Airbnb Analysis: Key Insights and Process

for

Warmer

By

Tatyana Parks

1. Dataset and Motivation

For this project, I analyzed the 2024 New York City Airbnb dataset from Inside Airbnb, containing over **36,000 listings** and **975,000 guest reviews**. I chose this dataset because it was strategically relevant to Warmer's platform dynamics. The analysis reminded me of how Warmer connects experts with clients: Airbnb hosts must leverage their hospitality expertise to provide pleasant stays, with certain indicators (communication scores, amenities, reviews) signaling trustworthiness and quality. Similarly, Warmer experts must apply their lived experience and expertise in specific issues to help clients navigate challenges they're struggling with. In Airbnb, success depends on demonstrating competence through tangible signals that build user confidence. Similar trust-building mechanisms could be valuable for Warmer's platform if not already implemented.

2. Questions Explored

1. **Primary Question:** What variables seem most strongly correlated with the outcome of interest?
2. **Supporting Question:** Which neighborhoods consistently outperform others?

3. Methodology and Data Cleaning

- **Outlier Management:** The dataset contained extreme values. For example, some listings reported maximum stays of **2,147,483,647 nights** (likely a data entry error or system default), while others listed nightly prices of **$50,184**. Following methodological practices from published research on short-term rental markets, I applied conservative filtering: listings above $1,000/night were excluded, and maximum stays were capped at 365 nights to focus on genuine short-term rentals rather than luxury estates or long-term leases.
- **Amenities Processing:** I transformed inconsistent amenity strings into standardized binary features through custom parsing functions. This two-step approach first cleaned the raw data and then normalized entries (e.g., all TV variants became "tv").
- **Sentiment Analysis:** I used VADER sentiment analysis on guest reviews to create listing-level sentiment scores, providing a quantitative measure of guest satisfaction beyond star ratings.
- **Statistical Approach:** I employed Spearman rank correlation to handle non-normal distributions typical in marketplace data.

4. Key Findings

- **Communication is the Foundation of Success:** Host communication was the strongest predictor of guest satisfaction ($\rho = 0.73$ with ratings, $\rho = 0.47$ with sentiment), highlighting responsiveness as central to positive outcomes.
- **Context Determines Value:** Amenity value was hyper-local. A hair dryer in Little Italy correlated strongly with revenue ($\rho \approx 0.78$, ~35% prevalence) but showed no effect in the Theater District (~95% prevalence). Similar effects appeared for coffee makers in Ozone Park ($\rho \approx 0.70$) and microwaves in Richmond Hill ($\rho \approx 0.59$).

- **The Revenue-Quality Paradox:** High revenue did not equal better reviews. All review metrics had weak negative correlations with revenue (-0.02 to -0.11), suggesting location or demand may outweigh satisfaction.
- **Geographic Hotspots and Coldspots:** Neighborhood mattered. Greenwich Village, Theater District, and Financial District topped $50k average revenue, while Throgs Neck and Morningside Heights fell below $11k. Research (Brinkman 2022) indicates proximity to tourist attractions likely explains these gaps. A regression including distance-to-attractions would clarify this further.
- **Trust Signals Matter:** Superhost status correlated with communication, ratings, and revenue ($\rho \approx 0.22–0.28$). While moderate, the consistency validated Airbnb's recognition system as a quality marker.

5. Process Insights and Challenges

1. **Iterative Refinement:** Early amenity parsing produced messy categories like "50 Inch HDTV." Refinement through normalization produced usable features, underscoring the value of iteration.
2. **Statistical Sophistication:** Most correlations were modest, reinforcing that marketplace outcomes hinge on complex interactions. IQR flagged outliers, but business logic often guided final cutoffs.

6. Broader Implications

The Airbnb analysis highlights patterns that may be relevant for understanding how trust and reliability function in marketplaces like Warmer.

- **Communication translates across platforms:** Just as Airbnb hosts often succeed by responding quickly and reliably, it may be worth exploring whether similar responsiveness from Warmer experts could help build stronger client relationships. The observed correlation ($\rho = 0.73$) between communication and satisfaction suggests that this kind of signal might be an indicator of success in marketplaces.
- **Context determines value:** The hair dryer example (valuable in Little Italy but meaningless in Theater District) parallels how expert specialization matters at Warmer. An expert tagged for "caregiving" may need different sub-expertise for young children versus elderly parents. The same credential has a different value depending on the client context.
- **Recognition systems work:** Even moderate correlations ($\rho \approx 0.22-0.28$) show Superhost status consistently predicts quality. This highlights the value of expert credibility systems. At Warmer, this could take shape through client feedback, tenure, or peer endorsement.
- **The quality-volume tension:** High-revenue Airbnb properties sometimes showed weaker review metrics, suggesting potential tradeoffs between scale and satisfaction. For Warmer, this raises questions about whether high-engagement topic categories maintain the same satisfaction levels as specialized, lower-volume areas.