

## ABSTRACT

This report unveils the insights gleaned from a comprehensive analysis of Airbnb listing descriptions, employing text mining and visualization techniques to decode customer preferences and sentiments. We spotlight pivotal terms and sentiment scores that influence guest experiences, offering Airbnb actionable strategies for enhancing listing appeal and market competitiveness.

Saichand Ghantasala

## EXECUTIVE SUMMARY

The data analytics initiative for Airbnb's marketplace has revealed pivotal insights into the customer language around property listings. Through text mining, sentiment analysis, and advanced visualizations, we've uncovered patterns and preferences that resonate with guests. These insights can guide strategic decisions for listing optimizations, targeted marketing, and improved customer experience.

## VISUALIZATION DESCRIPTIONS:

### **Frequency Analysis (Top 15 Words):**

A bar chart visualization identifies the most commonly used words in property descriptions, highlighting terms that resonate with guests. The frequency bar chart highlights the most commonly used words in Airbnb property descriptions. The word "city" tops the chart, indicating a significant emphasis on urban experiences. Other frequent terms such as "beach," "restaurants," and "bathroom" suggest guests are interested in properties that offer convenience, location-specific experiences, and well-furnished facilities. This data can guide hosts to emphasize such amenities and experiences that are likely to attract bookings.

### **Sentiment Analysis:**

Sentiment scores of words from property descriptions are depicted, offering a glimpse into the emotional undertone of the content Airbnb's hosts are providing. The sentiment analysis bar chart displays words that contribute positively to the listings' sentiment score. The presence of "available" and "free" at the top suggests that guests may have a strong preference for listings that offer availability and complimentary amenities. Words like "quiet," "enjoy," and "beautiful" also rank highly, reflecting the emotional desires guests may be seeking to fulfill with their stay. Hosts can use these insights to enhance their listing's emotional appeal by using language that has been shown to resonate positively with potential guests.

### **Bigram Analysis:**

A network graph that demonstrates the most frequent word pairings, illustrating how certain amenities or features are commonly described together. The network graph of bigrams demonstrates how certain words are often used together in property descriptions. The connection between "equipped" and "kitchen" indicates a preference for listings that offer ready-to-use cooking facilities. Similarly, "minutes walk" paired with "metro station" emphasizes proximity to transportation as a valuable feature. This visualization provides hosts with insight into crafting descriptions that align these amenities with guest preferences.

### **TF-IDF Analysis:**

This visualization ranks words by their importance across documents, helping to identify unique selling points of listings. The TF-IDF bar chart identifies unique terms that carry a significant weight in differentiating listings. The term "testtttt," which could be a placeholder or error in the data, appears at the top. Meanwhile, words like "noreen" and "bacana" (cool or nice in Portuguese) might be associated with specific, culturally unique listings or hosts. This indicates that incorporating local language or unique property features could be an effective strategy for standing out in a crowded market.

## **LDA Topic Modelling:**

The top terms for each derived topic are visualized, showing the thematic structure of listing descriptions and uncovering latent topics. The bar charts from the LDA model identify groups of words that form coherent topics across multiple Airbnb listings. For example, one topic cluster includes "pool," "ocean," and "beach," suggesting a theme around waterfront properties. Another cluster contains "city," "metro," and "minutes," which could relate to urban properties with convenient transportation options. These clusters assist in understanding the thematic composition of listings, which can inform targeted marketing campaigns.

## **DASHBOARD OVERVIEW:**

The interactive Airbnb Listings Dashboard provides a comprehensive exploration tool that allows users to fine-tune their search with specific filters such as city, number of bedrooms, and number of bathrooms. The dashboard integrates a detailed map view alongside data tables that display the top and bottom 10 rentals, based on review scores, providing a dual perspective that merges geographical insights with quality assessments. This juxtaposition enables a granular understanding of rental properties, facilitating informed decisions based on location, amenities, and guest satisfaction.

## **Key Findings and Strategic Business Insights from Text Mining Analysis:**

### **Word Frequency Analysis**

The visual representation of the 'Top 15 Words in Airbnb Description of houses' points to the salience of accessibility and amenities in the guest experience. The prevalence of terms such as "city," "beach," and "restaurants" underscores a significant trend where guests prefer listings that offer both natural and urban attractions. For Airbnb hosts, this emphasizes the importance of highlighting locational advantages in their listings. Strategic marketing that foregrounds these elements can elevate the property's appeal, potentially increasing occupancy rates.

### **Bigram Analysis**

The network diagram of 'Top 15 Most Frequent Bigrams' presents insightful pairings like "minutes walk" and "equipped kitchen," signaling that guests value convenience and functionality. This analysis suggests hosts should not only focus on the aesthetic appeal but also the practicality of their accommodations. Emphasizing walkability to key destinations and the availability of fully equipped kitchens can serve as pivotal decision factors for potential guests.

### **TF-IDF Scores**

The 'Top TF-IDF Scores in Airbnb Descriptions' graph sheds light on unique terms that set listings apart, such as "ses" and "marroquina." These may point to properties with distinctive cultural or design features, tapping into niche markets. Hosts can leverage these findings by crafting unique storytelling around these elements, thereby attracting guests interested in curated experiences or specific cultural aspects.

## **LDA Topic Modeling**

The 'Top Terms in Each Topic from LDA Model' graph exposes underlying themes within the listings, emphasizing elements like transit proximity and exclusive amenities such as pools. The insight gained here allows for segmentation strategies, where hosts can target specific audiences—like business travelers who need transit access or luxury seekers interested in premium features.

## **Sentiment Analysis**

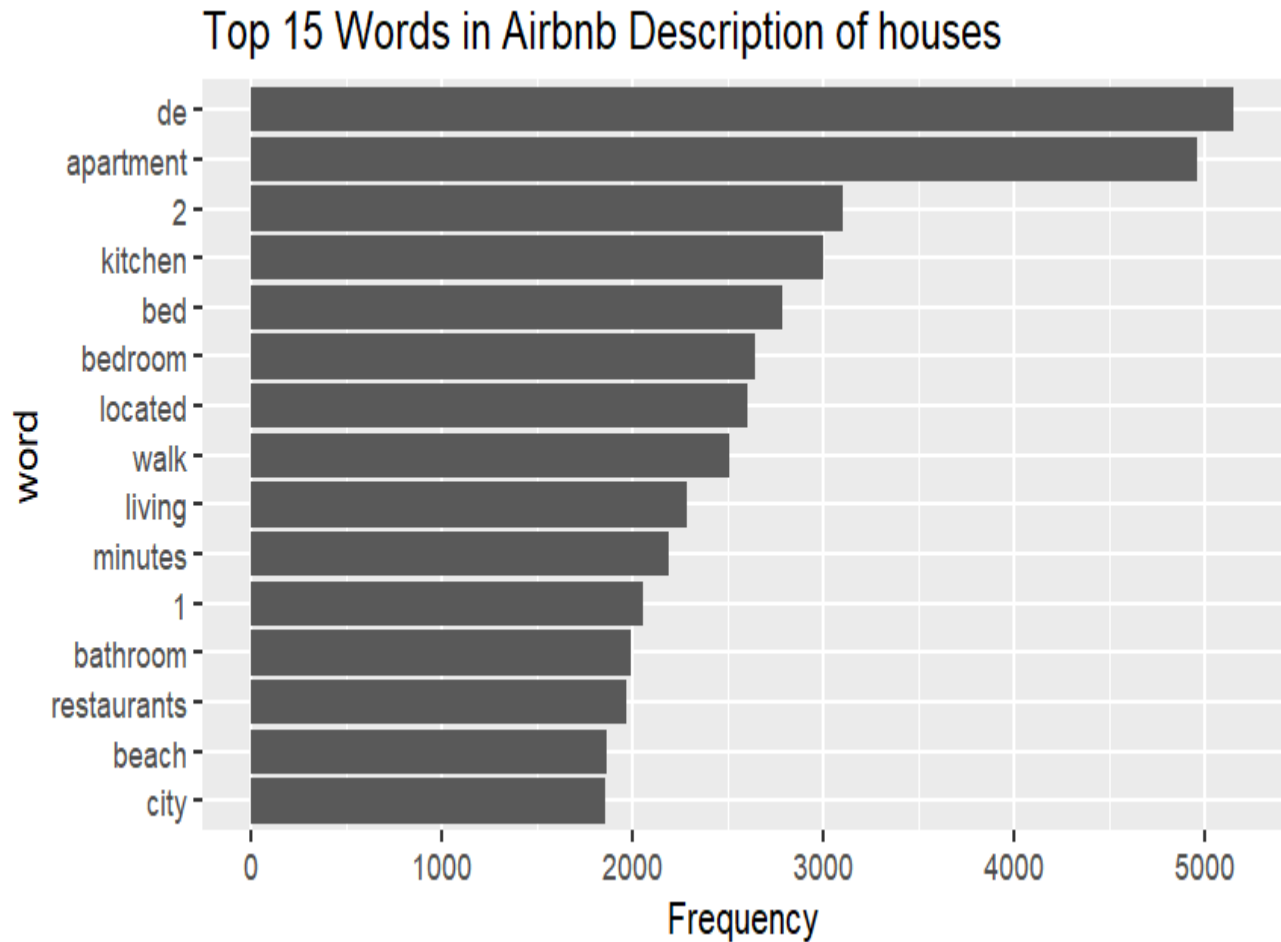
The sentiment analysis visual, which showcases overwhelmingly positive descriptors such as "comfortable," "spacious," and "modern," illustrates the importance of positive language in listings. For hosts and property managers, this signifies a direct correlation between the sentiment conveyed in a property description and the likelihood of attracting guests. There is a clear opportunity to audit and enhance property listings to ensure language use is effectively capturing and conveying the positive attributes of a listing.

Incorporating these data-driven insights can inform business strategies that not only align with current market demands but also anticipate emerging trends. Optimizing property listings based on these analyses can help hosts differentiate their offerings, ultimately leading to increased visibility, guest satisfaction, and market competitiveness. As the Airbnb market flourishes, the ability to harness and apply these insights will become a critical factor in maintaining a strategic edge.

## APPENDIX:

### Visualizations from Text Mining:

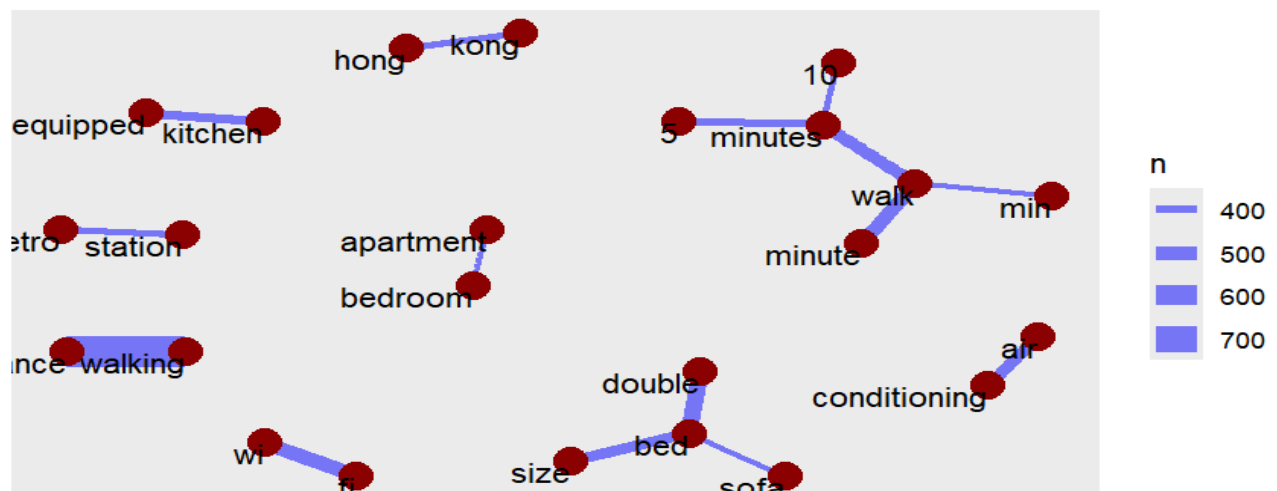
#### Frequency Analysis:



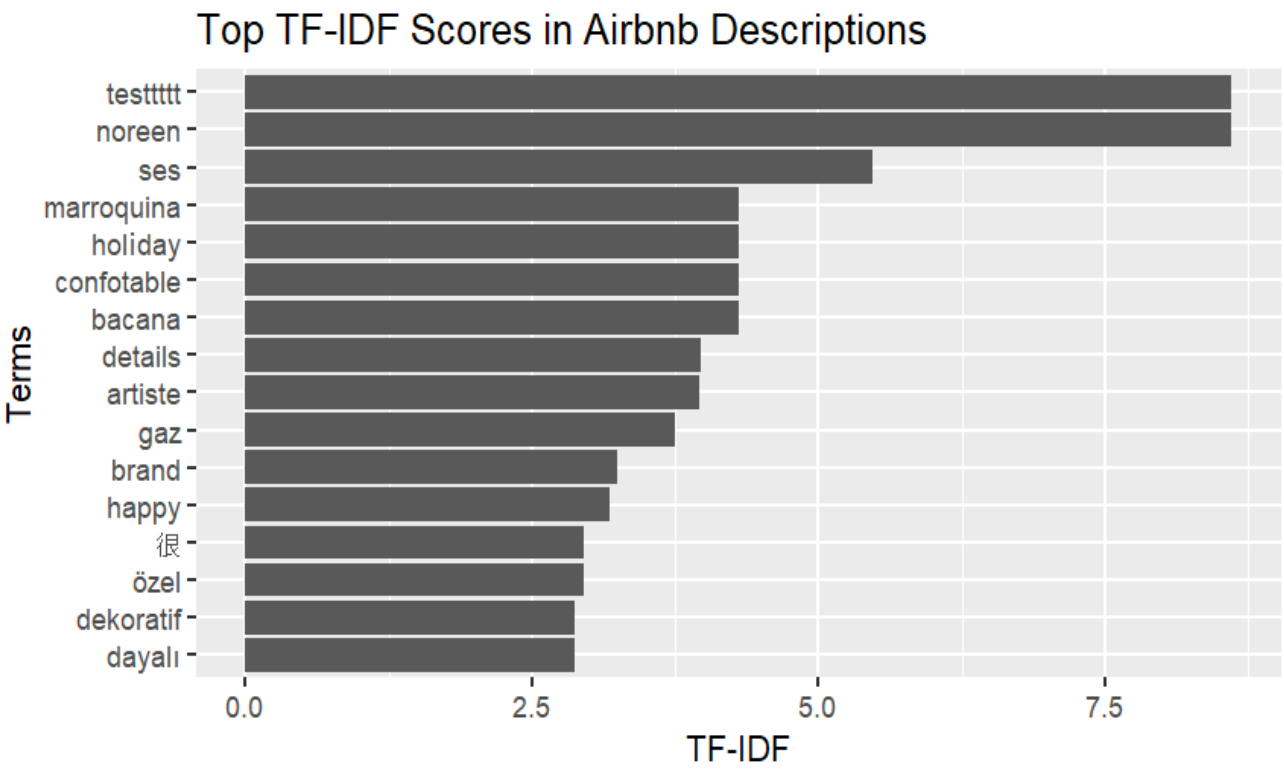
#### Bigram Analysis:

##### Top 15 Most Frequent Bigrams

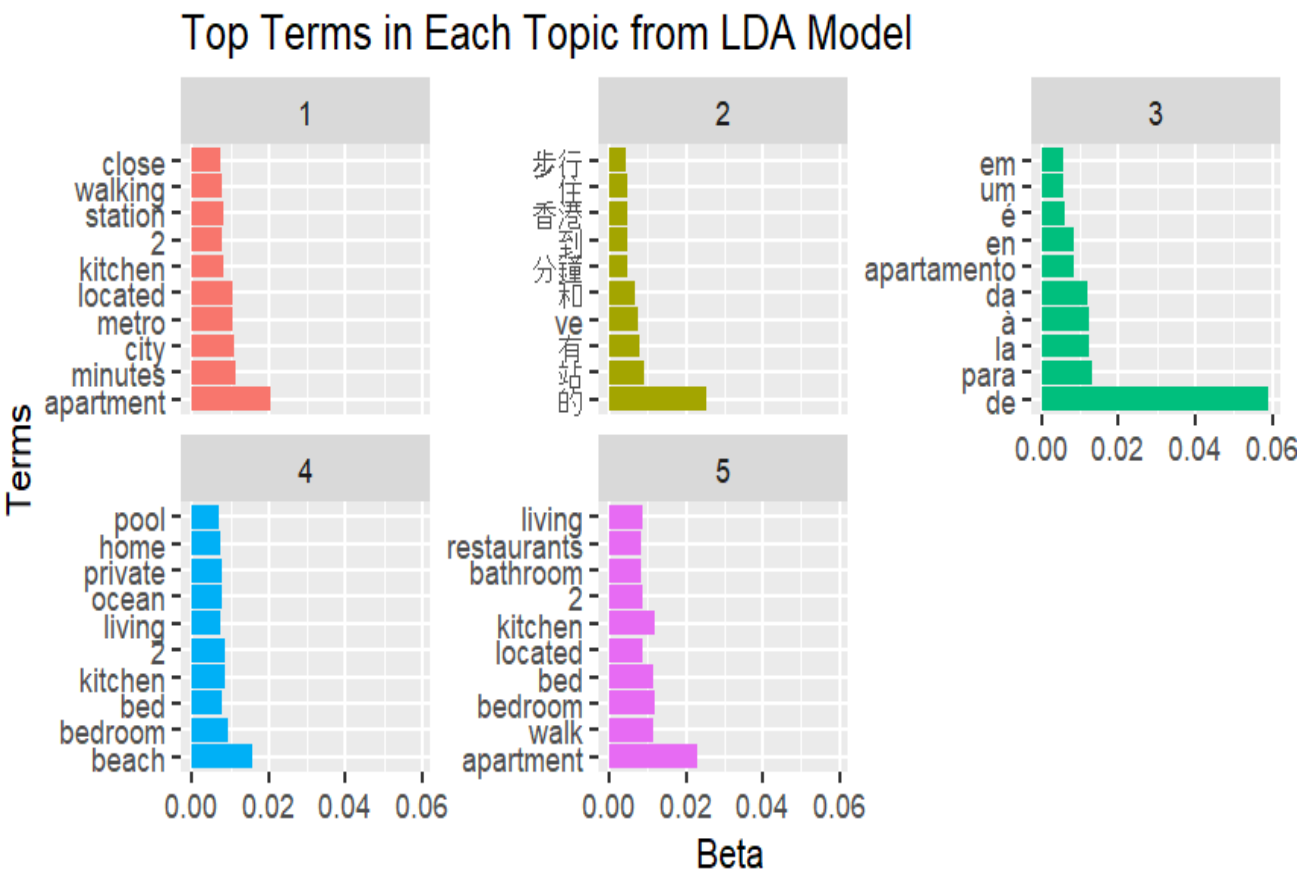
Filtered to exclude common stopwords



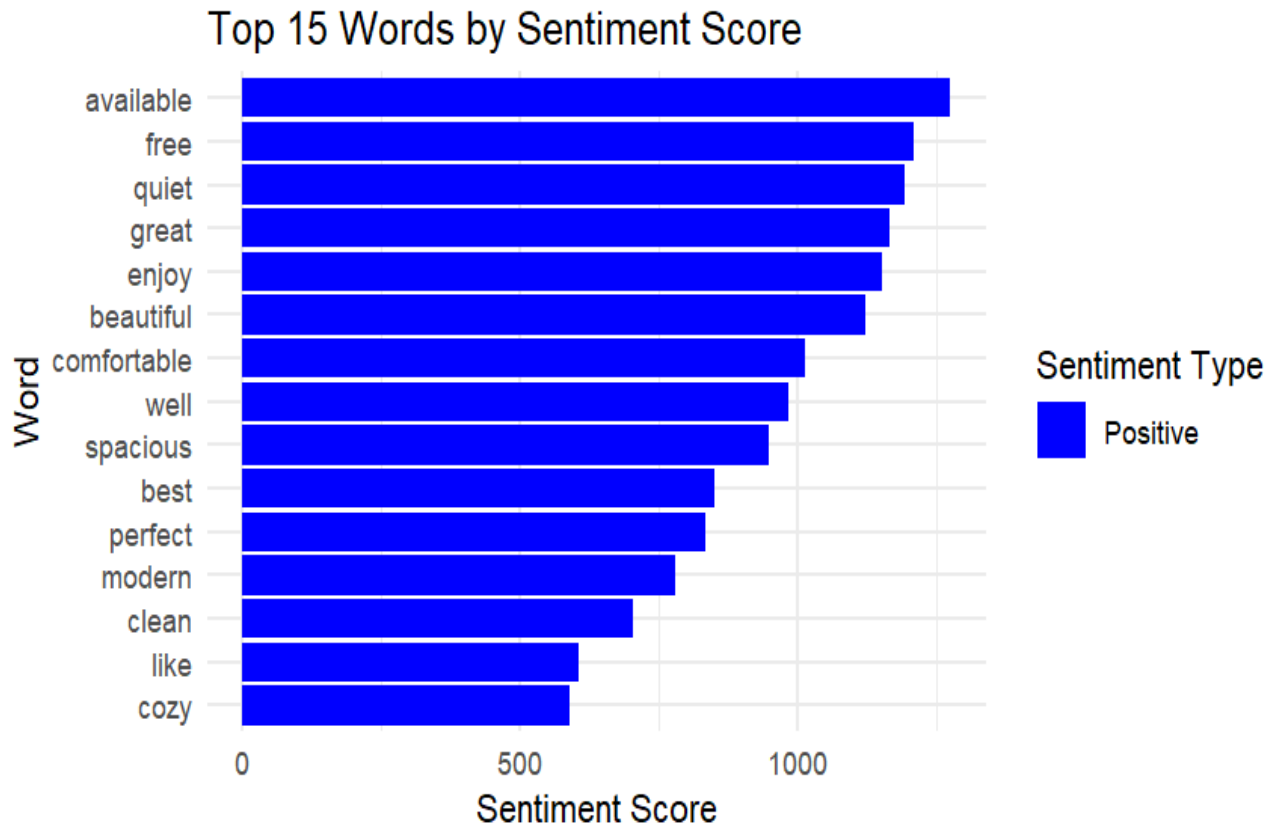
TF-IDF:



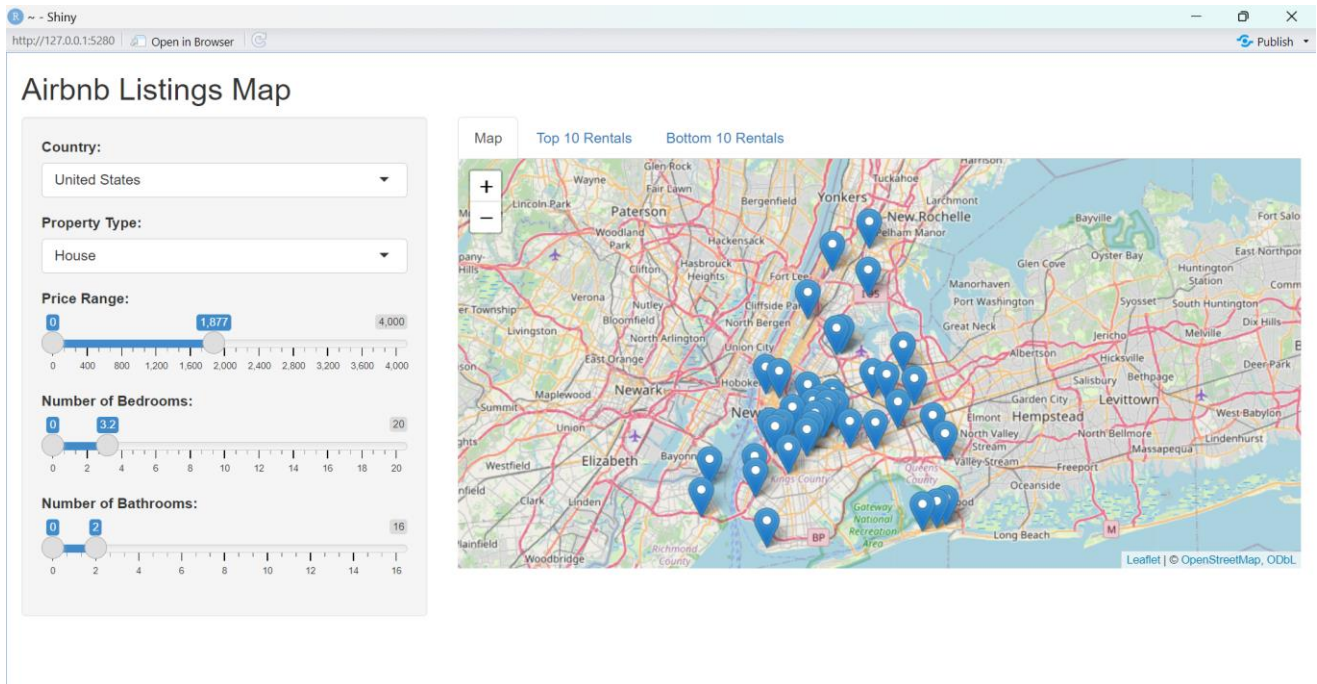
LDA Model Analysis:



## Sentiment Analysis:



## DASHBOARD SCREENSHOTS:



Shiny

http://127.0.0.1:5280 Open in Browser Publish

## Airbnb Listings Map

Country: United States

Property Type: House

Price Range: 0 1,877 4,000

Number of Bedrooms: 0 3.2 20

Number of Bathrooms: 0 2 16

Map Top 10 Rentals Bottom 10 Rentals

Show 10 entries

Search:

	url	name	summary	space	description	neight
1	<a href="https://www.airbnb.com/rooms/11935301">https://www.airbnb.com/rooms/11935301</a>	Spacious, Sunny Room in Park Slope	Spacious, sunny room in owner occupied house in the heart of Park Slope. Shared bathroom with shower/bath tub. Steps away from restaurants/bars and shopping. Within walking distance of major subway lines. Prospect	The bedroom is quite large. It is very sunny. There is a table, a small refrigerator and a coffee maker in the bedroom.	Spacious, sunny room in owner occupied house in the heart of Park Slope. Shared bathroom with shower/bath tub. Steps away from restaurants/bars and shopping. Within walking distance of major subway lines, Prospect Park, BAM and Barclay Center. The bedroom is quite large. It is very sunny. There is a table	

Shiny

http://127.0.0.1:5280 Open in Browser Publish

## Airbnb Listings Map

Country: United States

Property Type: House

Price Range: 0 1,877 4,000

Number of Bedrooms: 0 3.2 20

Number of Bathrooms: 0 2 16

Map Top 10 Rentals Bottom 10 Rentals

Show 10 entries

Search:

	url	name	summary	space	description	neight
2	<a href="https://www.airbnb.com/rooms/23839051">https://www.airbnb.com/rooms/23839051</a>	Kama'iki	Welcome to Kama'iki. We would love to share our little home near the beach with you. Kama'iki means small child, and Kama'iki is our small child. He sits on the property with two other houses. Kama'iki is the middle house, he is between a house on the road and a house on the beach. He has a gazebo on his makai side (Oceanside), where you can feel and hear the tropical breezes, smell the ocean mists, and hear the crashing waves. Perfect for just a peaceful, relaxing, quiet vacation.			

## Airbnb Text Analysis Dashboard

Number of Words:

1 10 50

1 6 11 16 21 26 31 36 41 46 50

Frequency

Sentiment

Bigrams

TF-IDF

LDA

