

Text Mining Analysis of Airbnb Listings in New York and Sydney

Authored by SHIH-HSUAN, YAN

Executive Summary

This analysis utilizes text mining techniques to analyze descriptions written by Airbnb hosts for listings in New York and Sydney, uncovering key insights that influence guest preferences and booking decisions. The analysis identifies critical factors such as location desirability, property features, and host-expressed sentiments which significantly influence booking behaviors. By encouraging hosts to enhance listings to emphasize these valued features and addressing common areas of concern, Airbnb can improve guest satisfaction, optimize listing performance, and increase booking rates. This report also provides actionable recommendations to refine marketing strategies and host management practices, ensuring they align with the distinct preferences identified in two of Airbnb's major markets.

Price and Listing Analysis (Figure 1)

An analysis of pricing variations across multiple locations revealed notable currency differences. Despite these discrepancies, the significant similarities in listing volumes and room type ratios between New York and Sydney were particularly striking. This similarity is critical as it provides a robust basis for direct comparison, allowing for more accurate cross-city analyses in subsequent text-based studies. Focusing on these two cities enables a deeper understanding of market dynamics in major urban centers, thereby facilitating more targeted and effective insights for Airbnb's strategic decisions.

Textual Content Analysis (Figure 2)

Text mining was used to show the most frequent words in listing discussions across both markets. Descriptions with words like 'apartment', 'amazing', and 'access' indicate general positive attributes valued in both New York and Sydney markets. City-specific analysis using TF-IDF methods to show key terms that differentiate New York and Sydney in the marketplace. In New York, tokens such as 'Manhattan', 'Brooklyn', and 'subway' appear frequently which emphasize the importance of urban experience and transportation accessibility. For Sydney, tokens like 'Bondi', 'harbour', and 'CBD' appear the most which show a strong preference for iconic locations and proximity to major attractions, suggesting that guests are particularly attracted to listings that highlight these features. These insights can guide Airbnb to prioritize and market properties that align with these distinct preferences, potentially increasing appeal and bookings.

Phrase Analysis (Figure 3)

Expanding on individual words, analysis of bigrams and quadrigrams offers deeper insight into the textual nuances within host descriptions. In New York, phrases like 'the heart of Williamsburg' and 'walk to Central Park' suggest that hosts value not only the location but also the cultural and recreational opportunities associated with them. In Sydney, phrases such as 'walk to Bondi Beach' and 'the heart of Sydney' indicate that hosts promote tourist-friendly areas that blend urban and natural experiences. This analysis highlights the importance of marketing properties based on their location and the lifestyle they offer, which can be a decisive factor for guests choosing between listings.

Sentiment Analysis (Figure 4)

Sentiment analysis using 'Bing words' identified a mix of positive and negative descriptors in the hosts' descriptions that provide insights into guest satisfaction drivers. Notably, the analysis revealed that positive terms significantly outnumber negative terms in both New York and Sydney, emphasizing a predominantly favorable portrayal of listings by hosts. Positive terms like 'quiet', 'spacious', and 'modern' across both cities suggest these are highly valued features that enhance guest experiences. Conversely, negative terms such as 'noise', 'limited', and 'complex' reveal common pain points that could deter potential guests. Encouraging hosts to address these negative aspects in their descriptions and improve the actual attributes could mitigate dissatisfaction and improve the guest experience. For instance, providing accurate descriptions and setting realistic expectations about property noise levels and space limitations can prevent negative reviews and increase transparency. The overall higher incidence of positive descriptors indicates that while there are areas for improvement, hosts generally succeed in creating appealing listings, which is beneficial for maintaining high guest interest and satisfaction levels.

Conclusion and Recommendations

This comprehensive text mining analysis highlights specific textual elements in Airbnb host listings that might affect guest choice and satisfaction. By examining property features and marketing strategies with the preferences revealed through text analysis, Airbnb can enhance its competitive edge. Identifying positive and negative attributes can help to improve guest satisfaction, increased bookings, and enhanced loyalty. Additionally, refining property listings to highlight desirable features, encouraging hosts to improve features that receive negative feedback, and using location-based marketing to attract guests looking for specific experiences in New York and Sydney.

Reference

OpenAI. (2023). Personal communication [with ChatGPT-4]

Appendix

Figure 1

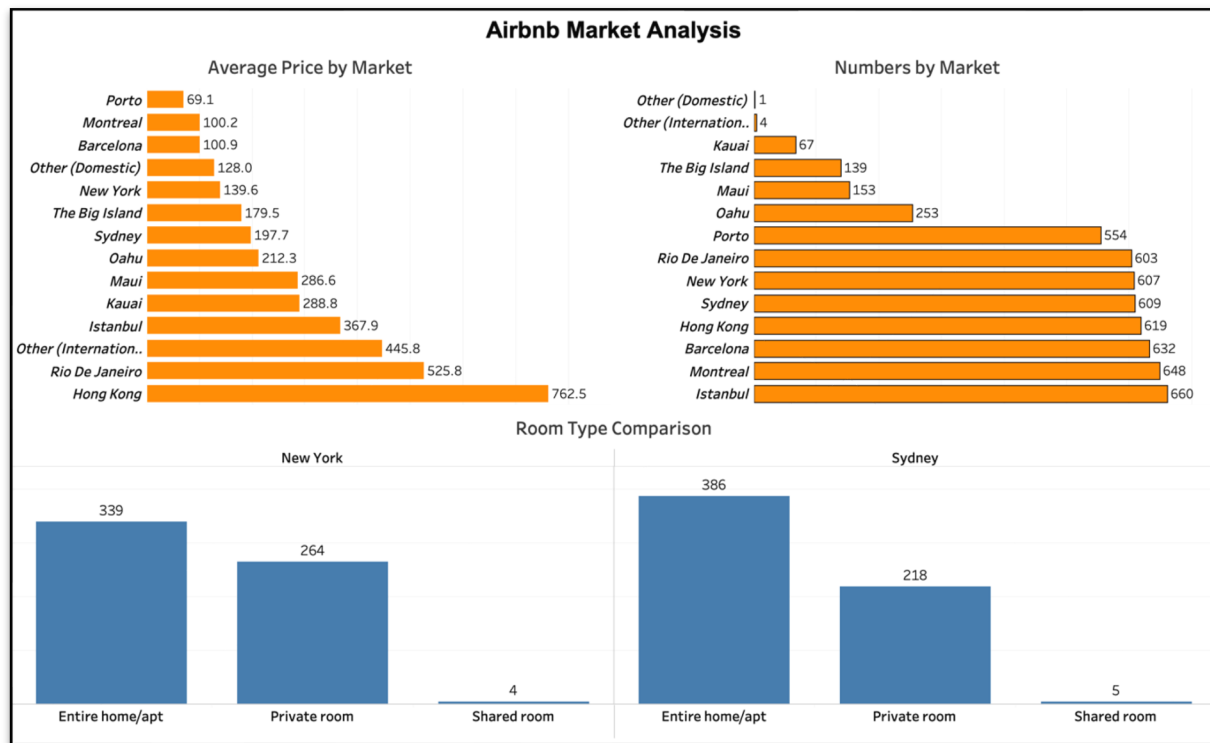


Figure 2

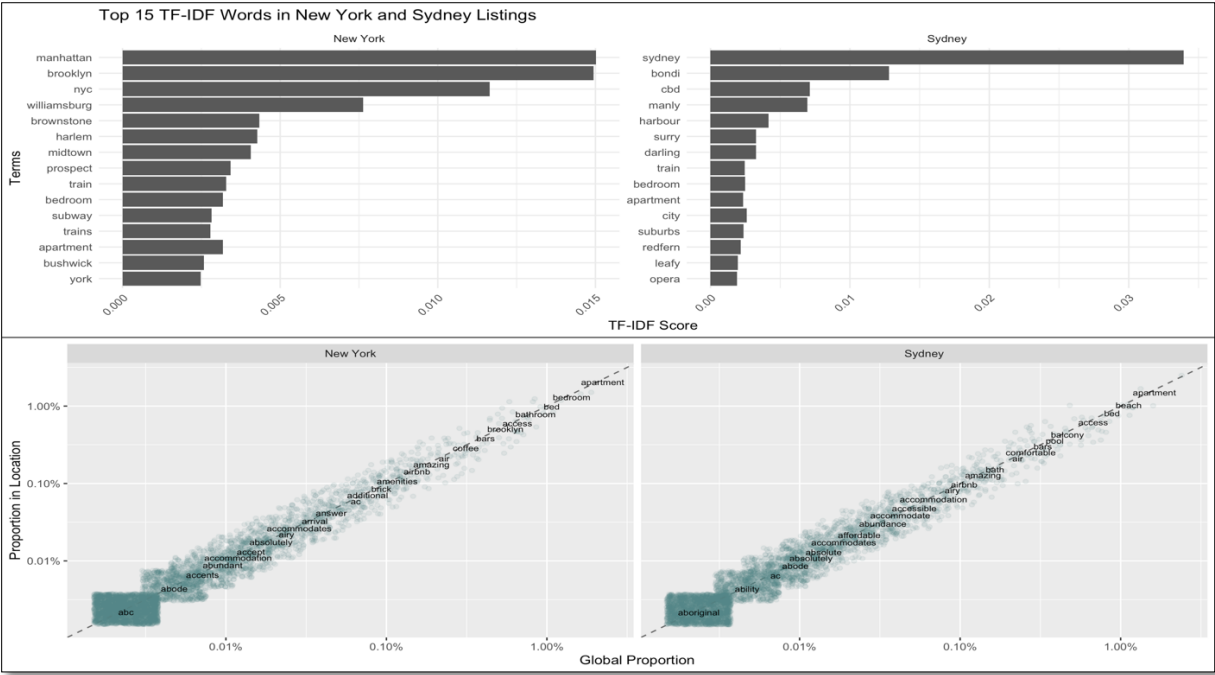


Figure 3

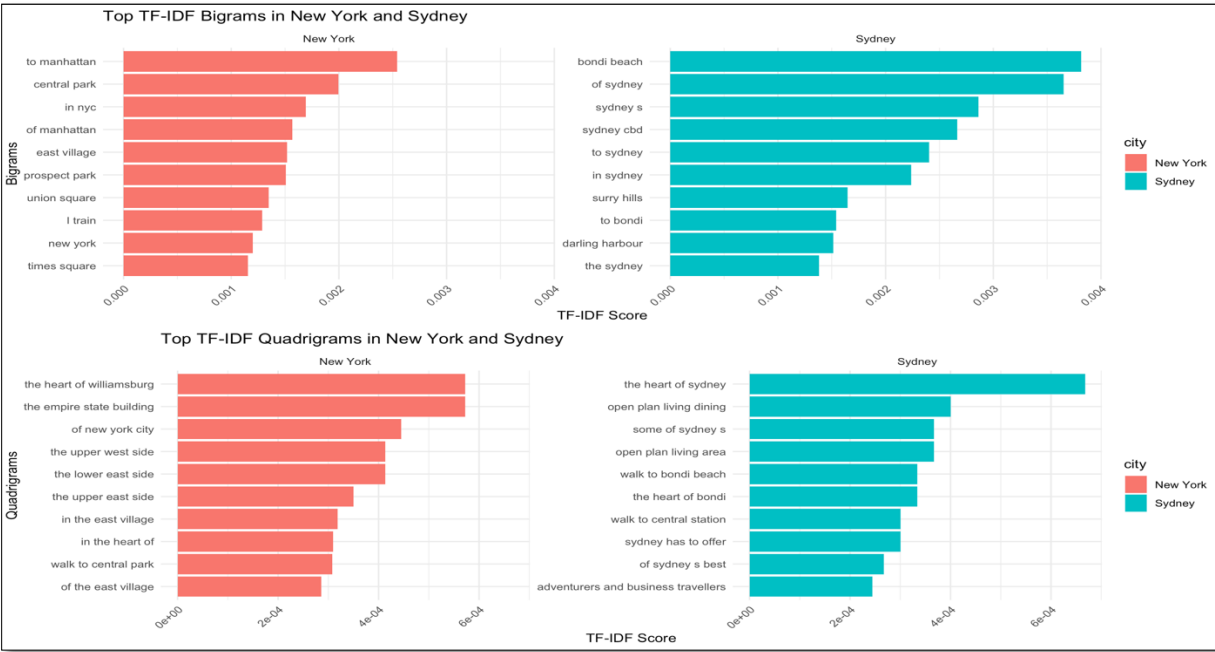


Figure 4

