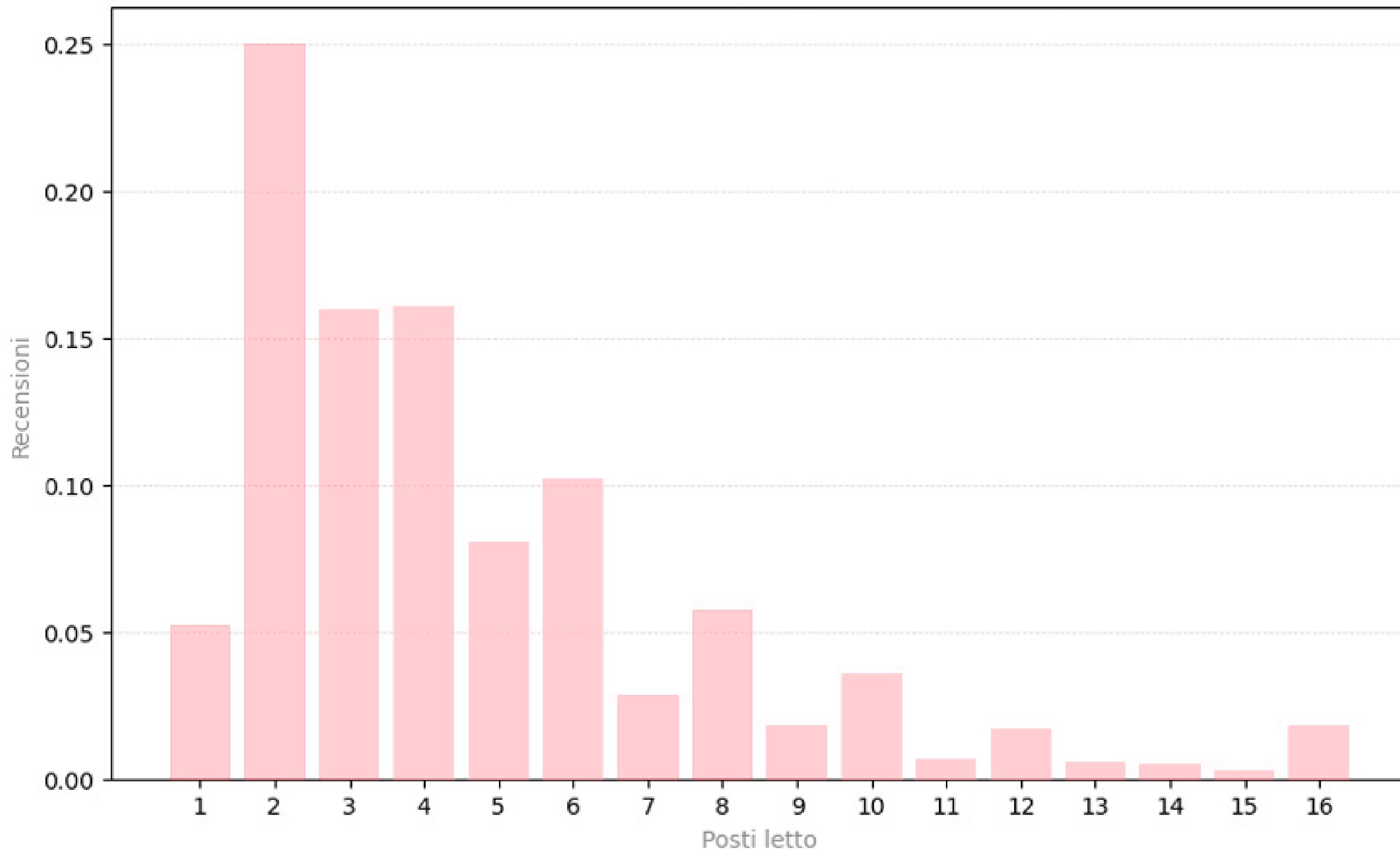
## Numero recensioni per posti letto - New York City



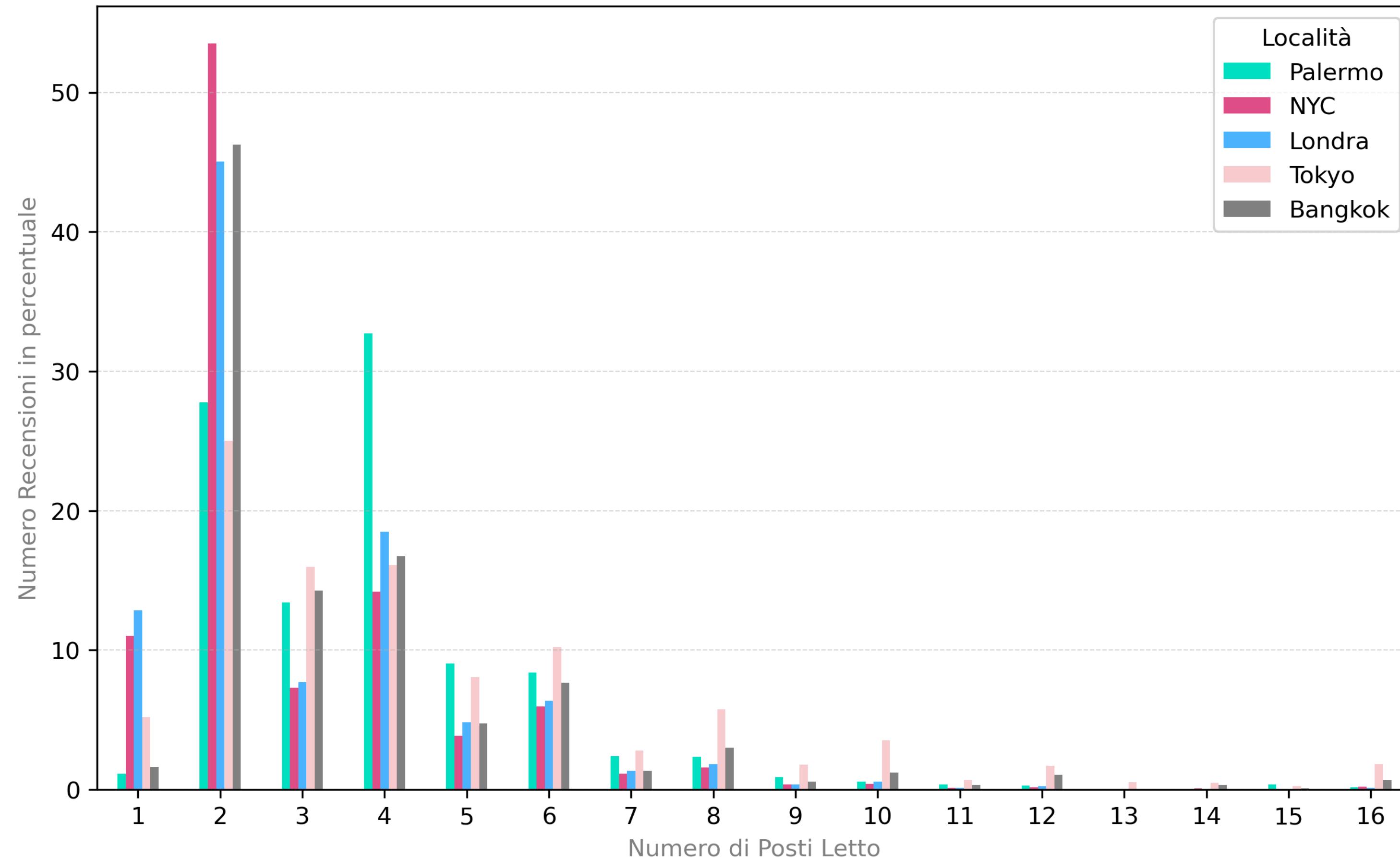
## Numero recensioni per posti letto - Bangkok



## Numero recensioni per posti letto - Tokyo



# **Numeri Recensioni e posti letto in percentuale - Palermo, NYC, Londra, Tokyo, Singapore**



# CONCLUSIONI

## WE BASTA



**TARGET DEMOGRAFICO:** LA MAGGIOR PARTE DEI TURISTI SU AIRBNB È COMPOSTA PRINCIPALMENTE DA COPPIE E SINGOLI.

**TREND OSSERVATO:**

PREFERENZA PER LOCALI CON DUE POSTI LETTO.

RAFFORZAMENTO DEL TREND IN DIVERSE CULTURE E CONTESTI GEOGRAFICI.

**ECCEZIONI:** NELLE CITTÀ SICILIANE, SI OSSERVA UNA PREFERENZA PER LOCALI CON QUATTRO POSTI LETTO, IN GIAPPONE INOLTRE VI SONO DELLE CRESCITE NEI LOCALI CON PIÙ POSTI LETTO.

**FATTORI ECONOMICI:**

GLI ALLOGGI CON DUE POSTI LETTO SONO GENERALMENTE PIÙ MODESTI RISPETTO AD ALTRE OPZIONI, RENDENDOLI PIÙ ACCESSIBILI.

**IMPLICAZIONI GLOBALI:**

LA PREFERENZA PER I LOCALI CON DUE POSTI LETTO È INFLUENZATA DA FATTORI ECONOMICI, CULTURALI E DI ESPERIENZA PERSONALE.



*The End*

A WALT DISNEY PRODUCTION