Spring 2024

Leveraging Sentiment Analysis for Tokyo Airbnb Hosts and Decision Makers

Shaikha Aldahmani
sa9895@rit.edu

Leveraging Sentiment Analysis for Tokyo Airbnb Hosts and Decision Makers

by

Shaikha Aldahmani

A Thesis Submitted in Partial Fulfilment of the Requirements for the Degree of Master of Science in Professional Studies: Data Analytics

Department of Graduate Programs & Research

Rochester Institute of Technology- RIT Dubai Campus

Spring 2024

RIT

Master of Science in Professional Studies:

Data Analytics

Graduate Thesis Approval

Student Name**: Shaikha Aldahmani**

Graduate Thesis Title**: Leveraging Sentiment Analysis for Tokyo Airbnb Hosts and Decision Makers**

Graduate Thesis Committee:

Name: Dr. Sanjay Modak Date:
Chair of committee

Name: Dr. Ioannis Karamitsos Date:
Member of committee

Acknowledgments

The completion of this thesis report marks a significant milestone, and it would not have been possible without the contributions and support of exceptional individuals who have enriched this journey. First and foremost, I extend my deepest gratitude to Dr. Ioannis Karamitsos, my thesis supervisor, whose expertise and guidance were invaluable in the development of this work. Also, I would like to express my heartfelt gratitude to my family and friends for their encouragement and belief in my abilities that fueled my determination and made this journey special and unforgettable.

Abstract

This study focused on the growing challenges that Tokyo Airbnb hosts and policymakers face in the holiday homes and home stays market. The study aimed to empower hosts with insights to optimize offerings and assist policymakers in informed decision-making for enhancing neighborhood experiences, with a focus on data-driven business strategies grounded in sentiment analysis. The project employed a robust methodology that leverages sentiment analysis techniques to gain insights from guest comments of Tokyo Airbnb listings spanning from September 2011 to June 2023. The CART algorithm and the Affin sentiment lexicon were used for data cleaning, consolidation, and sentiment analysis. The project aimed to bridge the gap between market popularity and data-driven strategies, ultimately benefiting Tokyo's tourism industry.

Keywords: Tokyo, holiday homes, Short-term vacation rentals, Airbnb, sentiment analysis, guest comments, data-driven insights, host strategies, policymaking, tourism industry, CART algorithm, Affin, sentiment patterns.

Table of Contents

List of Figures

List of Tables

Chapter 1

1.1 Introduction

The emergence of the Short-term vacation rentals (STR) and platforms have advanced the tourist destinations and the tourism industry as whole. Shor-term vacation rentals (STR) is a new kind of accommodation which means sharing economy, peer to peer consumption and the platform economy. They are privately-owned accommodation rental websites that link travelers with homeowners of available apartments, rooms, and homes for rent. One of the famous platforms is Airbnb. It provides a distinctive accommodation experience by offering a cultural or local experience from different hosts at a feasible price and encourages hosts and guests to express their opinions and post their experiences for the other hosts and guests (Agrusa, 2020).

With the surge in the number of peer-to-peer accommodation platform users, the large amount of customer reviews is perceived as a valuable source in understanding the customers' expectations. Reviews from customers are typically written in two formats: textual and numerical like the star ratings. While the numerical ratings from customers may not tell the whole story of the overall service quality. Textual reviews contain personal narrative of experiences and have a vital role in extracting nuanced sentiments (Lee, C. K. H., & Tse, Y. K., 2021).

Japan has undertaken many efforts to attract international tourists. In 2017, it welcomed 28.7 million foreign tourists, triple the number of international visitors in the last five years. Due to this instant growth, the government introduced in 2017 a new regulation permitting residents of Japan to rent out available rooms in their homes to tourists for revenue (Agrusa, 2020). In 2019, 41,000 listings in Japan operated on Airbnb and its features were used 5 million times by international tourists in 2018. This trend is expected to continue with the 2020 Olympics in Japan (Agrusa, 2020).

In contrast, Tokyo, as von Briel & Dolnicar (2020) noted, is one of the protective cities that firmly regulate short-term rentals. This includes forcing regulations on the permitted locations, operating hours and ways of rentals, without any revisions. Moreover, Tokyo initiated a council-managed system specialized in receiving residents' reports and complaints of short-term letting. Despite these rigorous regulations, the Tokyo holiday home market is thriving and adapting to these regulations.

Despite receiving millions of reviews daily, few figured out the significance of using this data in assessing the customer satisfaction, understanding the important features in short-term accommodation rentals to Airbnb customers and informing host decisions (Güçlü, Roche & Marimon 2019). Therefore, it is important to apply sentiment analysis and other text mining techniques to understand customers preferences and enhance decision making processes for hosts and decision makers.

1.2 Statement of the problem

The holiday homes and home stays market has gained huge popularity and demand in the last few years especially in Tokyo. However, hosts lack data-driven business strategies based on sentiment analysis to increase revenue, and tailor offerings to align with guest preferences based on past sentiment data. Additionally, policy makers lack the insights needed to make informed decisions to improve neighborhood experiences. To address these challenges, it is crucial to leverage sentiment analysis to empower both hosts and policy makers. The purpose of this analysis is to improve the profitability of holiday homes and home stays by providing hosts with insights to optimize their offerings and helping policy makers create more effective decisions based on guest sentiment.

1.3 Research Aim and Objectives

This project aimed to provide Tokyo Airbnb hosts with data-driven insights obtained from sentiment analysis to improve their business strategies, guest satisfaction, and increase revenue. It specifically aimed to understand sentiment patterns and tailor services to align with guests' preferences based on past sentiment data. This project aimed to bridge the gap between the growing popularity of the holiday home and home stay market and the lack of data-driven strategies for hosts and informed decision making. The following research objectives:

1. To provide Tokyo Airbnb hosts with actionable sentiment analysis insights. This will allow them to fine-tune their business strategies, improving service quality and increasing revenue.
2. Assist policymakers in making informed decisions about the regulation and development of the holiday home market. The insights gained from sentiment analysis can be used to guide the development of effective regulations to improve the overall experience in Tokyo neighborhoods.
3. Contribute to the growth and sustainability of Tokyo's tourism industry by creating a mutually beneficial environment for hosts, guests, and local communities.

1.4 Research Questions

In this thesis the following research questions are

RQ1: What are the common sentiments expressed by guests in reviews of Airbnb listings in Tokyo?

RQ2: How can sentiment analysis of Tokyo Airbnb listings' guest reviews (from Sep 2011 to Jun 2023) be used to find sentiment patterns, and equip hosts with tailored offerings based on guests' preferences?

RQ3: How can this analysis help policy makers in making informed decisions to improve neighborhood experiences for tourists?

1.5 Research Methodology

1.5.1 Sentiment analysis process

The sentiment analysis of the Airbnb comments process involved seven distinct stages, as outlined in Figure 1.

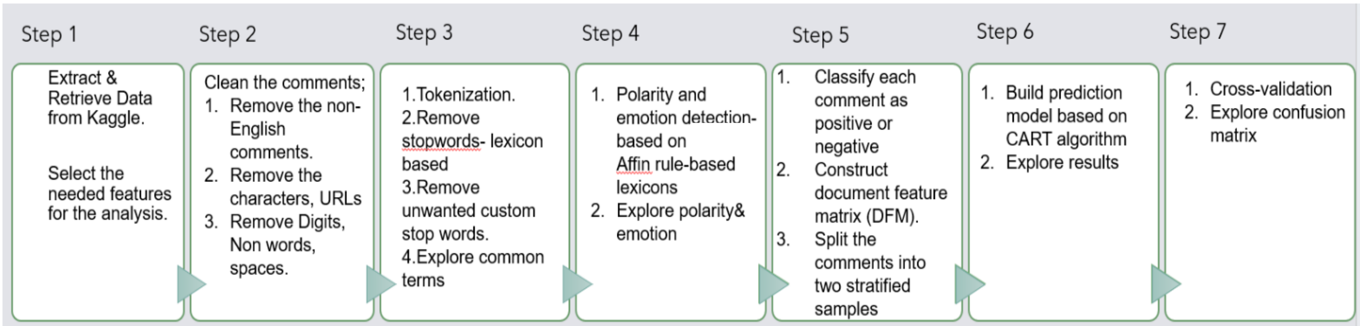

Figure 1: Sentiment Analysis methodology of Airbnb

STEP 1: To begin the analysis, missing values and data quality issues were addressed, by carefully cleaning the collected data.

STEP 2: To streamline the analysis, relevant variables such as ID, date, and comments were chosen, while irrelevant variables were removed. Next, the selected columns were consolidated, and their formats were standardized for consistency.

The comment text underwent a thorough cleaning, with hashtags, URLs, unwanted strings, digits, non-words, and extra spaces removed. Furthermore, the text was converted to lowercase, and only English comments were retained, while comments in other languages were discarded.

STEP 3-4: Tokenization was then applied, and the Affin lexicon dictionary containing negative and positive words was used to detect emotions.

STEP 5: Classify comments into positive and negative categories. The comments were divided into two groups, and a document feature matrix was constructed.

STEP 6-7: For robust sentiment analysis, a prediction model was developed using the CART algorithm and the evaluation of the algorithm using confusion matrix.

1.6 Limitations of the Study

Although the study provides valuable insights into sentiment analysis and classification within the scope of English language data, we cannot generalize the results. The non-English language comments were excluded from the analysis, and a limited dataset sample size was used for the sentiment classification of the comments and creation of the document-feature matrix (DFM) and machine learning prediction.

Chapter 2 – Literature Review

Since its inception in 2008, Airbnb has become a prominent player in the sharing economy, offering a platform for peer-to-peer accommodation. Airbnb's platform has brought numerous benefits not only to property owners but also to the local communities where these accommodations are located. Airbnb revenue contributes to the financial well-being of residents, helping them meet expenses and repay loans. Additionally, it provides local governments with increased tax revenue through sales and lodging taxes (Agrusa, 2020). Agrusa's (2020) study highlights the importance of striking a balance to sustainably develop a tourism industry that caters to both the preferences of residents and the needs of visiting tourists. Japan, one of Airbnb's fastest-growing markets, introduced new regulations in 2018 to manage the platform's expansion. Hosts are now required to register their accommodations and display their license numbers on their listing pages in order to continue operating on Airbnb. Furthermore, to protect both tourists and the community, the government implemented strict guidelines. Short-term rental periods were limited to less than 180 days per year, and Airbnb operations were restricted during peak tourist seasons. For instance, in Kyoto, a renowned tourist destination in Japan, home-sharing in residential areas was restricted to specific periods such as mid-January to mid-March or during non-tourist seasons.

To navigate these regulations and enhance the experience for both guests and hosts, it is essential to analyze guest reviews and leverage the power of sentiment analysis. This literature review aims to examine the key findings and insights from previous research, with a focus on the application of sentiment analysis in the context of the tourism industry, Airbnb, and other short-term accommodation platforms.

2.1 Sentiment analysis in the tourism industry

Sentiment analysis is one of the natural language processing techniques that can apply classification on sentiments based on online reviews. It enables the identification of positive, neutral, and negative opinions expressed in any document, paragraph, or sentence (Lee, & Tse

,2021). It is one of the effective techniques to be used in the tourism industry, sharing economy or the peer-to-peer accommodation development such as in developing business strategies, predicting demand and in policy making. The increasing interest in automated methods, such as text mining and sentiment analysis, is highlighted by He et al. (2017), they emphasized their efficiency in processing large amounts of user-generated data and extracting meaningful knowledge and insights. Text mining, as an emerging technology, aims to instantly extract meaningful information from many textual documents. Mohammad, Kiritchenko & Zhu (2013) created a sentiment lexicon generation scheme that employs SVM classifiers to detect sentiment at both the message and term levels, resulting in success in sentiment analysis tasks, particularly in social media contexts.

In hotel management, sentiment analysis helps identify patterns in guest reviews to understand factors influencing satisfaction and recommendations. Berezina et al. (2015), analyzed and compared 2,510 online reviews from satisfied and dissatisfied hotel customers in Florida using a text-mining approach. They discovered that satisfied customers are more likely than dissatisfied customers to mention intangible aspects of their hotel stay, such as staff members. Dissatisfied customers, on the other hand, frequently mention tangible aspects of their hotel stay, such as furnishings and finances. This emphasizes the significance of sentiment analysis in extracting nuanced patterns from customer feedback. Likewise, Güçlü, Roche & Marimon (2020) looked at Airbnb customer reviews to understand customer satisfaction. They focused on five top tourist cities in Europe. They found that overall satisfaction with Airbnb is high. Customers are drawn to cities for their brand or entertainment offerings. Barcelona and Istanbul are seen as entertaining, while London, Paris, and Rome are known for their city brands. The most desired features are housing and accommodation features. These findings are useful for governments, tourism organizations, and Airbnb.

Bagherzadeh et al. (2021), presents a novel and efficient sentiment analysis methodology for TripAdvisor hotel reviews, with the goal of overcoming limitations in existing methods. Their study collects a large amount of data from TripAdvisor and conducts sentiment analysis using a bag-of-words weighted approach, resulting in two hotel-specific lexicons (L1 and L2). The performance of these lexicons is superior to that of a public dictionary-based method and a complex machine-learning algorithm. The proposed methodology provides hotel managers with a dependable tool for analyzing guest feedback, understanding market attitudes, and anticipating changes. The study is unique since it uses a bag-of-words approach for sentiment analysis and creates a field-specific lexicon.

Ye, Zhang, & Law (2009), research focuses on sentiment classification of online travel reviews using supervised machine learning approaches. The study compares three supervised machine learning algorithms: Naive Bayes, SVM, and a character-based N-gram model, and incorporates sentiment classification techniques into mining travel blogs for seven popular destinations in the United States and Europe. The results show that SVM and N-gram approaches outperform Naive Bayes, with all three approaches achieving at least 80% accuracy when training datasets have many reviews. The study highlights the importance of sentiment classification in analyzing the vast amount of information available on travel blogs.

Santos et al. (2022), explore sentiments in online reviews of Airbnb users in Fortaleza, Ceará, Brazil. This is using sentiment analysis to identify positive and negative opinions after analyzing 2353 reviews from 2019 and 2020 related to 506 Airbnb accommodation offers. They discovered a predominance of positive feedback, with women, superhosts, private rooms, and residential neighborhoods.

Brochado, Troilo, and Shah (2017) studied the Airbnb customer experience in three countries: India, Portugal, and the United States. They identified themes such as stay, host, place, location, apartment, room, city, and home after analyzing 1776 text reviews from 24 Airbnb properties (8 from each country) using Leximancer, a text analysis software. The study's findings revealed evidence of convergence, indicating that the dimensions of post-experience shared in web reviews were highly similar across the three culturally distinct countries. This finding adds to the fact that the cultural norm of individualism has little impact on customer preferences for Airbnb experiences in these countries.

Sentiment analysis is critical not only in understanding customers' needs and expectations, but also in addressing complex challenges in the tourism industry, especially demand prediction. Colladona et al. (2019), undertook a pioneering effort to forecast tourism demand using social media data obtained from a TripAdvisor travel forum. The database comprises over 2.6 million posts from approximately 147,000 users. Using semantic analysis and social network techniques, the researchers were able to extract important variables from user-generated content predicting anticipated tourist arrivals in seven European cities. The findings, especially from the factor augmented autoregressive model and bridge model, demonstrated the significance of language complexity and communication network centralization as important predictors for precise forecasting. The study found that incorporating these extracted variables into conventional forecasting models outperformed models based solely on volume-based web search query data, highlighting the value of social media data in refining tourism demand forecasts.

Folgieri, Tea, & Maja (2018), shed light on the challenge of predicting tourist inflows early in the growing tourism industry, with a focus on optimizing local economies and developing tourist income. The authors highlight the ongoing difficulty in forecasting tourist demand due to heterogeneous data that traditional statistics struggle to analyze. The study introduces a backpropagation Artificial Neural Network (ANN) as an innovative Machine Learning approach and compares its performance in predicting tourist arrivals in Croatia to traditional linear regression methods. The study advocates for the novel application of algorithms, such as ANN, in the tourism industry by combining insights from Tourism Economics and Information Technology. The results show that the neural network model outperforms linear regression techniques in predicting tourist arrivals, highlighting the potential for proactive decision-making and service improvements in the tourism industry.

Hu. et al. (2022), focused their efforts on improving the forecasting accuracy of international tourist arrivals in Hong Kong from seven English-speaking countries using tourist-generated online review data. Their comprehensive framework combined weekly user-generated online reviews with monthly tourist arrival data from August 2012 to June 2019. Using the SARIMA-MIDAS Model, which incorporates high-frequency online review volume and average review rating, the study found significant improvements in forecasting accuracy when compared to traditional time-series models such as SARIMA and SNAIVE. The findings highlighted the importance of incorporating high-frequency online review variables into tourism demand forecasting, providing insights for informed decision-making in the tourism industry.

2.2 Enhancing Business Strategies through Sentiment Analysis

Various studies have used advanced analytical techniques to gain insights from online reviews and offer personalized recommendations to travelers to optimize business strategies in the tourism industry. Lee & Tse (2021), conducted a comprehensive study using text analytics techniques to improve peer-to-peer accommodation services and develop new service strategies. The researchers used topic modeling, sentiment analysis, and Process Chain Network (PCN) analysis on a dataset of 216,528 reviews from guests who used Airbnb services in London in 2017. Latent Dirichlet allocation (LDA) was chosen in their study because it considers each review as a mixture of latent topics, effective at identifying underlying topics in a large volume of documents and without being strained with grammatical rules. Moreover, the VADER model was applied for sentiment scores from -1 to +1, and reviews categorization into positive, negative or neutral. Their findings emphasized the critical role of non-human resources and technology in improving consumer experiences, highlighting key contributors to negative experiences such as a lack of hot water, poor sleep quality, and unpleasant check-in processes. The study emphasizes the importance of maintaining physical amenities, providing professional services, and improving interactive information formats on P2P platforms to meet guest preferences, demonstrating the benefit of text analytics in service improvement.

Madhi & Alhammad (2021), Sentiment analysis was conducted in Amsterdam to determine the factors influencing customer satisfaction with Airbnb. Using Airbnb guest reviews and listings from 2019, the study applied machine learning techniques such as sentiment analysis, word clustering, and ordinal logistic regression. Their findings revealed that factors such as property price, cleanliness, host communication, and the accuracy of property descriptions have a significant impact on guest satisfaction. Furthermore, precise location within neighborhoods emerged as a critical determinant of guest experience, emphasizing the importance of positive customer experiences and effective host interactions in increasing satisfaction levels.

Xiang et al. (2015), text analytics were used on a large dataset of Expedia.com consumer reviews to investigate the relationship between hotel guest experience and satisfaction. The study, which covered the 100 largest cities in the United States, gathered 60,648 reviews from 10,537 hotels. The researchers used factor analysis and linear regression to identify words related to guest experience that had high explanatory power for satisfaction ratings. Their findings revealed distinct semantic compositions of guest experience dimensions, emphasizing the importance of hygiene factors in hotel services. The study emphasizes the strong relationship between satisfaction and guest experience, as well as the importance of theoretical frameworks in guiding data analysis and informing service improvements.

Taecharungroj & Mathayomchan (2019), embarked on a thorough investigation aimed at improving business strategies by analyzing online reviews of tourist attractions in Phuket, Thailand. Using a large dataset of 65,079 TripAdvisor reviews covering various attractions such as markets, beaches, islands, temples, and pedestrian streets, the researchers used latent Dirichlet allocation (LDA) and naive Bayes modeling to uncover distinct dimensions for each attraction type. They used innovative tools such as lexical salience-valence analysis (LSVA) and dimensional salience-valence analysis (DSVA) to evaluate and interpret the findings, shedding light on factors such as review frequency, positivity, and the significance of the underlying terms. The study's findings helped the Tourism Authority of Thailand (TAT) make decisions about crowd management, commercialization, beach cleanliness, and visitor preferences for serene beachscapes. Furthermore, the study suggested alternative activities, such as hiking to viewpoints, to alleviate congestion at popular tourist destinations, emphasizing the importance of aligning marketing campaigns with visitor attitudes and leveraging data-driven promotional strategies in destination management.

On a similar note, Abbasi-Moud, Vahdat-Nejad, & Sadri (2021), introduced an innovative tourism recommendation system, which helped to advance business strategies. The system uses advanced text processing, semantic clustering, and sentiment analysis to extract user preferences from social media reviews, followed by feature extraction of tourist attractions based on factors such as weather conditions. The recommendation system uses contextual data such as time, location, and weather to compare user preferences with attraction features and make tailored recommendations. When tested with TripAdvisor data, the system outperformed comparable systems in terms of metrics such as f-measure, precision, and recall. The study emphasizes the importance of incorporating more contextual information and expanding the system to accommodate group scenarios, highlighting its potential to improve decision-making processes in the tourism industry.

Mavroforaki & Kolovou (2020), Use sentiment analysis to gauge tourists' attitudes toward Athens, Greece, across multiple online platforms. The study identifies positive traveler attitudes toward Athens by analyzing data from platforms such as Twitter, Google Maps, Foursquare, and Airbnb, highlighting the city's culinary and entertainment offerings while addressing issues such as noise pollution and cleanliness. Using machine learning algorithms and deep learning approaches, the study highlights the potential of big data analytics to improve tourism research, paving the way for data-driven practices to raise industry standards.

Raza,Hussain, & Varol, A (2022), Emphasize sentiment analysis of Airbnb reviews, acknowledging the critical role consumer feedback plays in shaping service quality and rental decisions. The study uses deep learning algorithms such as RNN, LSTM, and GRU to assess sentiment classification accuracy, with GRU emerging as the top performer. Their findings emphasize the importance of sentiment analysis in helping hosts effectively manage service aspects and reputation, which is critical for guests making informed lodging decisions.

Nazirkar & Kulkarni (2023), Contribute to sentiment analysis by using Long Short-Term Memory (LSTM) to predict sentiment in Airbnb reviews. They achieve impressive accuracy metrics by breaking down their study into sentiment analysis using LSTM and topic modeling, shedding light on the factors that influence customer satisfaction and dissatisfaction. The research, which employs LSTM and topic modeling techniques, provides valuable insights for Airbnb hosts to improve customer experiences and emphasizes the importance of comprehensive sentiment analysis in understanding guest sentiments.

Bandara, Charles, & Lekamge (2022), through sentiment analysis and topic modeling of online reviews, this article investigates the accommodation experience in the sharing economy, with a focus on Airbnb. The study employs a dataset of 401,964 Airbnb review comments to identify factors influencing the accommodation experience. Word cloud, frequency distribution, and topic modeling data analysis techniques reveal key factors such as location, safety, host-guest interaction, amenities, proximity to restaurants and transit options, and apartment uniqueness. According to the findings, these factors play a significant role in improving the services offered by peer-to-peer accommodation platforms, allowing them to better cater to customer preferences and streamline operations.

Ding et al. (2020), Investigate service quality attributes in the context of Airbnb accommodations in Malaysia, using a large dataset of 242,020 reviews spanning five years. Using structural topic modeling (STM), the study identifies 22 service-related topics, shedding light on the nuances of

Airbnb users' perceptions. Notably, the study introduces four novel service attributes that have previously been overlooked in Airbnb research. The study identifies distinct preferences between Malaysian and international users by examining the compatibility of a modified SERVQUAL questionnaire with Airbnb service quality attributes, with a focus on factors such as property appearance, location, and host communication. While the findings provide valuable guidance for improving service quality and customer satisfaction, the study emphasizes the need for future research to explore variations across platforms and investigate factors influencing changing perceptions.

Li, Chen, & Huang (2020), contribute to understanding guest sentiments in the Airbnb ecosystem versus traditional hotel experiences, using Sydney as a case study. The study uses text mining and sentiment analysis on extensive online review data to elucidate factors influencing Airbnb consumers' experiences, ranging from positive feelings about atmosphere and flexibility to concerns about value for money and cleaning responsibilities. The study not only demonstrates the theoretical value of big data analytics in tourism and hospitality studies, but it also provides practical insights for Airbnb developers, operators, and hosts to make more informed reservation decisions and strategic improvements. Despite its useful contributions, the study suggests opportunities for additional research, comparisons with other platforms, and the development of a specific lexicon for the tourism and hospitality industry.

2.3 Utilizing Sentiment Analysis to Make Informed Policy Decisions

Stylos, Zwiegelaar & Buhalis (2021), investigate the transformative potential of Big Data (BD) in dynamic service industries, with a special emphasis on tourism, and its implications for decision-making processes. Through an ethnographic study involving online focus groups with information technology and database professionals, the research reveals the profound impact of BD in predicting customer behavior and facilitating tailored propositions, thereby improving organizational agility, particularly in fast-paced industries. The study proposes an integrated BD framework that addresses critical dimensions such as stakeholder needs, spatial and temporal considerations, and utility contextual factors, providing a comprehensive roadmap for effective BD management and application in decision-making processes across diverse business domains.

Alonso-Almeida, Borrajo-Millán, & Yi (2019), provide critical insights into the intersection of social media data and overtourism, with a focus on Barcelona and Chinese tourists. Their use of big data techniques and sentiment analysis sheds light on how social media influences tourist behavior and worsens congestion in popular destinations. By proposing policy measures to mitigate the negative effects of overtourism, such as dispersing travelers and leveraging digital platforms for destination promotion, the study emphasizes the importance of proactive policymaking in response to changing tourism dynamics.

Park, Kim, & Seong (2021), contribute to urban policy making by using Twitter data to investigate the relationship between neighborhood conditions and sentiments in Detroit, Michigan. Their findings emphasize the importance of safety, amenities, and demographic composition in promoting happiness in urban communities. By identifying positive and negative sentiment hotspots, the study helps policymakers prioritize resources and interventions in shrinking cities. Despite the inherent limitations of social media data, the study emphasizes the value of social listening as an evidence-based policy making tool in urban development and revitalization efforts.

Bertrand et al. (2013), present a ground-breaking study on measuring public sentiment in New York City using Twitter data. Using sentiment analysis techniques tailored to the unique characteristics of tweets, the researchers created a classifier capable of determining an individual's mood based on their Twitter activity. The study uses user-provided geotagging to analyze sentiment on fine-grained spatial and temporal scales, providing insights into the city's mood at various locations and times. The study identifies daily and weekly sentiment fluctuations, with weekend tweets being more positive than weekday tweets and a daily sentiment peak occurring around midnight. This temporal analysis provides insights into how public mood changes over time, allowing policymakers to consider potential interventions to improve community well-being and satisfaction. This study shows how social media analytics can inform urban policy development by using Twitter data for sentiment analysis.

Kirilenko, Stepchenkova, Kim, & Li (2018), contribute to the discussion on sentiment analysis's applicability in tourism, hospitality, and marketing studies by comparing the performance of automated sentiment analysis classifiers to human raters. Using real-world datasets such as surveys, TripAdvisor reviews, and Twitter messages, the study compares the effectiveness of machine learning methods to human judgment. While certain classifiers, such as Naive Bayes and SVM, outperform human raters on specific datasets, the study emphasizes the importance of cautious interpretation and validation, considering variations in classifier performance across diverse datasets. With nuanced insights into the strengths and limitations of automated classifiers, the study emphasizes the importance of careful selection and rigorous evaluation to ensure the reliability and validity of sentiment analysis results when informing policy decisions.

Schmunk et al. (2013), conducted a study to extract decision-relevant insights from user-generated content (UGC), with a focus on the Swedish mountain tourism destination Are. The researchers used an innovative approach to compare various data mining techniques for accurate topics and sentiment detection in textual user reviews. The study uses four automated sentiment analysis tools, including lexicon-based and machine learning methods, to analyze reviews from ski resort visitors obtained from TripAdvisor.com and Booking.com. The study demonstrates the effectiveness of data mining methods, particularly lexicon-based classification approaches, in automatically extracting and analyzing customer feedback using a meticulous sentiment analysis methodology consisting of six tasks. The findings of this analysis provide useful inputs for decision-making processes, facilitating product optimization and CRM activities within tourism destinations. The study advocates for potential improvements, such as expanding wordlists, using larger sets of training data, using web services to retrieve reviews, and addressing spelling errors.

The article of Alamanda et al. (2019), contribute to the sentiment analysis discourse by researching traveler reviews on social media platforms, with a focus on Garut's tourist attractions in Indonesia. The study, which uses a dataset of 413,175 netizen comments from Instagram and Google reviews, sheds light on the economic significance of tourism in Garut as well as the challenges the industry faces. Despite the abundance of positive feedback, concerns about hygiene, particularly in beach tourism, emerge as significant deterrents. Using a descriptive qualitative approach combined with data mining techniques, the study emphasizes the importance of ongoing efforts to improve tourism experiences. By prioritizing infrastructure and cleanliness improvements that are consistent with the Sapta Pesona elements of tourism charm, the study advocates for a strategic roadmap for local governments to effectively address these challenges.The findings highlight the critical role of netizen sentiment in shaping tourism strategies and priorities, emphasizing the importance of a comprehensive approach to destination management.

De Marchi, Becarelli, & Di Sarli (2022), address the changing landscape of tourism sustainability measurement by introducing the Tourism Sustainability Index (TSI). By combining sentiment analysis of online content with open data sources, the TSI provides a georeferenced evaluation of tourism sustainability that outperforms existing frameworks such as the European Tourism Indicator System (ETIS). The Data Appeal Company created the TSI, which includes Destination Management, Overtourism, Social and Cultural Aspects, and Environment, to provide nuanced insights into sustainability at various geographical levels. The study emphasizes TSI's theoretical and managerial implications, highlighting the importance of continuous data updating and visualization for destination stakeholders to make informed decisions. By incorporating sentiment analysis into sustainability assessments, the TSI ushers in a paradigm shift in evaluating tourism sustainability, facilitating proactive interventions.

2.4 Main Takeaways from the Literature Review

1. Sentiment analysis is a powerful tool for understanding nuanced guest opinions, emphasizing the importance of both tangible and intangible factors in shaping customer satisfaction in the tourism sector.
2. Text analytics techniques used by researchers include Structural Topic Modelling (STM), Latent Dirichlet Allocation (LDA), and sentiment analysis using Naive Bayes, SVM, and GRU. This toolkit allows for the extraction of nuanced insights from massive amounts of textual data.
3. Tokenization is a fundamental step in natural language processing, and the construction of a document-feature matrix is a common approach in text analysis, particularly when using techniques such as topic modeling and sentiment analysis.
4. The utilization of sentiment lexicons enriches sentiment analysis by providing a comprehensive set of positive and negative words for emotion detection and comment classification.

Chapter 3- Project Description

This research adds to the body of knowledge in sentiment analysis through academic publications, providing insights into the unique challenges and opportunities presented by the sharing economy in the context of Tokyo's Airbnb ecosystem.

Offers an interactive and user-friendly visualization dashboard to present sentiment analysis results, giving hosts and decision makers intuitive insights into the prevalent sentiments surrounding Airbnb accommodations in Tokyo for improvement.

The dataset used in this research was retrieved from Kaggle, specifically from the Inside Airbnb project, on December 6, 2021. It is a sample of 10.7k listings and contains detailed listings data and review data files of current Airbnb listings in Tokyo as of October 28, 2021.

For the sentiment analysis, two datasets were joined to create a more comprehensive analysis of the holiday homes market. The Reviews dataset included reviews for each listing and was selected for the comment's analysis. It contains reviews from September 2011 to June 2023. The essential variables were listing ID, date, comments, and reviewer ID. The Summary Listings dataset included more details related to the reviews and listing details, such as the listing name, neighborhood name, location, and room type. The datasets were joined using the left_join function from the dplyr package in R based on the common column, which is "listing_id".

https://www.kaggle.com/datasets/lucamassaron/tokyo-airbnb-open-data-2023?select=reviews.csv

The data sets dictionary:
https://docs.google.com/spreadsheets/d/1iWCNJcSutYqpULSQHlNyGInUvHg2BoUGoNRIGa6Szc4/edit#gid=360684855

Chapter 4- Analysis

4.1 Preprocessing Data

In sentiment analysis, one of the initial steps after retrieving datasets and selecting relevant features or variables is data cleaning, as shown in (Figure 1). For the review dataset, the following important variables were selected: Listing ID, comments, date, and reviewer ID. During the cleaning process, missing values were removed from the comments, selecting only English comments, and non-English comments were discarded. The resulting dataset contained 277024 English comments. These comments were then cleaned by removing characters, URLs, digits and spaces.

4.2 Tokenization of comments

The crucial step in the sentiment analysis methodology is the applications of tokenization methods. The total word count was 5,587,843 and after removing stop words, the most common or frequent words were analyzed, as shown in Figure. 2. Additionally, the wordcloud() function in R was used to create a word cloud with the most frequent words (Figure 3).

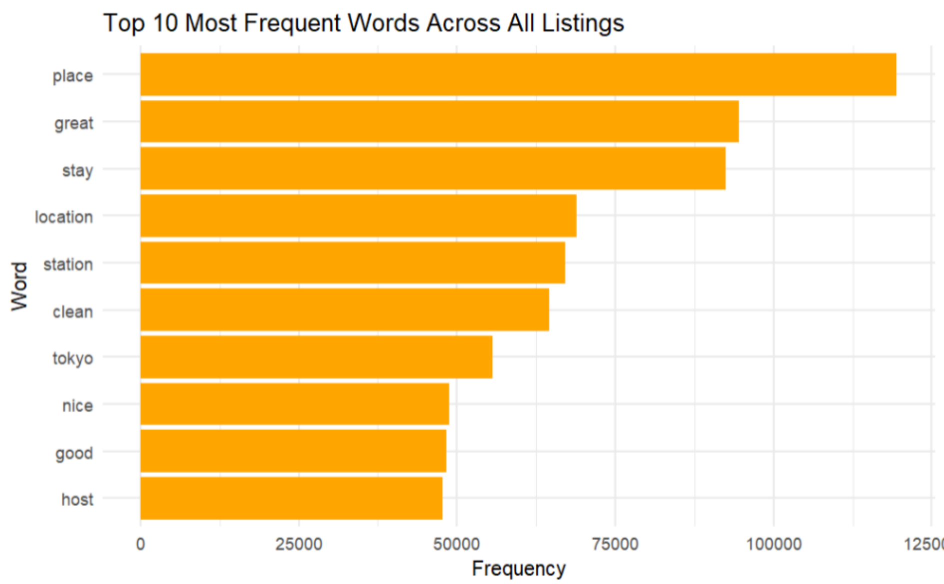

Figure 2: Most 10 frequent words across listings.

Figure 3: Word Cloud of most frequent words across listings.

4.3 Polarity and Emotion Detection using AFINN Lexicon

One common tool used in polarity and emotion detection is the AFINN sentiment lexicon. The AFINN lexicon created by Finn Arup Nielsen and it is a list of English words that are rated with a whole number between minus five(negative) and plus five (positive). The ratings are assigned manually and are intended to capture the sentiment of the word based on the meaning.

Affin sentiment lexicon was used for polarity and emotion detection. As shown in (Figures 4 & 5), the distribution of positive emotions and sentiments is greater than the distribution of negative emotions and sentiments across all the listings.

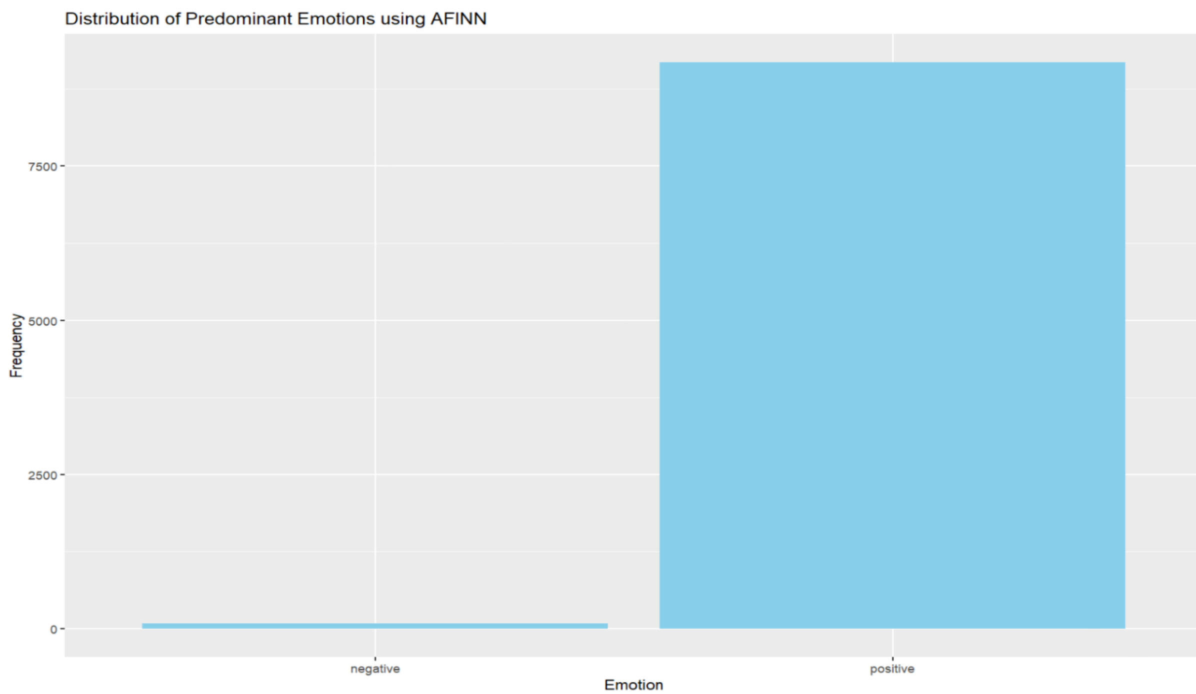

Figure 4: The distribution of emotions/sentiments across listings

In the Figure. 5 shows the changes in the sentiment score over time between 2011 and 2023 was then plotted.

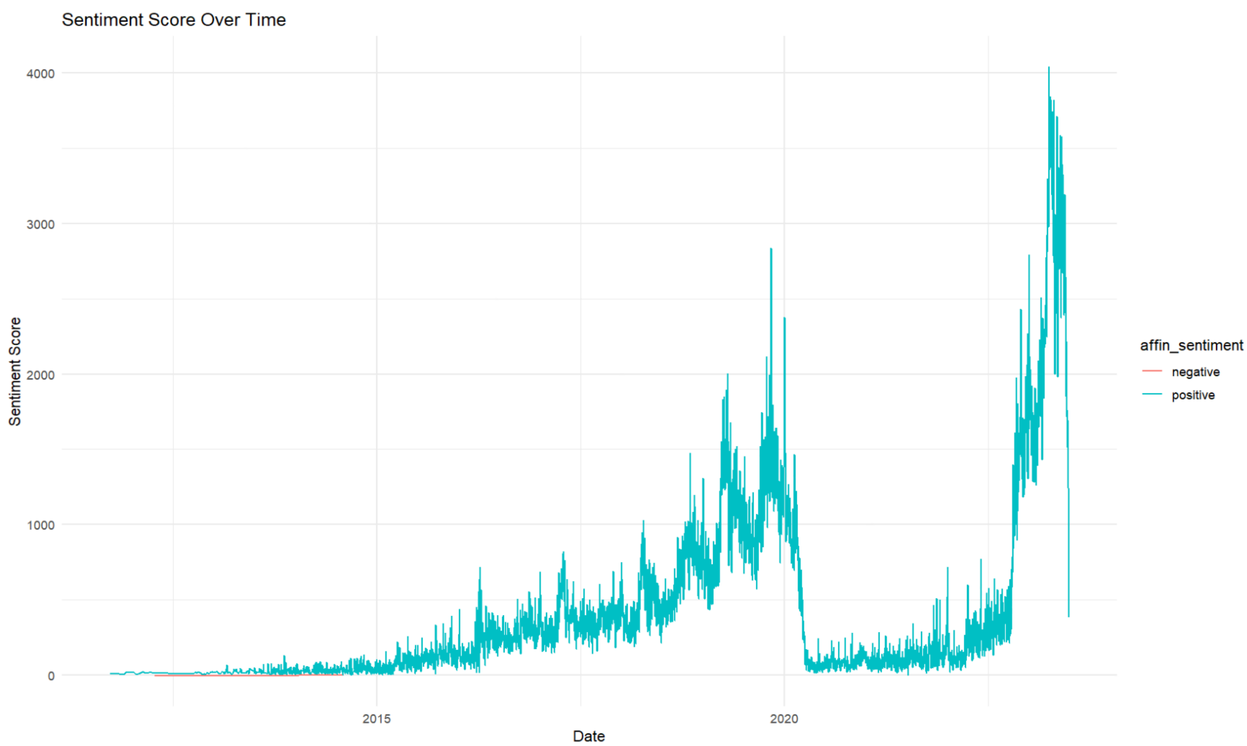

Figure 5: Sentiment score using Affin from (Sep 2011 to Jun 2023)

Figure. 6 and Table.1 show the average sentiment score across the neighborhoods. The top five neighborhoods with the highest average sentiment scores are:

1. Okutama Machi (2.71)

2. Musashimurayama Shi (2.29)

3. Ome Shi (2.20)

Figures 6 & 7 illustrate the most common positive and negative words used in the comments across all listings. The most common positive words (great, clean, nice, good, recommend) are shown in the word clouds, while the most frequent negative words are (problem, hard, stop, noisy, bad, dirty).

Table 1: Sentiment score per Neighborhood

A tibble: 46 x 3

neighbourhood <chr>	affin_score <dbl>	affin_sentiment <chr>
Adachi Ku	11225	positive
Akiruno Shi	619	positive
Akishima Shi	346	positive
Arakawa Ku	24542	positive
Bunkyo Ku	26628	positive
Chiyoda Ku	15949	positive
Chofu Shi	3482	positive
Chuo Ku	43549	positive
Edogawa Ku	21767	positive
Fuchu Shi	1519	positive
Fussa Shi	446	positive
Hachioji Shi	2298	positive
Hamura Shi	316	positive
Higashimurayama Shi	4280	positive
Hino Shi	2898	positive
Itabashi Ku	16193	positive
Katsushika Ku	41565	positive
Kita Ku	36294	positive
Kodaira Shi	587	positive
Koganei Shi	2244	positive
Kokubunji Shi	2012	positive
Komae Shi	1798	positive
Koto Ku	36333	positive
Kunitachi Shi	314	positive
Machida Shi	4541	positive
Meguro Ku	19620	positive
Minato Ku	79211	positive

1-27 of 46 rows   Previous 1 2 Next

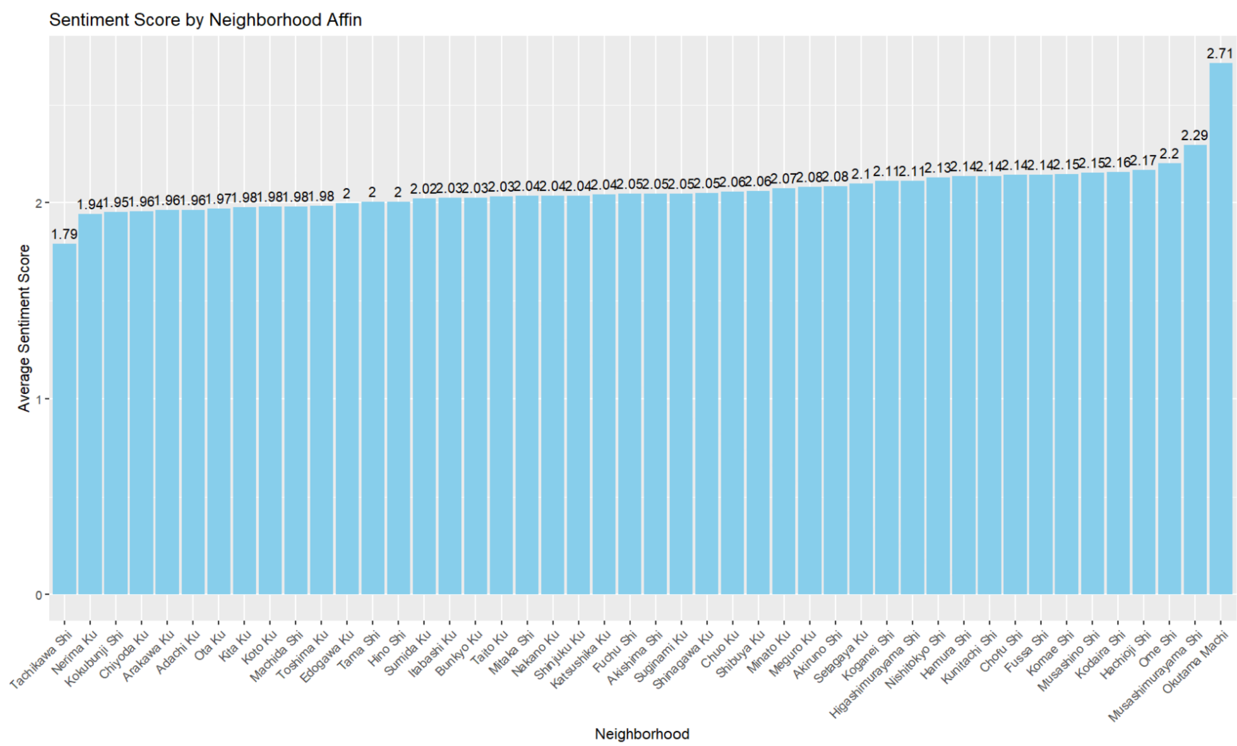

Figure 6: Affin average sentiment score per Neighborhood.

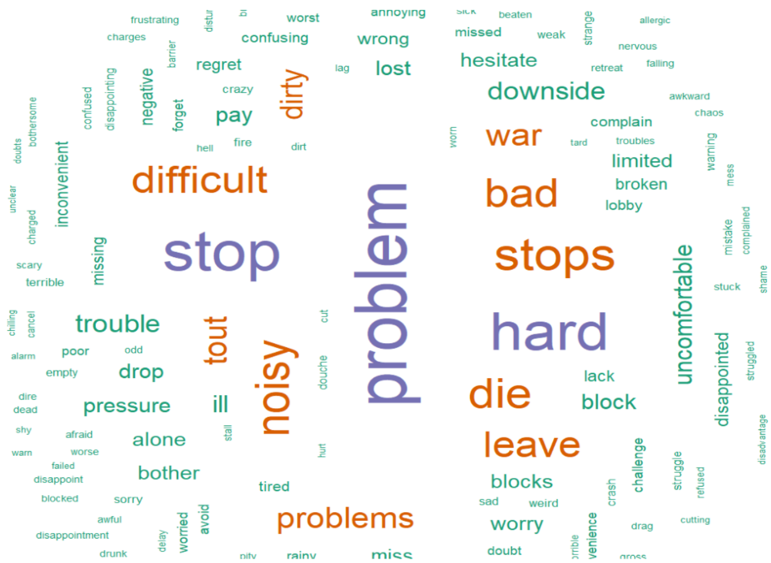

Figure 7: The frequency of negative words in comments.

Figure 8: The frequency of positive words in comments.

4.4 The Document Frequency Matrix (DFM)

The Document Frequency Matrix (DFM) stores the frequency of each word in each comment, which allows us to compare all tokens across comments. Due to memory constraints, it was applied to a random sample size of 100 from the dataset, as shown in (Table 2). In this matrix, each row represents a document (a listing), while each column describes a term or word used in the listings' comments.

However, the sparsity of the DFM was very high (96%), as shown in (Figure 9). This is determined by the number of comments in the corpus and the number of words in each comment. Training the prediction model with a machine-learning algorithm would be extremely difficult and time-consuming.

To address this issue, a new vector space model was introduced using a Latent Semantic Analysis (LSA) technique. LSA extracts the relationship between comments and terms. The assumption is that terms with similar meanings will appear in similar text passages. LSA utilizes a Singular Value Decomposition (SVD) factorization of the term-document matrix to extract these relationships, as shown in the following equation:

$$SVD = V \sum W^T$$

Where:

V contains the eigenvectors of the term correlations XX^T.

W contains the eigenvectors of the comment correlations X^TX.

Machine learning algorithm is used to create prediction models after adding a lexicon classification-based label feature and splitting the comments into 70%-30% stratified splits (Table 4).

	document	term	count
1	10001115	advantage	1
2	10158167	advantage	1
3	10322230	advantage	1
4	10366183	advantage	1
5	10459562	advantage	1
6	10534006	advantage	1
7	10582789	advantage	1
8	10600978	advantage	1
9	11167120	advantage	2
10	11219498	advantage	1
11	11575556	advantage	1
12	11589890	advantage	1
13	11679419	advantage	2
14	11981272	advantage	1
15	12009565	advantage	1
16	12071858	advantage	1
17	12114845	advantage	1
18	12117412	advantage	2
19	12137847	advantage	1
20	12257424	advantage	1

Table 2: Extract sample from the Document frequency matrix (DFM)

```
<<DocumentTermMatrix (documents: 95, terms: 140)>>
Non-/sparse entries: 542/12758
Sparsity           : 96%
Maximal term length: 12
Weighting          : term frequency (tf)
```

Figure 9: The DFM results

	comments	affin_score	sentiment
1	had a great nights stay with my wife and yr old kid clean an...	3	positive
2	such a lovely location it felt residential and not touristy at all...	2	positive
3	mitsukosan was incredibly welcoming accommodating and ...	-1	negative
4	we are travelers from america and have had a great stay we ...	2	positive
5	if youre looking for a place without the bother of communic...	3	positive
6	as im business man i would say that im having quite alot of ...	1	positive
7	brit was a side street but quiet and the alleys were beautiful...	-1	negative
8	this place was lovely there were big umbrellas phone charge...	2	positive
9	apartment is in a very convenient spot but tucked away off t...	1	positive
10	very convenient location less than a minute from metro stati...	2	positive
11	located at peaceful neighborhood near to the station with w...	2	positive
12	masami place was really amazing the amenities were really c...	2	positive
13	perfect location near the subway shops and restaurants des...	2	positive
14	a good hotel at a central location many attractions nearby ...	3	positive
15	taku is very hospitable and prompt with his responses he ev...	3	positive
16	han is an amazing host always here at anytime when we nee...	2	positive

Table 3: Sample of comments sentiment classification

4.5 Prediction Model Analysis

To predict comment sentiment on a sample of 100 from the dataset, a Decision Tree/CART model was applied. Decision trees are supervised learning models used for both classification and regression tasks. Classification and Regression Trees (CART) is a decision tree algorithm used for both classification and regression tasks. Recursively splitting the data into subsets according to the values of input variables is how it operates. At every step, the algorithm optimizes a selected criterion, like Gini impurity, to determine which variable best separates the data into homogeneous groups. In this analysis, Classification and Regression Trees (CART) were built to predict whether a given variable (comment) was positive or negative. After that, the tree would use patterns found in the text data to assign new comments to a sentiment category.

One method for evaluating a model's generalization performance is cross-validation. The process entails dividing the dataset into k folds, training the model on k-1 folds, and assessing its performance on the remaining fold. Every fold serves as the test set exactly once during the k repetitions of this process. Five-fold cross-validation was used to achieve the reported accuracy of 93.33% and a cp value of 0. In CART, a cp (complexity parameter) signifies the cost of adding another variable split to the tree. A cp value of 0 indicates that the model does not penalize the addition of new splits, possibly leading to a complex tree structure. Besides, the 93.33% score indicates that the model successfully categorized the majority of test data sentiments as either positive or negative. However, it's important to acknowledge that the model's performance may differ when applied to unseen data. The model's overall prediction accuracy across all classes is measured by the confusion matrix accuracy. The confusion matrix reveals that while the model correctly predicted positive sentiments, it failed to accurately classify any negative sentiments (see Figure 10). The confusion matrix accuracy is defined as follows:

$$\text{Accuracy} = \frac{TP+TN}{TP+FP+FN+TN} = \frac{a+d}{a+b+c+d}$$

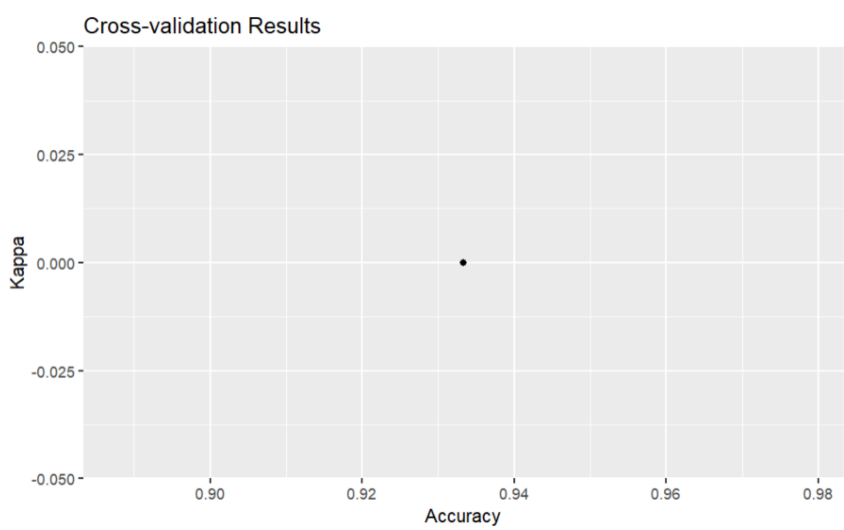

Figure 10: Cross- validation

```
predictions positive
   negative        0
   positive        3
```

Figure 11: Confusion matrix

4.6 Dashboard- Visualization

Data visualization uses charts, graphs, and maps to show information. These tools are vital for understanding Big Data. But visuals must be well-crafted to convey the message effectively. Among data visualization software, Tableau stands out (Patel, 2021). Therefore, after applying the machine learning model an interactive and user-friendly visualization dashboard was created using Tableau to present the sentiment analysis results, giving hosts and decision makers intuitive insights into the prevalent sentiments surrounding Airbnb accommodations in Tokyo (Figure 12).

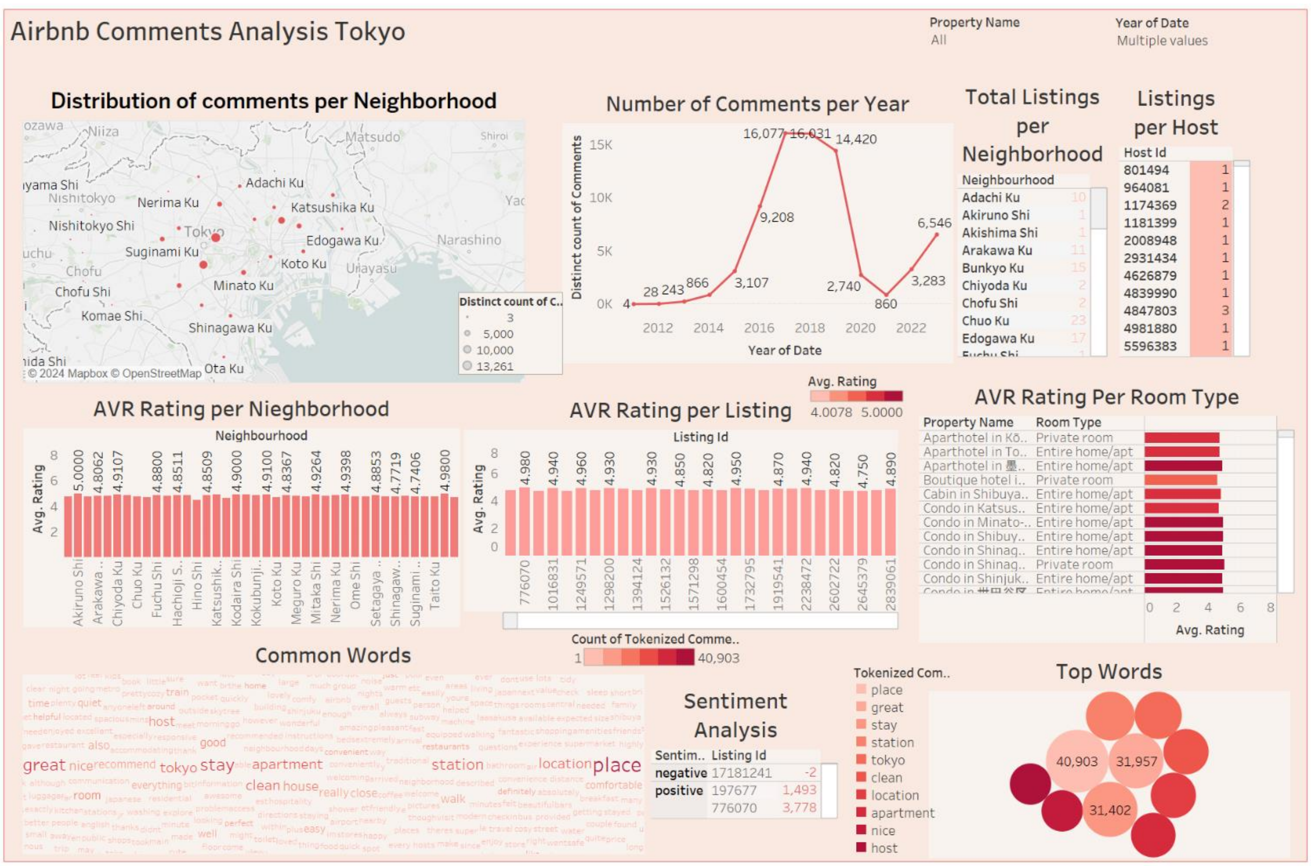

Figure 12: Comments Analysis Dashboard

Chapter 5 Results

Using the Affin sentiment lexicon, positive emotions and sentiments were found to outweigh negative ones across all listings. The sample dataset was divided into sub-samples using the k-fold cross-validation method. This method randomly divides the original sample into k equal-sized sub-samples and repeats the cross-validation procedure k times. Each time, one of the k sub-samples is used as the test set, while the rest are combined for training. The average error for all k trials is calculated. A 5-fold cross-validation was implemented on the Train dataset, followed by the CART Decision Tree Classification Algorithm. To address efficiency reduction, the Vector Space Model (Latent Semantic Analysis) was applied to the dataset for the CART.

Using 5-fold cross-validation, the CART model had an accuracy of 0.93 and a cp value of 0. This demonstrates that the model correctly classified the majority of test data sentiments as positive or negative. However, it is important to note that the model's performance may vary when applied to unseen data. The confusion matrix reveals that while the model correctly predicted positive sentiments, it failed to accurately classify any negative sentiments (see Figure 11).

Chapter 6 Conclusion

6.1 Conclusion

In this study, we used sentiment analysis techniques to extract insights from guest reviews of Tokyo Airbnb listings between September 2011 and June 2023. Using the Affin sentiment lexicon, we discovered that positive emotions and sentiments dominated all listings, indicating overall guest satisfaction. Using a 5-fold cross-validation method, we trained a CART model on a sample of the data set to classify sentiments with an accuracy of 0.93. However, the model had difficulty accurately classifying negative sentiments. Despite this limitation, the findings demonstrate the potential of sentiment analysis in enabling Airbnb hosts and policymakers to make informed decisions.

6.2 Recommendations and Future Work

Sentiment analysis of guest reviews can help hosts identify areas for improvement and tailor their services to better suit guest preferences. Policymakers can use sentiment analysis to gain insights into community experiences and make informed decisions to improve tourism experiences. Collaboration among hosts, policymakers, and data analysts can help develop data-driven strategies to improve the overall tourist experience.

Future work will focus on expanding the scope of research to other cities and sentiment analysis to incorporate non-English comments and integrate sentiment analysis with demand forecasting models for more accurate predictions of tourist arrivals. This will involve developing algorithms and natural language processing techniques for effectively analyzing non-English comments. Moreover, future research will involve creating predictive models that utilize sentiment analysis insights from guest comments to forecast demand for holiday homes and homestays. These models will undergo rigorous validation and evaluation to ensure their accuracy and effectiveness, with ongoing refinement and optimization based on performance metrics.

References

1. Agrusa, J. (2020). The Airbnb effect: how tourist destinations are addressing short-term vacation rentals. *International Journal of Qualitative Research in Services*, *4*(1), 20-29.
2. Alamanda, D. T., Ramdhani, A., Kania, I., Susilawati, W., & Hadi, E. S. (2019). Sentiment analysis using text mining of Indonesia tourism reviews via social media. *Int. J. Humanit. Arts Soc. Sci*, *5*(2), 72-82.
3. Alonso-Almeida, M. D. M., Borrajo-Millán, F., & Yi, L. (2019). Are social media data pushing overtourism? The case of Barcelona and Chinese tourists. *Sustainability*, *11*(12), 3356.
4. Bagherzadeh, S., Shokouhyar, S., Jahani, H., & Sigala, M. (2021). A generalizable sentiment analysis method for creating a hotel dictionary: using big data on TripAdvisor hotel reviews. *Journal of Hospitality and Tourism Technology*, *12*(2), 210-238.
5. Bandara, H. M. A. P., Charles, J., & Lekamge, L. S. (2022). Using Sentiment Analysis to Explore the Accommodation Experience in the Sharing Economy through Topic Modeling. In *2022 4th International Conference on Advancements in Computing (ICAC)* (pp. 156-161). IEEE.
6. Berezina, K., Bilgihan, A., Cobanoglu, C., & Okumus, F. (2016). Understanding satisfied and dissatisfied hotel customers: text mining of online hotel reviews. *Journal of Hospitality Marketing & Management*, *25*(1), 1-24.
7. Bertrand, K. Z., Bialik, M., Virdee, K., Gros, A., & Bar-Yam, Y. (2013). Sentiment in new york city: A high resolution spatial and temporal view. *arXiv preprint arXiv:1308.5010*.
8. Brochado, A., Troilo, M., & Shah, A. (2017). Airbnb customer experience: Evidence of convergence across three countries. Annals of Tourism Research, 63, 210-212. https://doi.org/10.1016/j.annals.2017.01.001
9. Colladon, A. F., Guardabascio, B., & Innarella, R. (2019). Using social network and semantic analysis to analyze online travel forums and forecast tourism demand. *Decision Support Systems*, *123*, 113075.
10. De Marchi, D., Becarelli, R., & Di Sarli, L. (2022). Tourism Sustainability Index: Measuring Tourism Sustainability Based on the ETIS Toolkit, by Exploring Tourist Satisfaction via Sentiment Analysis. *Sustainability*, *14*(13), 8049.
11. Ding, K., Choo, W. C., Ng, K. Y., & Ng, S. I. (2020). Employing structural topic modelling to explore perceived service quality attributes in Airbnb accommodation. *International Journal of Hospitality Management*, *91*, 102676.

12. Güçlü, B., Roche, D., & Marimon, F. (2020). City characteristics that attract Airbnb travellers: evidence from Europe. *International Journal for Quality Research*.
13. He, W., Tian, X., Tao, R., Zhang, W., Yan, G., & Akula, V. (2017). Application of social media analytics: A case of analyzing online hotel reviews. *Online Information Review*, *41*(7), 921-935.
14. Hu, M., Li, H., Song, H., Li, X., & Law, R. (2022). Tourism demand forecasting using tourist-generated online review data. *Tourism Management*, *90*, 104490.
15. Kirilenko, A. P., Stepchenkova, S. O., Kim, H., & Li, X. (2018). Automated sentiment analysis in tourism: Comparison of approaches. *Journal of Travel Research*, *57*(8), 1012-1025.
16. Lee, C. K. H., & Tse, Y. K. (2021). Improving peer-to-peer accommodation service based on text analytics. *Industrial Management & Data Systems*, *121*(2), 209-227.
17. Li, Z., Chen, H., & Huang, X. (2020). Airbnb or hotel?: A comparative study on the sentiment of airbnb guests in sydney – text analysis based on big data. International Journal of Tourism and Hospitality Management in the Digital Age, 4(2), 1-10. doi:https://doi.org/10.4018/IJTHMDA.2020070101
18. Madhi, H. A. B., & Alhammad, M. M. (2021). What Drives Airbnb Customers' Satisfaction in Amsterdam? A Sentiment Analysis. *International Journal of Advanced Computer Science and Applications*, *12*(6).
19. Mavroforaki, D. M., & Kolovou, A. (2020). Sentiment Analysis in Tourism: Athens's vibe through online platforms. Retrieved from https://pergamos.lib.uoa.gr/uoa/dl/object/2898493/file.pdf
20. Mohammad, S. M., Kiritchenko, S., & Zhu, X. (2013). NRC-Canada: Building the state-of-the-art in sentiment analysis of tweets. *arXiv preprint arXiv:1308.6242*.
21. Nazirkar, S., & Kulkarni, S. (2023). Sentiment Analysis and Customer Satisfaction Factors Based on LSTM and Topic Modeling. *Indian Journal of Science and Technology*, *16*(28), 2126-2132.
22. Park, Y., Kim, M., & Seong, K. (2021). Happy neighborhoods: Investigating neighborhood conditions and sentiments of a shrinking city with Twitter data. *Growth and Change*, *52*(1), 539-566.
23. Raza, M. R., Hussain, W., & Varol, A. (2022). Performance analysis of deep approaches on airbnb sentiment reviews. In *2022 10th International symposium on digital forensics and security (ISDFS)* (pp. 1-5). IEEE.

24. Santos, A. I. G. P., Perinotto, A. R. C., Soares, J. R. R., Mondo, T. S., & Cembranel, P. (2022). Expressing the Experience: An Analysis of Airbnb Customer Sentiments. *Tourism and Hospitality*, *3*(3), 685-705.
25. Schmunk, S., Höpken, W., Fuchs, M., & Lexhagen, M. (2013). Sentiment analysis: Extracting decision-relevant knowledge from UGC. In *Information and Communication Technologies in Tourism 2014: Proceedings of the International Conference in Dublin, Ireland, January 21-24, 2014* (pp. 253-265). Springer International Publishing.
26. Stylos, N., Zwiegelaar, J., & Buhalis, D. (2021). Big data empowered agility for dynamic, volatile, and time-sensitive service industries: the case of tourism sector. *International Journal of Contemporary Hospitality Management*, *33*(3), 1015-1036.
27. Taecharungroj, V., & Mathayomchan, B. (2019). Analysing TripAdvisor reviews of tourist attractions in Phuket, Thailand. *Tourism Management*, *75*, 550-568.
28. Xiang, Z., Schwartz, Z., Gerdes Jr, J. H., & Uysal, M. (2015). What can big data and text analytics tell us about hotel guest experience and satisfaction?. *International journal of hospitality management*, *44*, 120-130.
29. Yaraghi, N., & Ravi, S. (2017). The current and future state of the sharing economy. *Available at SSRN 3041207*.https://harryphillipsaic.com/wp-content/uploads/2020/08/sharingeconomy_032017final.pdf
30. Ye, Q., Zhang, Z., & Law, R. (2009). Sentiment classification of online reviews to travel destinations by supervised machine learning approaches. *Expert systems with applications*, *36*(3), 6527-6535.
31. Patel, A. (2021). Data Visualization Using Tableau.