Twitter and Airbnb: Sentiment and Perceptions

@justgrossi
https://github.com/justgrossi/Portfolio.git

***Abstract*—A sentiment analysis was performed to assess Twitter users' attitude and perceptions towards Airbnb. Tweets retrieved via API were tokenized and through natural language processing techniques positive, negative, and neutral feelings were categorized. Findings revealed the '*Airbnb*' brand inspired a diverse range of sentiments offering insights into customers' perceptions. Such analyses can benefit a multitude of companies in improving both their brand image and customers' experience. This review also underscores the significance of social media sentiment analysis in deciphering public perceptions for strategic decision-making.**

***Index Terms*—Sentiment Analysis, Social Media, Natural Language Processing.**

I. Introduction

Social media platforms have become valuable sources for understanding public sentiment and perceptions towards various products, services, and brands. Among these platforms, Twitter stands as a significant forum where users express opinions, share experiences, and engage in discussions encompassing diverse topics. The hospitality industry, characterized by its dynamic nature and reliance on customer satisfaction, is no exception to the influence of social media. This paper delves into the analysis of Twitter users' discussions pertaining to Airbnb, a prominent player in the sharing economy and accommodation rental space. Since its inception Airbnb has revolutionized the way people find lodging by providing a platform for individuals to rent out their properties or book unique accommodations worldwide. The primary objective of this analysis was to scrutinize and comprehend sentiments, experiences, and perceptions expressed by Twitter users. Leveraging natural language processing techniques on a data set collected via API, the aim was to uncover prevalent sentiments—be they positive, negative, or neutral—associated with the short-term rental platform. Understanding user sentiments on platforms like Twitter is crucial for companies to gauge customer satisfaction, identify pain points, and adapt their services to meet evolving consumer preferences. By analyzing Twitter users' sentiments towards Airbnb the analysis endeavors to contribute valuable insights into the public perception landscape of the hospitality industry and aid in strategic decision-making for service enhancement and customer satisfaction improvement. The subsequent sections will outline the methodology employed for data collection, sentiment analysis techniques utilized, presentation of findings, and a comprehensive discussion of the implications derived from the analysis.

II. Related Work

Understanding sentiments expressed on social media platforms has garnered considerable attention, as it offers invaluable insights into public perceptions across diverse domains. Sentiment analysis studies focused on Twitter data have been pivotal in decoding user sentiments and opinions. Hu et al. [1] demonstrated the efficacy of such methodology in extracting opinions and sentiments, highlighting its relevance in deciphering public attitudes towards various subjects and laying the groundwork for opinion extraction from customer feedback. Pak and Paroubek [2] examined sentiment classification techniques on Twitter, revealing the challenges posed by the platform's brevity and informal language shaping subsequent strategies for mining sentiments from social media's dynamic and concise content. Jansen et al. [3] highlighted the role of tweets as electronic word of mouth establishing Twitter as a powerful medium for opinion dissemination and influence. Authors also emphasized the rapid evolution of sentiment analysis techniques and the necessity for context-aware methodologies.

Additionally, Wang et al. [4] conducted a sentiment analysis study on consumer perceptions of sharing economy platforms, shedding light on user sentiments towards platforms like Airbnb and emphasizing the need for such analyses in understanding evolving consumer behaviors. Their work explored risk perception in the sharing economy unveiling nuanced insights within this evolving economy landscape. Furthermore, Chen and Lin [5] did the same in the context of online reviews, showcasing the application of cutting-edge deep learning techniques highlighting the importance of such analyses in uncovering valuable insights from user-generated content. Within the framework of this study, which focuses on sentiment analysis of Twitter discussions related to Airbnb, the literature underlines the significance of understanding public sentiments towards such platforms. This comprehension plays a vital role in informing strategic decisions and fostering positive user experiences in the continually evolving hospitality landscape. This array of literature exemplifies multifaceted methodologies in sentiment analysis across customer reviews, social media, and the sharing economy. These studies collectively deepen our understanding of sentiment analysis techniques and their applications across varied domains, paving the way for enhanced opinion mining and sentiment interpretation.

III. Methodology

A. Data Collection

The data collection process began with the creation of a Twitter developer type of account [6]. Once the access was granted, the next step was to create an environment and a dedicated application within it. All the required consumer keys and tokens necessary to set up the connection to the

API were generated, as to ensure the authentication would successfully go through once the script began pulling tweets. The extraction process was limited to tweets including these keywords, also encompassing what appeared to be some of the most common misspellings: *'airbnb', '@airbnb', '#airbnb', 'Airbnb', '@Airbnb', '#Airbnb', 'Air BnB', 'Air B N B', 'AirB&B', '#AirBnb', 'Air−bnb', 'Air_bnb'*.

The API documentation suggested a ten-word limit per request, as too complex queries could potentially result in errors, hindering the retrieval of information. The keywords, therefore, were looped through to get the best outcomes possible while respecting the guardrails applied to developer accounts in terms, for instance, of number of overall requests in a given time frame. In addition, since the API limited the access to tweets generated during the previous seven days, the data collection phase lasted six weeks to capture a comprehensive snapshot of discussions for a total of 4,412 tweets. That was necessary as the quantity of tweets obtained would frequently fall short of the specified amount in the search query. This discrepancy can occur due to several reasons. For instance, insufficient tweets containing the specified keywords are posted within the designated time frame; the allotted rate limits associated with a specific token are reached; recent actions by Twitter users such as tweet deletions can affect the available data.

Only English tweets tweeted from the New York area in the U.S.A. were retrieved as the city is consistently listed among company's key markets [7]. Retweets were excluded.

B. *Preprocessing*

Prior to analysis a comprehensive pre-processing phase was undertaken. The first step was to consolidate the data in a unique database. This also involved data cleaning procedures such as eliminating duplicates, handling missing or irrelevant content, filtering out noise, spam, and promotional tweets. To standardize the corpus text normalization techniques including lowercasing, eliminating leading and trailing spaces, removing special characters such as emoji and emoticons were applied as well. Lastly, through *tokenization*, each tweet was broken down into its fundamental components or words. During the *tokenization* process an input, commonly a string, and a token type, a meaningful textual unit like a word, are employed to partition the input into segments, known as tokens, aligning with the specified type [12]. The next step was to remove *'stop words'*, common words that carry little or no meaningful information when performing a sentiment analysis. It is common advice and practice to remove stop words for various natural language processing tasks. One of the main reasons for removing them is to decrease the computational time for text mining, hence, it can be regarded as a dimensionality reduction technique. Such principle was commonly applied especially to search engines so as to obtain better results [8] [9]. The task was carried out leveraging the R *tidytext* package which includes a list of 1,149 stop words divided upon three different lexicons *"SMART"*, *"snowball"* and *"onix"*.

C. *Sentiment Analysis Techniques*

The approach to the analysis was lexicon based, hence, it involved the use of sentiment dictionaries to assign scores to the individual tokens (i.e. words) composing the tweets. The *Bing Liu* [1], *Loughran* [10] and *AFINN* [11] lexicons were used:

- ***Bing Liu***:

This lexicon assigns polarity scores based on words' emotional connotations. It contains a list of words and their corresponding sentiment scores, typically ranging from strongly negative to strongly positive. For instance, words like "happy," "love," or "excellent" might have high positive scores, while words like "sad," "hate," or "terrible" might have significant negative scores. To assess the sentiment of a piece of text (like a sentence or document), the *Bing Liu* lexicon adds up the scores of all the words present in the text. The cumulative score reflects the overall sentiment of the text.

- ***Loughran-McDonald***:

Unlike others, this dictionary focuses specifically on financial and economic contexts. It contains a comprehensive list of words categorized with sentiment scores based on their relevance to financial markets, especially within corporate financial documents. Words are labelled with six possible sentiments in financial contexts: "negative", "positive", "litigious", "uncertainty", "constraining", or "superflous".

- ***AFINN***:

In this lexicon the value assigned to each word falls within a spectrum spanning from negative five to positive five. Profane or offensive language carries a rating of negative five, while highly positive terms are rated as positive five. Words or phrases conveying neutrality receive a score of zero. The *AFINN* method computes the overall score of a sentence by summing up the assigned values of its individual words.
Results were analyzed using statistical methods and visualization techniques to identify prevailing sentiments, sentiment distribution, and significant recurring themes within Twitter discussions involving Airbnb.

IV. Evaluation

The vast majority of tweets composing the analysed sample carried a neutral sentiment toward the company; while the proportions of tweets expressing either a positive or negative sentiment were almost equivalent (Fig.1). As shown in Table I 'neutral' tweets accounted for almost 62% of the sample, while the remaining were almost equally split between 'positive' and 'negative' type of messages.

TABLE I
SENTIMENT DISTRIBUTION

	Positive	*Neutral*	*Negative*
Percentages	*18.3%*	*61.7%*	*20%*

Positive and negative type of tweets had a skewed distribution with a median polarity score not too distant from zero, representing a text carrying a 'neutral' attitude toward the company (Fig.2). The long tails were due to a few outliers expressing overwhelmingly positive and negative sentiments. Emotion scores by sentence were extracted as well leveraging the *'sentimentr'* R package, which is able to recognize eight different emotions: anger, anticipation, disgust, fear, joy, sadness, surprise and trust. Those scores were then used to derive the frequency of each emotion within the analysed sample. The top three most detected resulted to be *'Trust'*, *'Anticipation'* and *'Anger'* (Fig.3). The word cloud depicted in figure number 4 shows both the most frequently used positive and negative words within the sample. Figure number 5, 6 and 7 show the contribution to the overall sentiment of the words which were used at least 30 times according to the *Bing Liu*, *Loughran-McDonald* and *AFINN* lexicons respectively.

V. Conclusions and Future Work

The analysis unveiled that Twitter users discussing Airbnb predominantly used words associated with a positive sentiment, indicating overall satisfaction with the platform. This positivity indicates a general satisfaction with the Airbnb experience, highlighting aspects contributing to positive perceptions and areas that might require attention for improvement. Limitations of the analysis encompassed the potential bias in Twitter data, limited access to certain types of tweets, and challenges associated with accurately interpreting sentiments due to sarcasm or contextual nuances in text. Lexicon-based sentiment analyses offer a straightforward and quick method for gauging sentiment in text, however, have several limitations that can affect accuracy and reliability. Lexicon-based methods often struggle to grasp the nuances of language and context. They might misinterpret sentiment due to sarcasm, irony, or complex sentence structures, leading to inaccurate conclusions. Lexicons in addition have finite word lists, and they might not encompass newly coined words, slang, or domain-specific terminology. As language evolves, these collections can become outdated and miss out on newer sentiment-bearing words. Analyses solely based on individual words may overlook the impact of word combinations or phrases. Some phrases might convey a different sentiment than the sum of their individual words' (e.g., 'not bad' might be positive). Negation words (like 'not', 'no', 'never') and modifiers ('very', 'extremely') can reverse or intensify the sentiment of adjacent words. Lexicon-based methods might struggle to accurately interpret such linguistic cues. Also, words can have different meanings in various contexts and lexicon-based approaches may misinterpret such words with multiple meanings, leading to incorrect sentiment classification. In addition, sentiment analysis may overlook the tone, emotions, or sentiments expressed through emojis, or visual elements leading to an incomplete understanding of the overall sentiment. While lexicon-based sentiment analyses serve as a useful starting point these limitations highlight the need for more sophisticated approaches to overcome these challenges and provide more accurate results. Future research could delve deeper into incorporating multimodal data or longitudinal studies to capture evolving sentiments. Additionally, exploring sentiment dynamics across diverse demographics could provide more targeted insights.

References

[1] Minqing, H., Bing, L., "Mining and summarizing customer reviews.", Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery & Data Mining (KDD-2004), Seattle, Washington, USA, Aug 22-25.

[2] Pak, A., Paroubek, P., "Twitter as a Corpus for Sentiment Analysis and Opinion Mining". In Proceedings of the Seventh International Conference on Language Resources and Evaluation (LREC'10), Valletta, Malta. 2010. European Language Resources Association (ELRA).

[3] Jansen, B.J., Zhang, M., Sobel, K., "Twitter Power: Tweets as Electronic Word of Mouth". Journal of the American Society for Information Science & Technology, 60, 2169-2188. 2009. https://doi.org/10.1002/asi.21149

[4] Wei-Lun, C., Wang, J.Y., "Mine is yours? Using sentiment analysis to explore the degree of risk in the sharing economy". Electron. Commer. Rec. Appl. 28, C. Mar 2018. 141–158. https://doi.org/10.1016/j.elerap.2018.01.014

[5] Cheng, L.C., Tsai, S.L., "Deep learning for automated sentiment analysis of social media". 1001-1004. 2019. 10.1145/3341161.3344821.

[6] Twitter developer account: https://developer.twitter.com/en/apply-for-access

[7] Airbnb News: https://news.airbnb.com

[8] Hvitfeldt, E., Silge, J., Supervised Machine Learning for Text Analysis in R. 2015. Available at: https://smltar.com/

[9] Huston, S., and Croft, W. B., "Evaluating Verbose Query Processing Techniques." In Proceedings of the 33rd International ACM SIGIR Conference on Research and Development in Information Retrieval, 291–298. SIGIR '10. 2010. New York, NY: ACM. http://doi.acm.org/10.1145/1835449.1835499.

[10] Loughran, T., McDonald, B., "When is a Liability not a Liability? Textual Analysis, Dictionaries, and 10-Ks". (March 4, 2010). Journal of Finance, Forthcoming, Available at SSRN: https://ssrn.com/abstract=1331573

[11] Nielsen, F.A., "A new ANEW: Evaluation of a word list for sentiment analysis in microblogs". 2011. Proceedings of the ESWC2011 Workshop on 'Making Sense of Microposts': Big things come in small packages (2011) 93-98. Available at: http://ceurws.org/Vol718/paper_16.pdf

[12] Manning, C. D., Raghavan, P., and Schütze, H. 2008. Introduction to Information Retrieval. New York, NY: Cambridge University Press.

Appendix

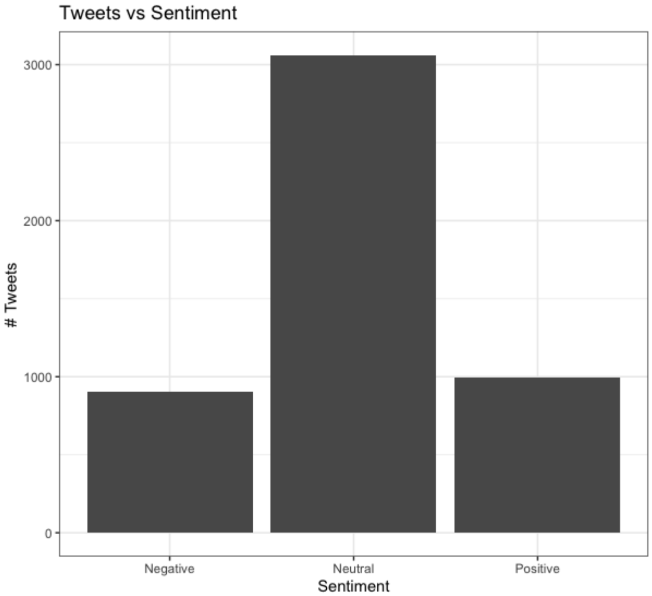

Fig. 1. Sentiment Distribution

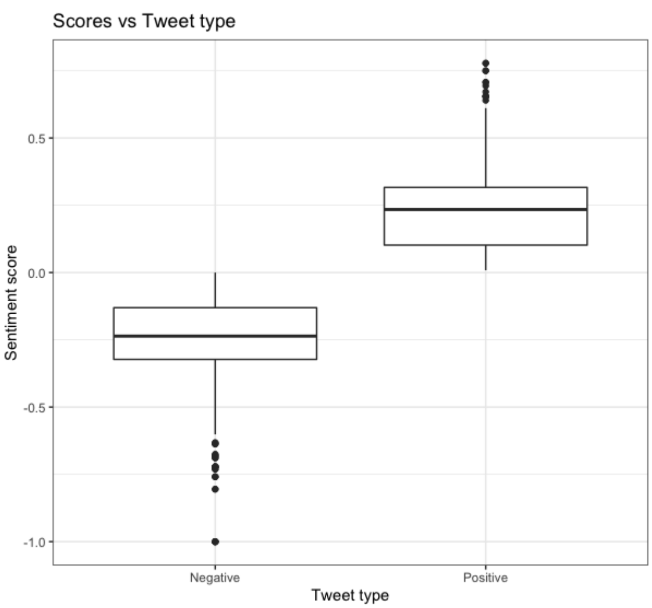

Fig. 2. Positive vs Negative Tweets Distribution

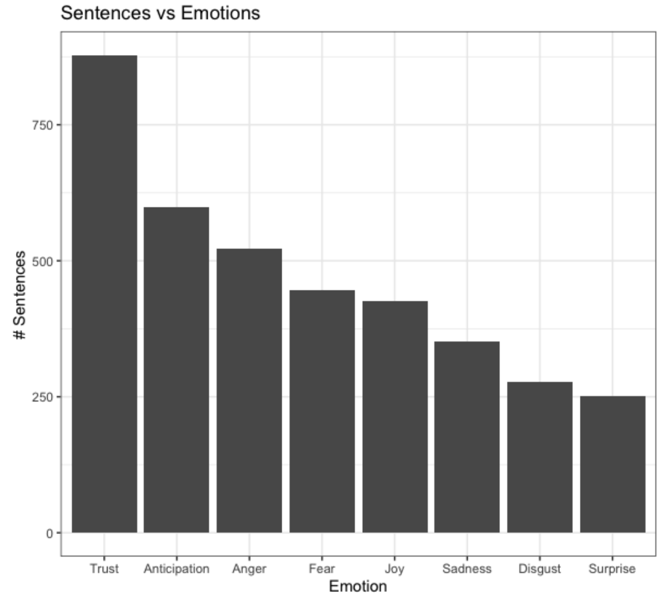

Fig. 3. Emotions Distribution

Fig. 4. Word cloud

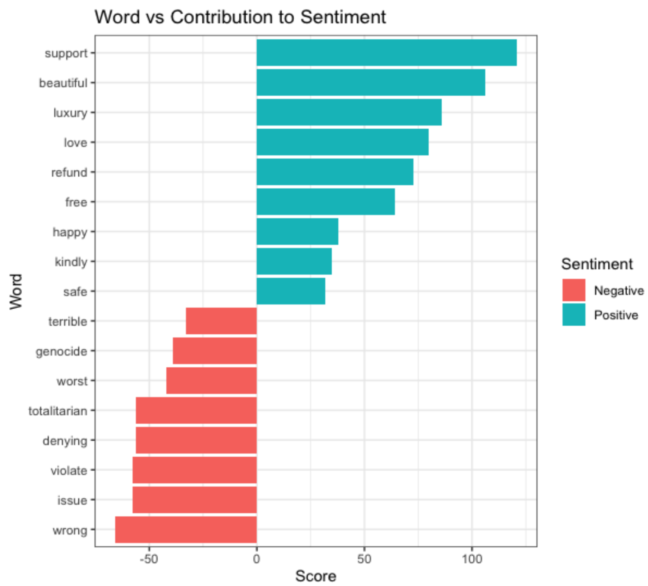

Fig. 5. Contribution to Sentiment: Bing Liu

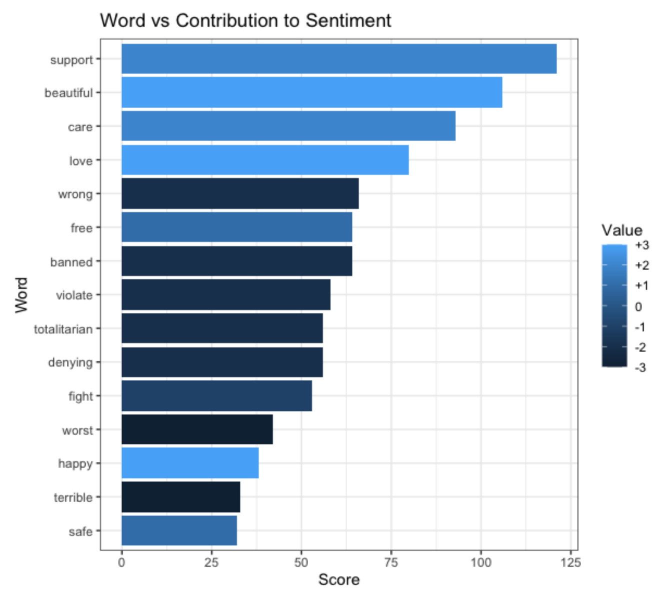

Fig. 7. Contribution to Sentiment: AFINN

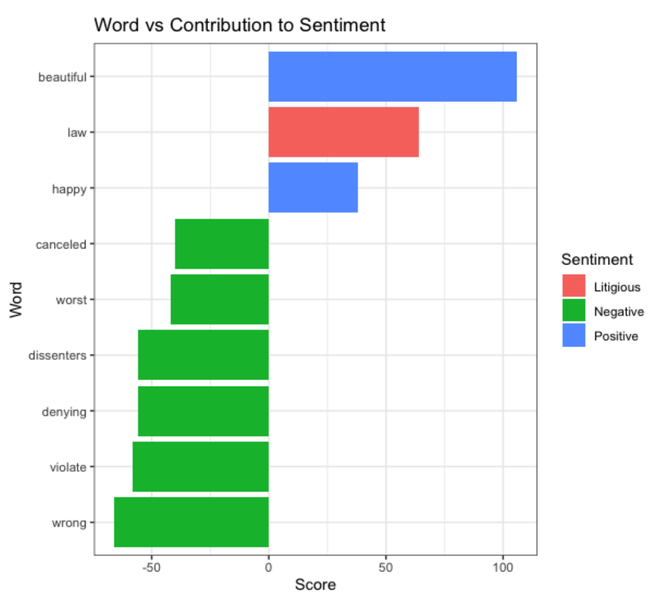

Fig. 6. Contribution to Sentiment: Loughran-McDonald