Rome Airbnb Market Analysis Report

This report provides a detailed analysis of Rome's Airbnb market, leveraging data from over 34,000 listings to explore the factors influencing pricing and guest preferences. The findings highlight the impact of location, property features, and guest reviews on pricing dynamics, offering insights into the market's structure and trends.

By examining the interplay of these factors, the report aims to provide actionable recommendations for hosts to optimize their offerings and for policymakers to support sustainable and equitable tourism. This analysis serves as a comprehensive resource for understanding Rome's evolving short-term rental market.

Preprocessing Summary

Objective:

The preprocessing phase ensured the dataset was clean, structured, and suitable for in-depth exploratory data analysis (EDA) and modeling. Key actions included handling missing data, addressing outliers, and engineering new features to enhance analytical depth.

Data Overview:

- **Original dataset:**
  - **34,061 rows** and **75 columns**
  - **Columns of Interest: 21**
    - **Listing Details:** *id*, *host_id*, *host_name*, *latitude*, *longitude*, *room_type*, *accommodates*, *bathrooms*, *bedrooms*, *price*, *amenities*
    - **Review Metrics:** *number_of_reviews*, *review_scores_rating*, *review_scores_accuracy*, *review_scores_cleanliness*, *review_scores_checkin*, *review_scores_communication*, *review_scores_location*, *review_scores_value*, *first_review*, last_review

- **Data Cleaning and Transformation:**
  - Converted *id* and *host_id* to string format
  - Transformed *price* to numeric by removing currency symbols
  - Reformatted *first_review* and *last_review* to datetime objects
  - Mapped listings to neighborhoods and quarters using GeoJSON data, replacing missing geographical data with "Unknown" for neighborhoods and "Unknown/NotCentralArea" for quarters in central Rome (column names: ***neighbourhood***, ***quarter***)

- **Handling Missing Data:**
  - **df_cleaned_basic** (keeping the integrity of the total count of listings):
    - removed rows with missing identifiers (id)
  - **df_cleaned_complete** (for further analysis and visualization):
    - Imputed numerical values (e.g., bathrooms, review scores) with medians
    - Flagged columns with >5% missing data (*review_scores_**) for exclusion or further analysis

Note: Missing values in the *review_scores** columns are closely linked to listings with *number_of_reviews* equal to 0, indicating that properties without reviews lack these scores.

Missing Values in Columns of Interest
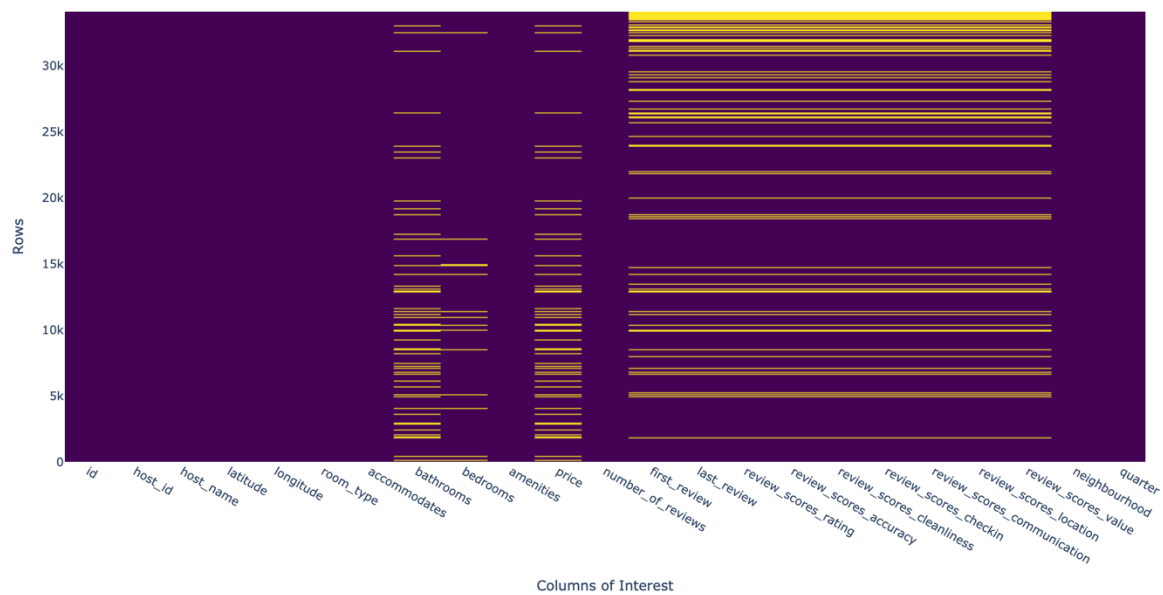
- **Outlier Detection and Removal:**
  - Removed listings with prices >€10,000, >10 bathrooms, or >1,000 reviews to ensure dataset reliability and ensure the dataset still reflects the expectedly high variability

- **Feature Engineering:**
  - *cost_per_person*: derived by dividing the *price* by the *accommodates* value to analyze affordability per guest.
  - *amenities* Transformation: Split *amenities* into a list for each listing and counted the frequencies of individual amenities for later analysis.

Summary:
Preprocessing steps successfully cleaned the dataset by addressing missing values, outliers, and data type inconsistencies.
Two versions of the cleaned dataset were created:
- *df_cleaned_basic:* Maintains the original number of listings with minimal cleaning.
- *df_cleaned_complete:* Prepared for analysis and visualization with missing values solved, outlier removal and feature engineering.

Report: EDA and Model Insights

Dataset Characteristics
- The dataset includes **34,061 Airbnb listings** managed by **18,588 hosts**, with a market heavily dominated by individual operators (**73%** manage a single property). Professional hosts, managing over 10 properties, make up only **1.3%** of the market.

Top 10 Hosts by Number of Listings
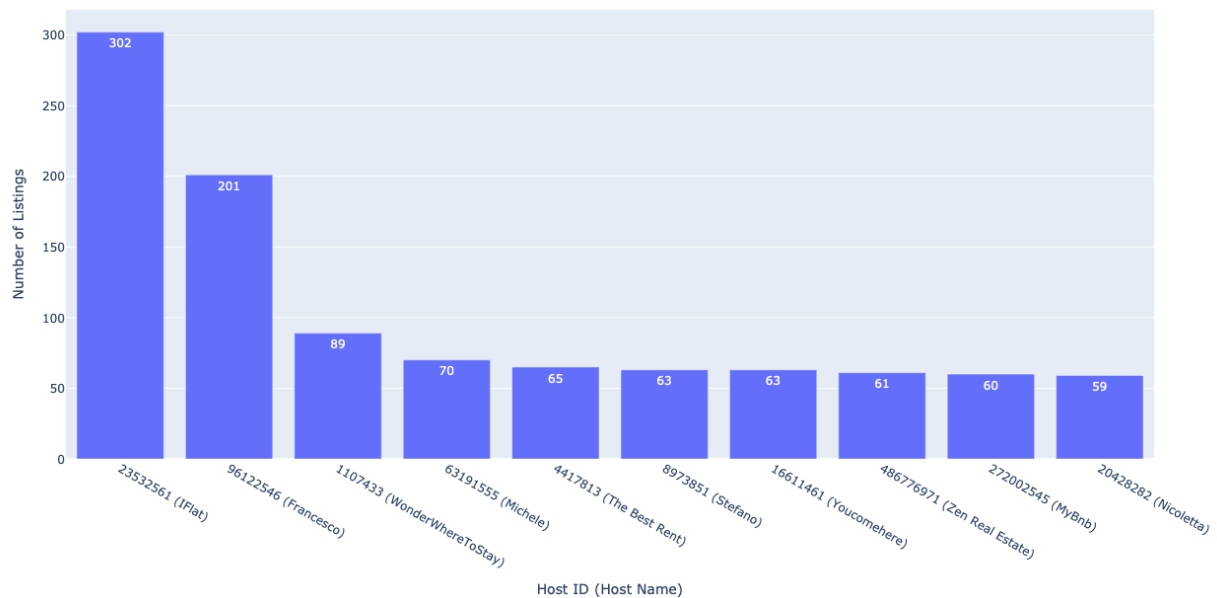
Room Types
- **Entire Homes/Apartments** dominate the market (**73.1%**), catering to families and groups, with an average price of **€227/night (€53/person** for groups).
- **Private Rooms** (**24.9%**) and **Shared Rooms** (**0.46%**) appeal to budget-conscious travelers, with respective average prices of **€150** and **€69** per night.

Neighborhood and Quarter Insights

- **Premium Neighborhoods: Centro Storico**, hosting **50.5%** of listings, commands the highest prices, especially in areas like Pigna and Trevi (up to 57% premium).
- **Affordable Options:** Peripheral areas such as San Giovanni and Parioli offer more budget-friendly accommodations.

Listings per Neighborhood
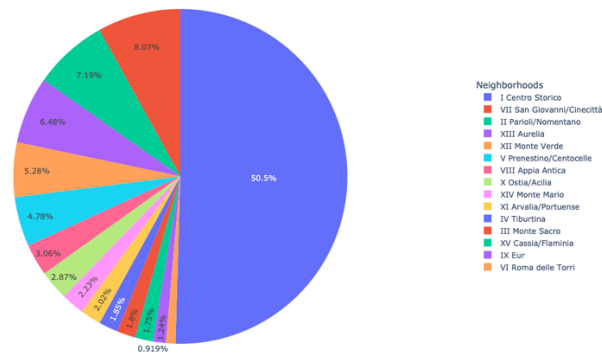
Reviews and Ratings

- Average guest rating is **4.77/5**, with **cleanliness** and **overall rating** being key factors influencing pricing.
- Smaller operator excel in engagement (e.g., MyBnb averages 15.22 reviews/month), while larger hosts face challenges in maintaining value ratings (<4.5).
- A weak positive correlation ($r = 0.07$, $p < 0,00$) exists **between price and review scores**.

Price vs. Review Scores Rating by Popularity
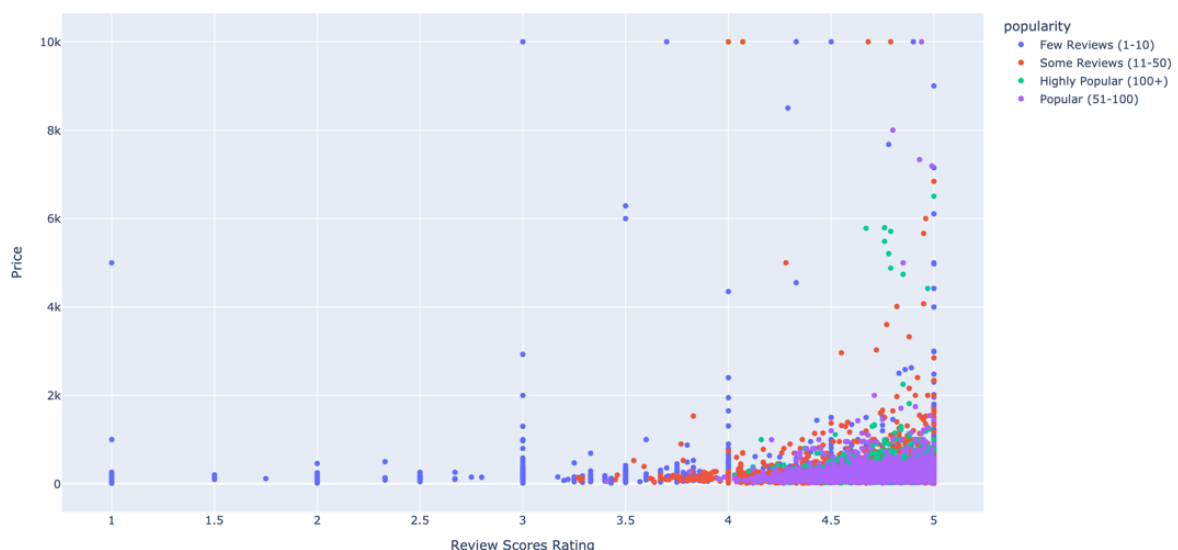
- A weak negative correlation ($r = -0.16$, $p < 0,00$) suggests that **listings with more reviews often receive slightly lower ratings.** This may reflect greater guest scrutiny as more reviews accumulate over time.

Correlation between Number of Reviews and Review Scores Rating
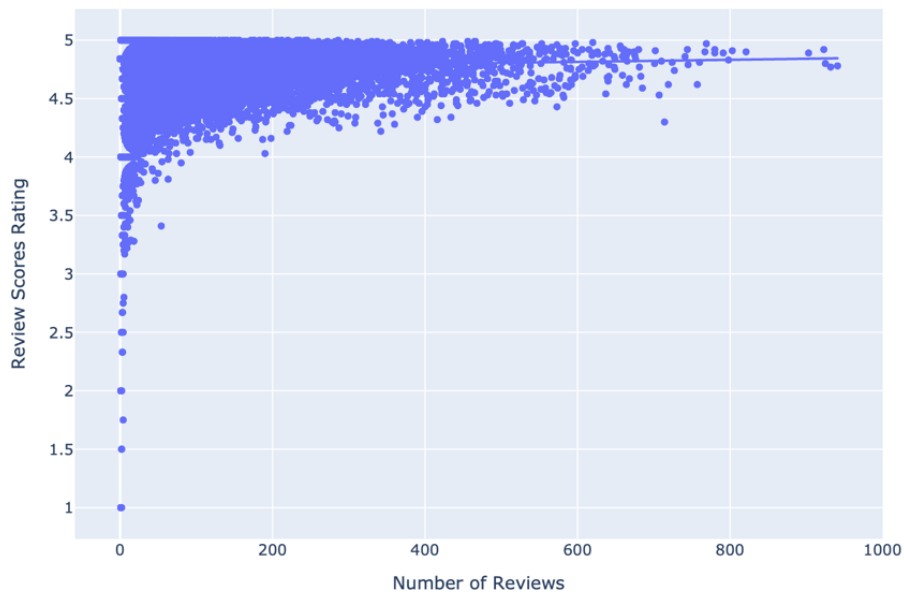
Amenities
- Common amenities like Wifi and kitchens are ubiquitous but do not significantly influence price.
- Premium features like **air conditioning** (+5%) and **shampoo** (+7.6%) are associated with higher pricing, while budget-associated amentities like hot water (-11.9%) and kitchens (-10.3%) correlate with lower prices.

Regression Model Analysis: Detaills

Model Overview

The OLS regression analysis explains **46,4%** of the variance in log-transformed price ($R^2 = 0,465$, $F = 210,5$, $p < 0,001$) using **55 predictors**. This indicates moderate explanatory power.

Below a breakdown of additional significant results:

Property Characteristics

- **Accommodates (+14.1%):** Each additional guest a property can accommodate results in a 14.1% increase in price, reflecting the premium for larger properties catering to families or groups.
- **Bathrooms (+12.4%) and Bedrooms (+7.2%):** Bathrooms contribute more strongly to pricing than bedrooms, highlighting their importance for comfort and utility, particularly in group stays.

Review Metrics

- **Overall Rating (+4.5%):** Each unit increase in overall rating corresponds to a 4.5% increase in price, emphasizing the value of guest satisfaction.
- **Cleanliness (+4.3%):** High cleanliness ratings are strongly associated with price premiums, reinforcing the importance of hygiene for guest preferences.
- **Value (-2.5%):** Lower value scores reduce prices, reflecting guest sensitivity to perceived affordability.
- **Number of Reviews (-8.7%):** Surprisingly, more reviews are associated with slightly lower prices. This may reflect older properties using competitive pricing to maintain bookings in saturated areas.

Amenities

- **Positive Influences:**
  - **Air Conditioning (+5.0%):** A crucial feature during Rome's warm months, significantly boosting pricing.
  - **Shampoo (+7.6%):** A small but impactful addition that signals attention to guest comfort.
  - **Dedicated Workspace (+2.9%):** Reflecting demand for remote work-friendly accommodations.
  - **First Aid Kit (+2.8%):** Enhances perceived safety and convenience.
- **Negative Influences:**
  - **Hot Water (-11.9%) and Kitchens (-10.3%):** Associated with budget listings, these features reduce perceived premiums.
  - **Dishes and Silverware (-7.2%):** Reflects their prevalence in lower-priced properties.
- Non-Significant Amenities: Standar features like WI-Fi and Essentials fo not significantly affect pricing, as they are common across all property types.

Location

- **High-Premium Quarters**: **Pigna (+57.9%), Trevi (+57.2%),** and **Campo Marzio (+50.4%)** command the highest premiums due to their historic and central locations.

- **Affordable Quarters: Testaccio (+3.9%)** reflects its less central position, offering more budget-friendly options.

Room Types
- **Entire Homes/Apartments (+18.6%):** Priced higher than shared accommodations, appealing to families or larger groups seeking privacy.
- **Shared Rooms (-56.1%):** Significantly cheaper, catering to budget-conscious travelers.

Conclusion

The analysis highlights the multi-faceted nature of pricing in Rome's Airbnb market. Larger properties, central locations, and premium amenities drive higher prices, while features common in budget listings (e.g., kitchens, hot water) tend to lower them. Review metrics, especially cleanliness and overall ratings, further influence pricing, with guest satisfaction playing a pivotal role. However, the weak correlation between price and number of reviews suggests that popularity alone does not determine pricing power.

Actionable Insights

For Hosts

- **Focus on Cleanliness and Ratings:**
    - Why It Matters: Cleanliness and overall ratings are significant predictors of price, with high cleanliness ratings adding up to a 4.3% premium. Guests consistently value hygiene, especially in post-pandemic travel.
    - Action: Regularly invest in professional cleaning services and communicate these efforts in your listing. Encourage satisfied guests to leave reviews to improve overall ratings and attract higher-paying travelers.

- **Invest in Premium Amenities:**
    - Why It Matters: Features like air conditioning (+5%), shampoo (+7.6%), and dedicated workspaces (+2.9%) enhance guest convenience and justify higher prices. In Rome's warm climate, air conditioning is especially valued, while workspaces cater to digital nomads.
    - Action: Install or upgrade air conditioning systems and consider small but impactful additions like luxury toiletries. If targeting business travelers, create a designated workspace with a desk, chair, and fast Wi-Fi.

- **Highlight Location Benefits:**
    - Why It Matters: Listings in premium neighborhoods (e.g., Pigna, Trevi) can command up to 57% higher prices due to their centrality and historic appeal. Even properties in more affordable neighborhoods like Testaccio (+3.9%) can benefit from emphasizing unique local attractions.
    - Action: For central properties, showcase proximity to landmarks, transportation hubs, and cultural attractions in your listing description. For peripheral locations, highlight value-added aspects like quieter surroundings, spacious accommodations, or access to hidden gems.

- **Target Specific Traveler Segments:**
    - Why It Matters: Entire homes and apartments cater to families and groups, commanding 18.6% higher prices, while private rooms are better suited for solo travelers or couples.
    - Action: Customize listings to appeal to target demographics. For family-friendly properties, emphasize child-friendly amenities like cribs or play areas. For couples or solo travelers, focus on cozy spaces with privacy and proximity to romantic or cultural experiences.

For Researchers and Policymakers

- **Promote Undervalued Areas:**
    - Why It Matters: Peripheral neighborhoods like Testaccio and Parioli provide more affordable accommodations but may be underutilized by tourists focused on central areas.
    - Action: Develop marketing campaigns or local guides highlighting the unique cultural, culinary, and historical aspects of these areas. Improve infrastructure and accessibility to enhance their appeal to visitors.

- **Incorporate Seasonal Trends in Future Models:**
    - Why It Matters: Seasonal demand fluctuations can significantly impact prices and booking patterns.
    - Action: Conduct further analysis to assess how peak and off-peak seasons influence pricing and incorporate these insights into planning for both hosts and local tourism policies.
- **Evaluate Location-Based Price Drivers:**
    - Why It Matters: While central neighborhoods like Pigna and Trevi command higher prices, proximity to secondary attractions or transport hubs could also influence guest decisions.
    - Action: Use geographic data to identify emerging areas with growth potential. Provide incentives for hosts to improve listings in these neighborhoods, potentially reducing pressure on central areas.

- **Monitor and Address Equity in the Market:**
    - Why It Matters: The dominance of premium neighborhoods in pricing trends may exclude lower-income travelers or reduce opportunities for local residents.
    - Action: Encourage the development of affordable short-term rental options in less saturated neighborhoods. Balance tourism growth with local housing needs through thoughtful policy interventions.