

Combined assessment : Text Mining / NLP

Zarah Fwa

Business Analysis with Unstructured Data
MBAN 1

21st April 2024

Table of Content

Airbnb Posting Content Analysis Report	2
1. Executive Summary	2
2. Descriptions of visualizations created	2
3. Dashboard Description	3
4. Text Mining Framework	4
5. Recommendations	4
References	5
Appendix	6

Airbnb Posting Content Analysis Report

1. Executive Summary

As we have been tasked to know if there are any business insights that can be found in the posting's text and numerical content, Airbnb's listings provide crucial data about visitor preferences and allow Airbnb to strategically place its properties across the world. The textual content of postings, particularly the frequency of certain keywords and the prominence of facilities in relation to positive and negative emotion, is critical for selling Airbnb homes to potential guests or producing targeted advertising. The numerical data analysis reveals a difference in pricing methods across property types, allowing Airbnb to assist hosts in optimizing listings for increased income and guest happiness.

2. Descriptions of visualizations created

After visualization of the characteristics of the Airbnb listing, the four visualizations offer insights to Airbnb posting text and numerical content.

The first visualization focuses on the terms that appear most frequently in the description column. The bar chart shows that the most commonly used terms by Airbnb hosts to describe their accommodations are "apartment," "kitchen," and "bedroom." This suggests that these elements are important selling aspects in Airbnb property descriptions.

The second graphic shows the most frequently used facilities. The bar chart shows that the top ten facilities requested by guests for Airbnb rentals are "dryer" and "tv," as well as "friendly" and "free," which are most likely related with "pet-friendly" and "free-parking" or "free-wifi." These findings suggest that guests may prefer practical house convenience, which

is why hosts should emphasize such things in their descriptions or aim to have these amenities.

The third graphic compares the most often used terms in Airbnb listings from Australia, Brazil, and the United States. Australia's listings frequently include "beach" and "Sydney," showing a preference for beach vacations and urban experience. Amenities such as surfboard storage may attract guests. Brazil's words commonly mention "beach" and "Rio," indicating a desire for houses near tourist attractions such as Copacabana. In the United States, the terms "apartment" and "private" indicate a desire for secluded, self-contained quarters. While "beach" is mentioned, it is less prevalent than in Australia or Brazil, indicating a diverse market in which seclusion is an important selling advantage.

The property type distribution indicates a preference for "apartments" and houses, likely due to their familiar and spacious qualities, and often more favorable pricing compared to hotels. Serviced apartments, houses, and condominiums fetch higher prices, suggesting guests value space and luxury amenities. Hosts could enhance their offerings by upgrading properties in line with these market trends.

3. Dashboard Description

The dashboard has interactive filters of key metrics for property rentals. . This allows Airbnb to see real-time assessment of market dynamics and identification of trends. It enables stakeholders to isolate specific segments for deeper analysis. It breaks down the average pricing by property type, indicating a range in expected costs for guests. The property type distribution graph highlights the prevalence of various accommodations, with apartments being the most common. The amenities bar chart ranks features by frequency, showcasing the most valued aspects of rentals, such as Wi-Fi and essentials. This visualization aids hosts in identifying trends and optimizing their listings to align with market demands.

4. Text Mining Framework

Text analysis with three frameworks (tokenization, sentiment analysis, and n-grams) yielded various findings. Tokenization found that the most often used single terms in descriptions and amenities were "beach," "apartment," and "kitchen," implying that guests value location, lodging type, and facilities. Sentiment analysis using three lexicons (AFINN, Bing, and NRC) generated a sentiment score for tokens. Words with high sentiment ratings, such as "comfortable" and "private," might boost a posting's attractiveness, implying that guests will be drawn to content that conveys a sense of comfort and exclusivity. N-grams provide insight on common words such as "walking distance" and "equipped kitchen," implying that accessibility to attractions and a well-appointed kitchen are important selling qualities.

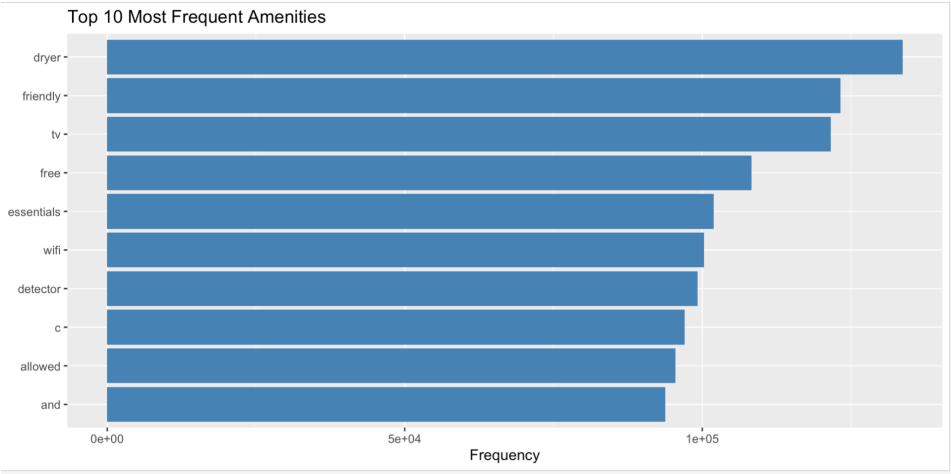
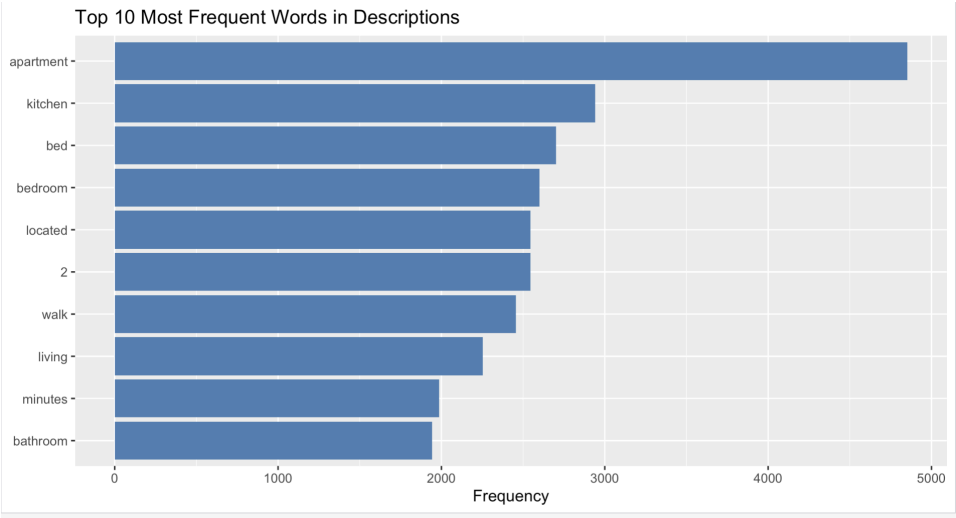
5. Recommendations

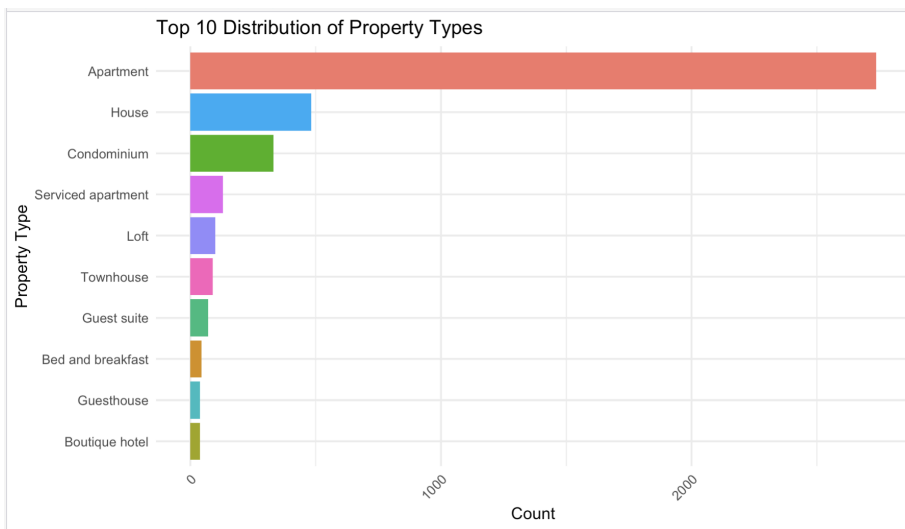
To maximize Airbnb listings, hosts should highlight crucial features such as "wifi" and "dryer," and use positive-connoting language in descriptions. Customize listings to showcase regional attractions, such as beaches in coastal locations, and change price to suit property type demand and amenities offered. Implement dynamic pricing throughout high seasons to ensure competitive prices when compared to similar local listings. Consider enhancements that appeal to tourists, such as contemporary kitchens, and clearly explain proximity to popular attractions or the availability of unique services, such as beach equipment for rent. These tactics can improve visitor experiences, reservations, and reviews.

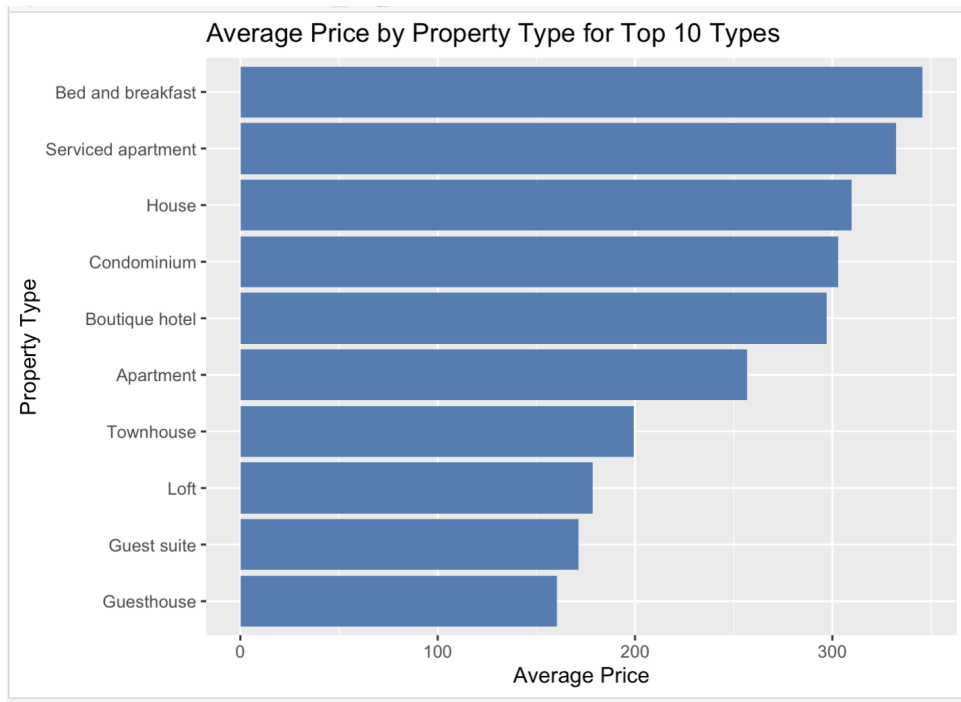
References

- MongoDB Atlas. (n.d.). *Cloud: MongoDB Cloud*. Account.mongodb.com. Retrieved April 22, 2024, from <https://cloud.mongodb.com/v2/6621d9b9880a46347d8f8d32#/clusters>
- Robinson, J. S. and D. (n.d.). 2 Sentiment analysis with tidy data | Text Mining with R. In *www.tidytextmining.com*. <https://www.tidytextmining.com/sentiment>

Appendix







My Dashboard

