**Executive Summary**

With the use of the Airbnb dataset's analysis of posting descriptions, this report—which was ordered by Airbnb—aims to provide business insights. This study provides a thorough overview of trends, host behaviors, and market dynamics in the Airbnb ecosystem through a thorough analysis of text and numerical material and rich data visualization. Significant new information is provided by the results about the density of Airbnb listings by nation, pricing distribution for various locations and room kinds, host quality indicators, response rates, and customer preferences for certain room and bed types, cancellation policies, and listing sentiment. In this investigation, sentiment analysis, bigram frequency analysis, and term frequency-inverse document frequency (TF-IDF) were the main text mining frameworks used. In order to improve visitor pleasure and maximize listing performance, Airbnb and its hosts may be able to make strategic decisions based on trends and preferences found in listing descriptions.

**Visualizations Description**

Airbnb Density by nation (Appendix 1): This section shows the concentration of listings globally using a nation heat map, illustrating the differences in the regulatory environments and popularity of Airbnb between different countries. Appendix 2 displays the median price distribution by room type and location. created a stacked bar chart to show how location and room type affect pricing strategies by comparing the median costs for various room types across many nations. Highest Caliber Hosts (Appendix 3) made a tabular representation to show the relationship between host performance and guest satisfaction by ranking hosts according to the quantity of reviews and average review scores. Host Response Rate Analysis (Appendix 4): To illustrate the significance of responsiveness in host ratings, a heat bar chart was utilized to link host response times with average review scores and the amount of reviews. Appendix 5 & 5: Frequency Distribution of Room and Bed Types used eye-catching bar graphs to illustrate market trends and guest preferences by showing the frequency of different room and bed configurations in listings. Cancellation Policy Frequency Distribution (Appendix 6): A bar graph was used to examine the popularity of various cancellation policies among hosts, providing insight into how booking conditions strike a compromise between certainty and flexibility. Leading Terms and Bigrams in Listings on Airbnb (Appendices 7 & 8): In order to provide insights into the main qualities and facilities that hosts emphasize, bar graphs were made to show which words and two-word phrases appear most frequently in listing descriptions. Emotional Distribution in Listings (Appendix 9): A bar graph that shows the favorable view of Airbnb lodgings by measuring the sentiments found in listing descriptions. Word Length against Score for TF-IDF (Appendix 10): Displayed as a scatter plot, this study looks at the relevance of word length in relation to listing descriptions.

**A thorough study of Airbnb data is offered by the Tableau dashboard through a variety of visualizations (Appendix 11):**

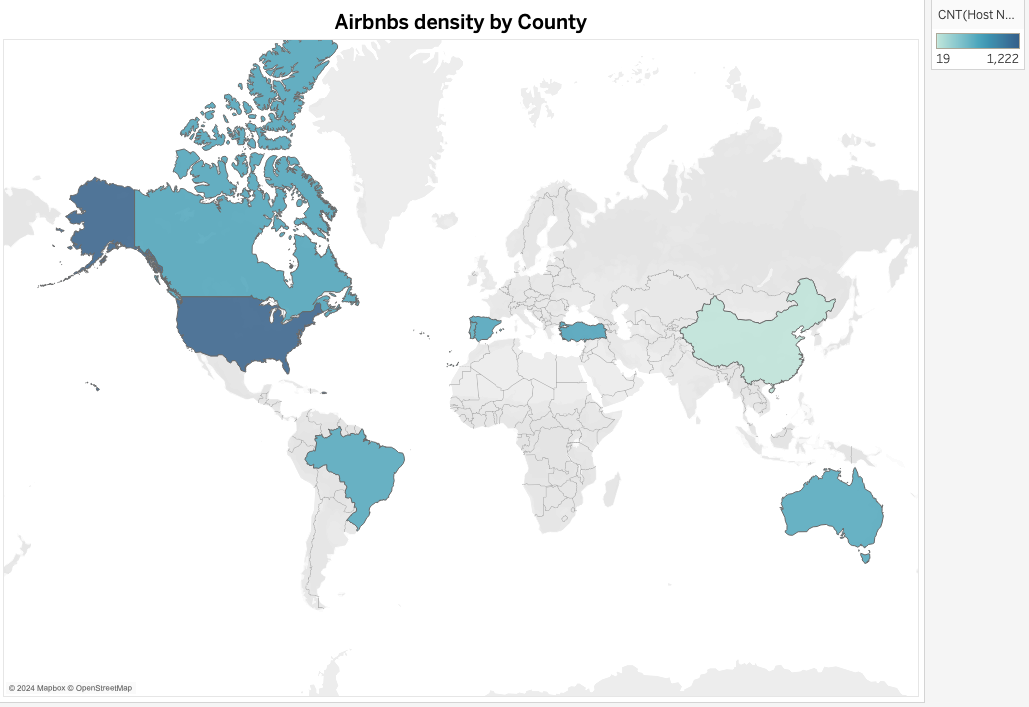
1. The global map with the nations highlighted in various hues of blue illustrates Airbnb's Density by County. The density of Airbnbs is greater in darker shades.
2. Median Price Distribution by Room Type and Location: This is shown as a vertical bar chart that shows the median costs for various room kinds across a number of nations. Separate rooms are colored blue, shared rooms are colored grey, and the entire house or apartment is orange. There are dropdown menus for Country and Room Type filters available.
3. Top-Rated Hosts based on Happiness of Guests: This table presents hosts according to how satisfied guests are. The number of reviews received, the average review score out of 1,000, and the host names (first name only) are shown in columns.
4. Analysis of the Host Response Rate: This is a horizontal bar graph that compares the quantity of reviews received to the host's response times. The bars show the various response time frames: within one hour, within a few hours, within a day, Null, and more. There is an Action filter above the graph that is set to "(Host Response Time)" and has options for Number of Reviews, which ranges from 704 to 117.329.

**Key Findings and Business Insights from Text Data**

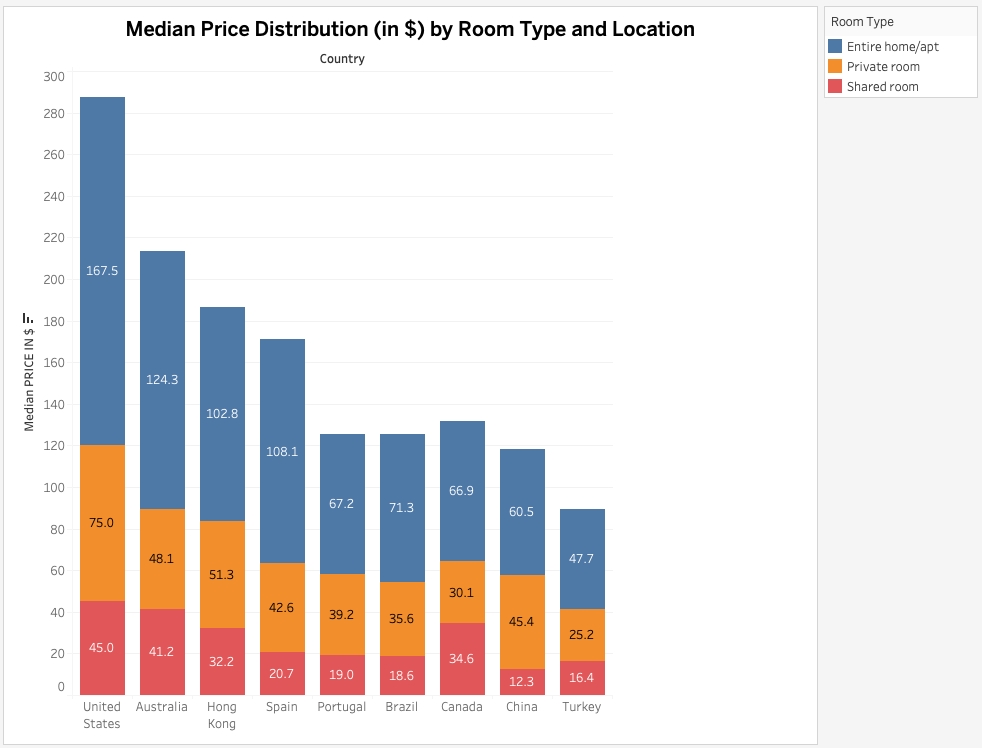
* Key Descriptor Prevalence in Listings (Appendices 7 & 8): Airbnb listings strategically emphasize terms like "stay," "apartment," and "location," along with phrases like "in the," "to the," and "the apartment," according to an analysis of the top words and bigrams in the listings. This suggests that hosts are acutely aware of the features that visitors value most, in this case the type of lodgings and their advantageous locations. Hosts may increase the appeal of their listings to prospective visitors by highlighting these features. This information implies that in order to increase its market appeal, Airbnb should concentrate on marketing listings that emphasize these important characteristics as they may increase visitor interest and reservations.
* Positive Emotions Are Highlighted by Sentiment Analysis (Appendix 9): There is a noticeable inclination towards positive emotions, especially happiness and trust, according to the Distribution of Emotions in Airbnb Listings. This suggests that customers have a generally positive opinion of Airbnb listings, which is a major selling factor for the website. This insight emphasizes the significance of preserving and improving the elements of the service that lead to these favorable feelings for company strategy. These elements include consistent customer service, truthful and transparent listing descriptions, and visitor experiences that either meet or surpass expectations.
* Preferences for Room and Bed Type (Appendices 5.1 & 5.2): The room and bed types frequency distribution in listings shows a distinct preference for 'Entire home/apt' and 'Real Bed' alternatives when it comes to Airbnb offers. This inclination suggests that Airbnb guests have a market need for comfort and privacy. This presents an opportunity for Airbnb and its hosts to modify inventory to accommodate these preferences; for example, they may give priority to whole houses or flats that have actual beds in their listings and search results to match visitor expectations.
* Trends in Cancellation Policies (Appendix 6): The examination of cancelation rules reveals that hosts often favor "Strict" regulations. This may indicate a desire for booking security, but it could also turn off travelers who are looking for flexibility. Through encouraging a combination of cancellation rules that address both host security and guest flexibility—possibly through rewards for hosts who provide more flexible restrictions—Airbnb has the chance to strike a balance between both demands.

**APPENDIX**

APPENDIX 1 : Airbnbs density by County



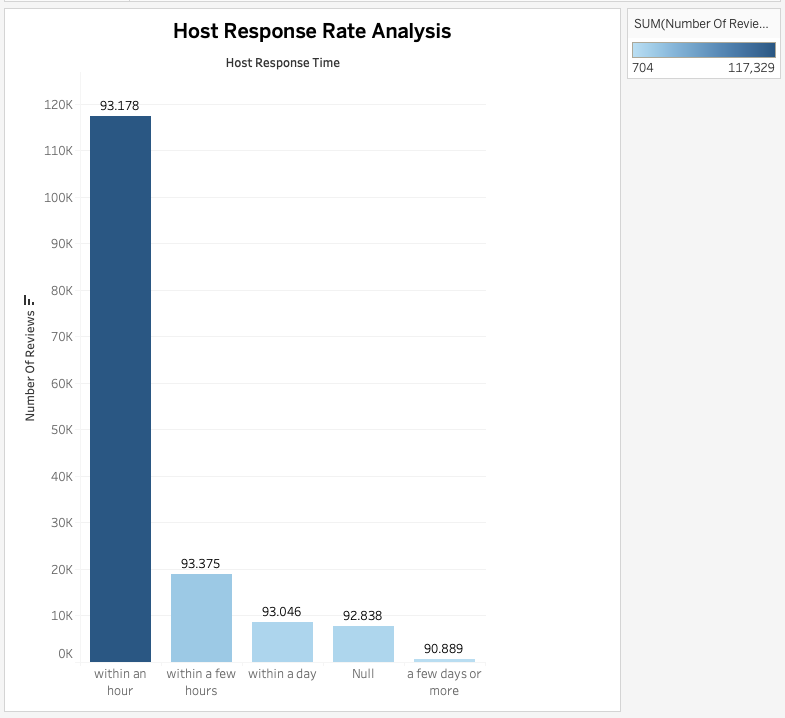
APPENDIX 2: Median Price Distribution by Room Type and Location



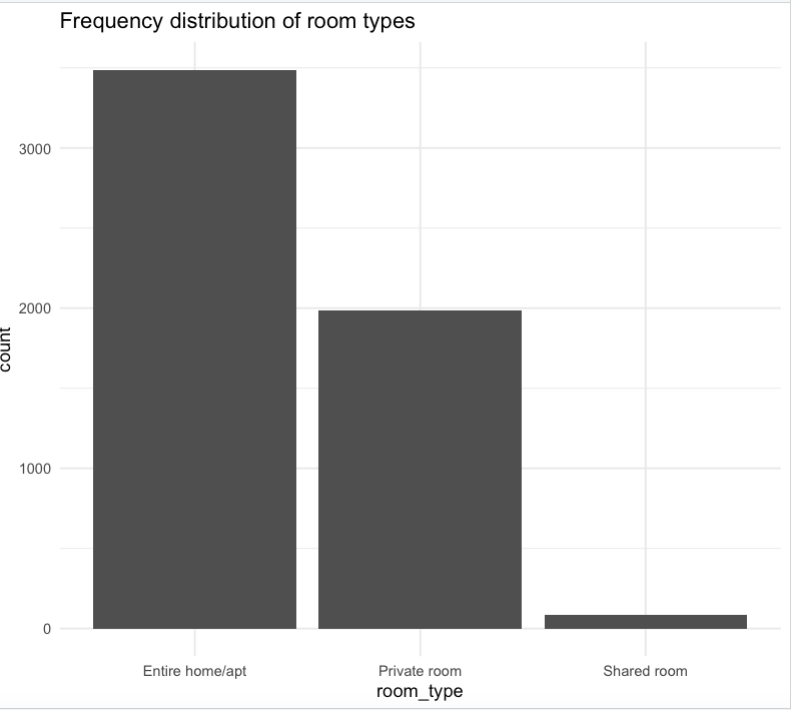
APPENDIX : 3 TOP QUALITY HOSTS

| **TOP QUALITY HOSTS** | | |
| --- | --- | --- |
| **Host Name** | **Number Of Reviews** | **Avg. Review Scores Rating** |
| David | 1,244 | 93 |
| Sarah | 905 | 91 |
| Maria | 900 | 95 |
| Paul | 786 | 96 |
| Ana | 758 | 93 |
| Sam | 735 | 91 |
| Luis | 728 | 91 |
| Cristina | 709 | 93 |
| Koni | 667 | 93 |
| João | 662 | 95 |
| Jose | 645 | 91 |
| Dana | 644 | 97 |
| Michael | 616 | 94 |
| Jason | 612 | 95 |
| Susan | 604 | 91 |
| Carlos | 578 | 91 |
| Miguel | 563 | 91 |
| Paula | 557 | 98 |
| Francisco | 555 | 87 |

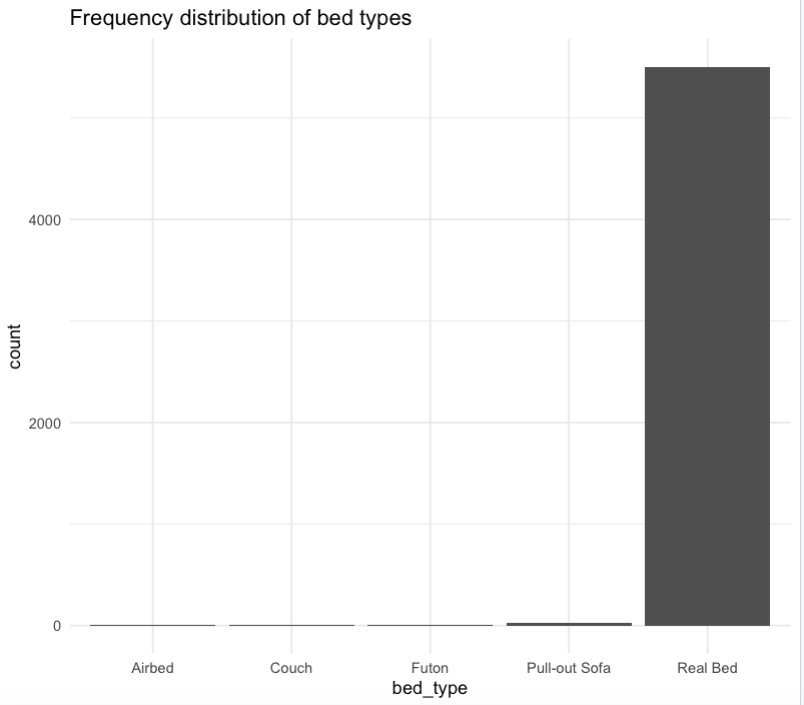
APPENDIX 4 : Host Response Rate Analysis



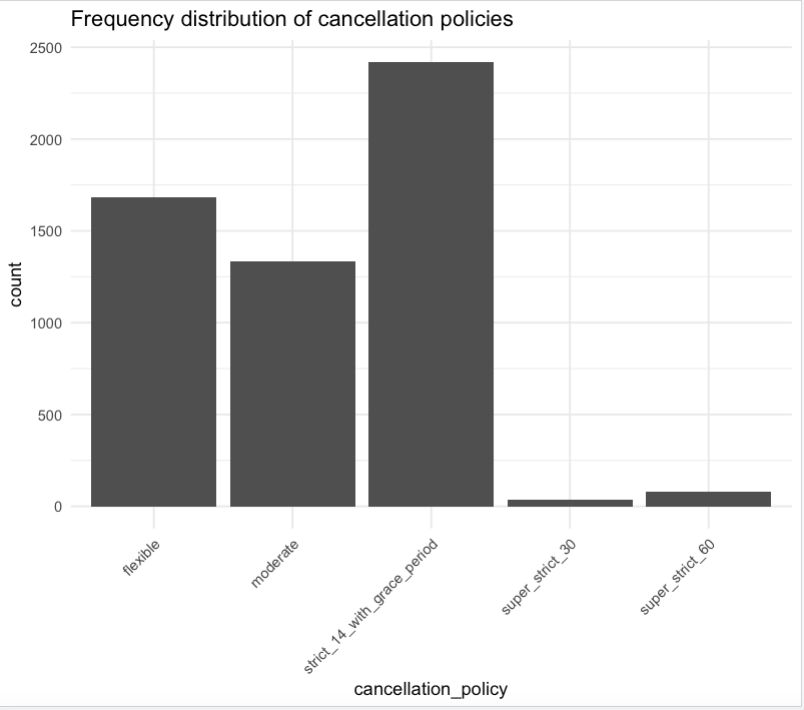
APPENDIX 5.1 : FREQUENCY DISTRIBUTION OF ROOM TYPES



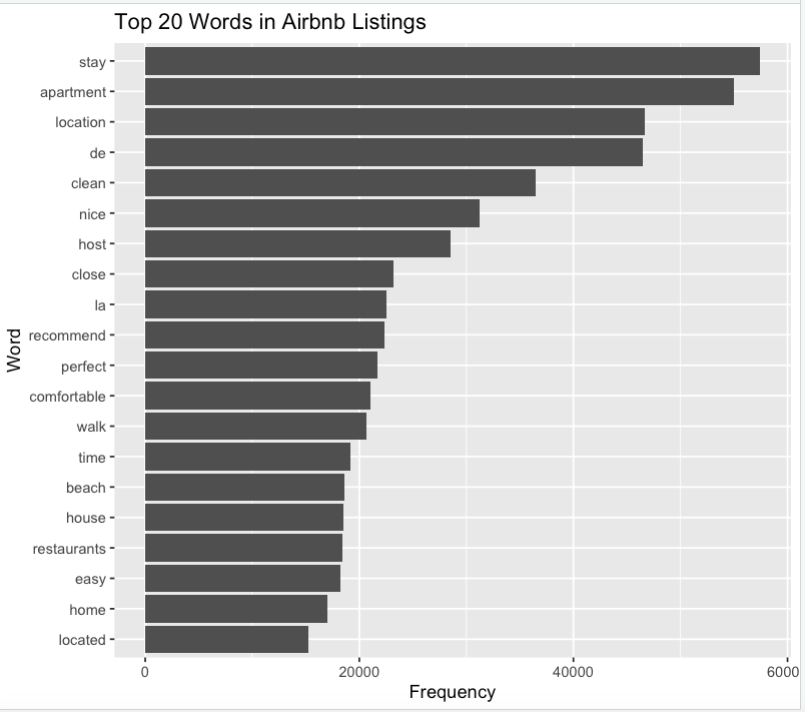
APPENDIX 5.2 : FREQUENCY DISTRIBUTION OF BED TYPES



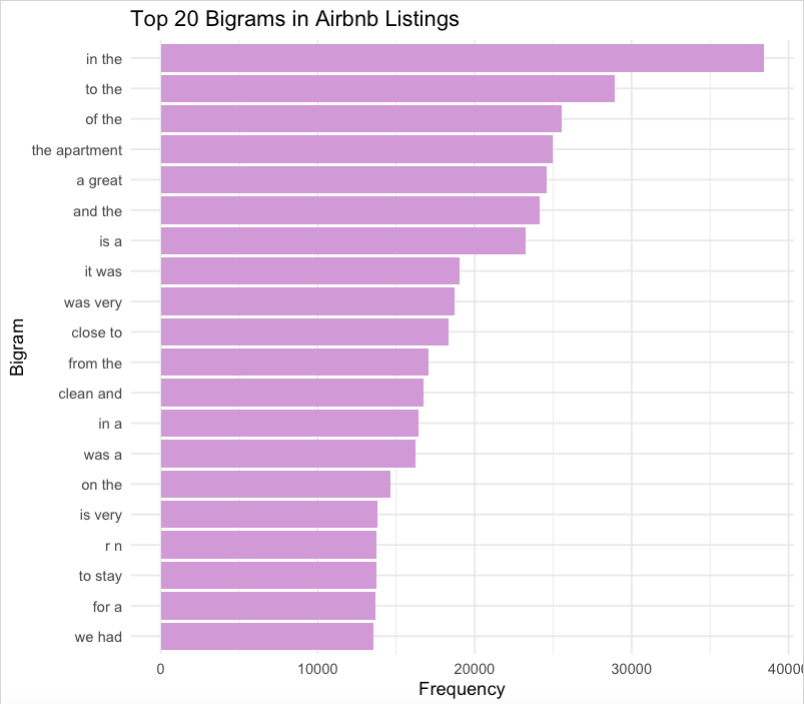
APPENDIX 6 : FREQUENCY DISTRIBUTION OF CANCELLATION POLICIES



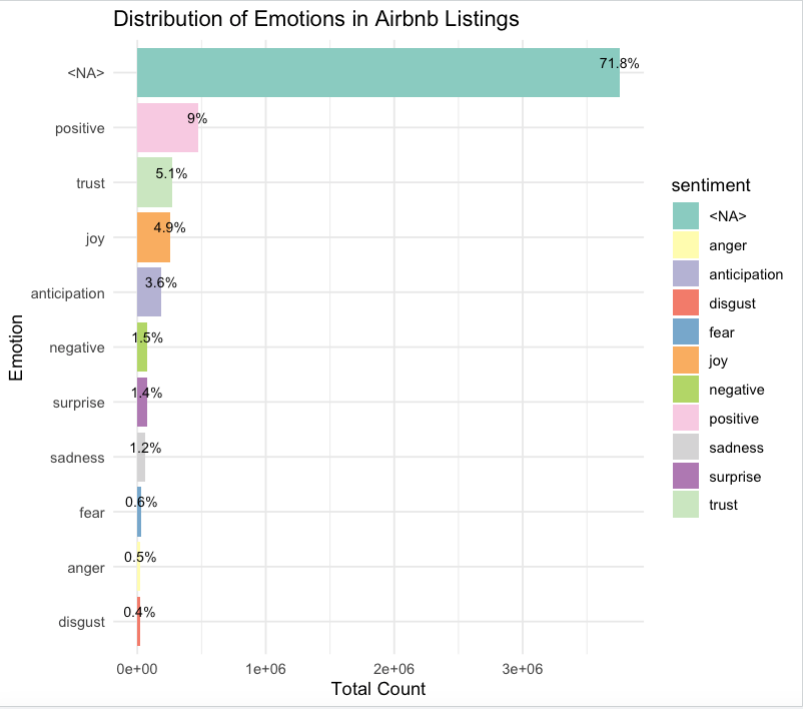
APPENDIX 7 : TOP 20 WORDS IN AIRBNB LISTINGS



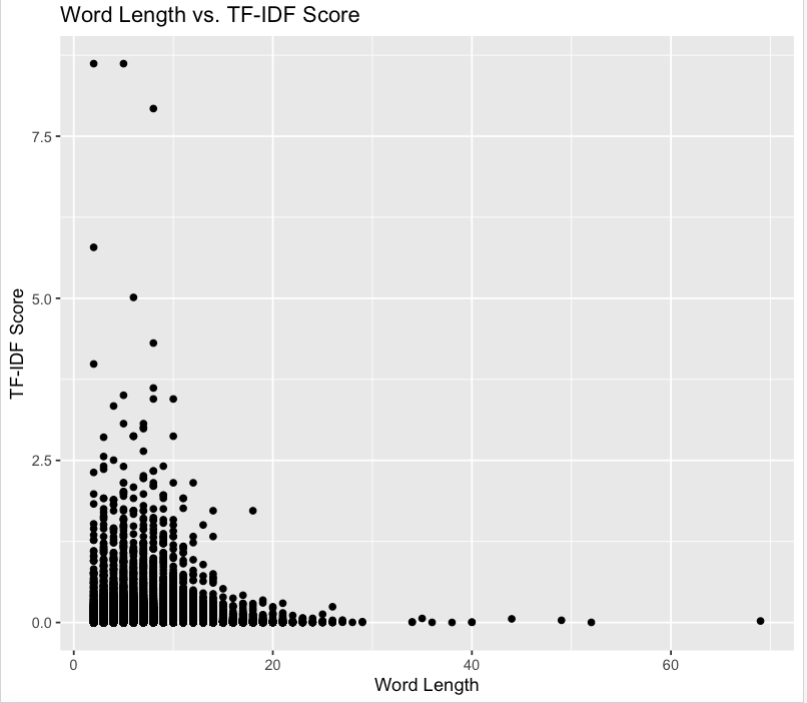
APPENDIX 8 : TOP 20 BIGRAMS IN AIRBNB LISTINGS



APPENDIX 9 : DISTRIBUTION OF EMOTIONS IN AIRBNB LISTINGS



APPENDIX 10 : WORD LENGTH VS TF-IDF SCORE



APPENDIX 11 : DASHBOARD