Journal of Indigenous Culture, Tourism, and Language (JICTL)
Vol 1, No 1, 2025, 61-73 https://doi.org/10.35912/jictl.v1i1.3256

Big data analytics to understand guest sentiment: Time series study of TripAdvisor reviews for luxury hotel in Bali

Ayu Trisna Sanjiwani[1], Putu Diah Sastri Pitanatri[2], Cahyo Purnomo Loanata[3]
Bali Tourism Polytechnic, Bali, Indonesia[1,2,3]
ayutrisnasanjiwani8@gmail.com[1], diahsastri@ppb.ac.id[2], cahyoloanata@yahoo.co.id[3]

Article History
Received on 04 January 2025
1st Revision on 09 February 2025
Accepted on 13 March 2025

Abstract
Purpose: This study analyzes guest sentiment from TripAdvisor reviews and examines its relationship with room occupancy at the Raffles Bali Hotel. It explores key factors influencing guest satisfaction and their correlation with occupancy trends, using a time-series forecasting approach to predict future hotel performance.
Research/methodology: The research utilized TripAdvisor review data from 2020–2024, which was scraped, cleaned, and classified for sentiment using Python and Julius AI. A Seasonal Autoregressive Integrated Moving Average model was applied to sentiment data, while a SARIMAX model incorporated sentiment as an exogenous variable to forecast occupancy rates.
Results: Findings indicate that most guest reviews were positive, contributing to high overall satisfaction levels. Although occasional declines in sentiment occurred (e.g., March 2024), trends remained favorable. Time-series analysis revealed a significant positive influence of sentiment on occupancy, with slight negative short-term fluctuations. This suggests that while sentiment strongly supports long-term occupancy growth, short-term variations are less predictable.
Conclusions: Positive guest sentiment is a key driver of occupancy rates in luxury hotels. Although Granger Causality testing did not confirm a short-term causal link, long-term trends highlight the importance of managing guest sentiment to sustain occupancy levels. Hotel managers can use these insights to optimize service quality, improve guest experiences, and refine marketing strategies.
Limitations: The SARIMA models excluded external factors such as marketing campaigns, seasonal events, and competitor data. Guest demographics were not segmented in this study.
Contribution: This study introduces a novel integration of sentiment analysis and time-series forecasting, providing actionable insights to enhance service quality and improve hotel occupancy performance.

Keywords*: Big Data Analytics, Guest Satisfaction, Sentiment Analysis, TripAdvisor Reviews*

How to Cite: Sanjiwani, A. T., & Pitanatri, P. D. S., & Loanata, C. P. (2025). Big data analytics to understand guest sentiment: A time series study of Tripadvisor reviews for luxury hotels in Bali. *Journal of Indigenous Culture, Tourism, and Language,* 1(1), 61-73.

1. Introduction

Tourism plays a vital role in Bali's economy and is one of the region's main economic drivers. The island relies heavily on tourism as its primary source of income, and this sector is one of Indonesia's largest contributors to foreign-exchange earnings. A significant portion of the Balinese population depends on tourism for their livelihoods (Chin, 2017). Bali's tourism sector has substantial potential for further development, making it a valuable source of regional income (Wiranatha, Suryawardani, Purbanto, Yudiastina, & Bantacut, 2024). In addition to its various tourist destinations, Bali is home to many star-rated hotels that offer complete facilities and exceptional services to visitors. The attraction of Bali for both domestic and international tourists lies in its rich cultural heritage, complemented by natural landscapes such as beaches and well-established infrastructure that supports the tourism industry. Among the most influential elements are hotels, resorts, and villas, ranging from non-starred accommodations to five-star standards (Artini, Antara, Susrusa, & Ambarawati, 2020).

In recent years, luxury tourism has significantly transformed Bali's hospitality landscape. The island now boasts a wide array of luxury hotels and resorts that cater to high-end travellers seeking exclusivity, premium services, and immersive experiences. With the increasing number of luxury hotels in Bali, competition among premium properties has intensified. Each establishment strives to differentiate itself through unique branding, exceptional service quality, and personalized guest experience. Given the saturated market, luxury hotels must maintain high service standards and continually evaluate and improve their performance through guest feedback and data-driven strategies (Williady, Wardhani, & Kim, 2022).

Raffles Bali is a five-star hotel in Bali and is part of the Accor hotel group. Raffles Bali was selected as the focus of this study not only because of its distinguished position within Bali's luxury hospitality market but also because of its commitment to premium guest experiences and its active online presence, particularly on platforms such as TripAdvisor (Phillips, Barnes, Zigan, & Schegg, 2017). This combination makes it a compelling case to explore how guest sentiment correlates with room occupancy performance in a luxury setting. The room occupancy rate at Raffles Bali has shown a significant increase over five years, from 29.00% in 2020 to 54.00% in 2024. However, the growth between 2023 and 2024 is only 0.80%, indicating a slowdown in occupancy growth compared to previous years (Chang, Chen, Lai, Lin, & Pai, 2021). This trend has become a concern for management, highlighting the need for more effective strategies to improve room occupancy consistently. The room occupancy rate is a critical performance indicator in the hospitality industry, reflecting the proportion of rooms occupied relative to those available (Kim & Han, 2023). An in-depth analysis of room occupancy trends from 2020 to 2024 confirmed a steady increase; however, the slowdown between 2023 and 2024 suggests a plateau in growth. Room occupancy directly contributes to hotel revenue because rooms are the primary product of hospitality. Higher occupancy leads to higher profit margins. According to Mattila and O'Neill (2003), there is a connection between room pricing, occupancy, and guest satisfaction. They noted that higher prices do not always equate to higher satisfaction, emphasizing the importance of balancing pricing with service quality.

To address this challenge, it is crucial to design effective strategies based on a deeper understanding of the guest feedback. Guest reviews, especially on platforms such as TripAdvisor, are valuable tools for evaluating service quality and informing improvements (Toker, 2024). By analyzing these reviews, hotels can gain insights into guest preferences and expectations, ultimately helping them formulate more targeted service enhancements. Some negative reviews of Raffles Bali highlight issues such as delayed welcomes, tension in staff demeanors, inconsistency in butler services, and lack of proactive responses from management (Zhang, Xu, Gou, & Chen, 2021). Although certain staff members are praised for their dedication, several operational shortcomings have negatively impacted guest experiences. One review reported significant disturbances caused by ongoing construction, despite prior requests for a quiet villa location. These experiences reflect service quality gaps and a mismatch between brand expectations, particularly for luxury hotels, and actual guest experiences (Peterhans, 2010).

As more users express their opinions via Web 2.0, online review platforms such as TripAdvisor have gained considerable influence. These platforms provide a space for travelers to reflect on and discuss their experiences, contributing to the wider dissemination of electronic word-of-mouth (e-WoM) (X. Liu, Lin, Jiang, Chang, & Lin, 2024; Mokgehle & Fitchett, 2024). Despite the growing academic interest in TripAdvisor data, many studies still rely on manual data collection, which is labor-intensive and prone to human errors. Automated methods, such as web scraping, offer a more scalable and accurate alternative for sentiment analysis (Barbera, Araujo, & Fernandes, 2023). Although Raffles Bali offers luxurious experiences and impressive services, there are areas for improvement. The minimal occupancy growth between 2023 and 2024 indicates the need to leverage guest reviews, particularly TripAdvisor, as a tool for service optimization. Based on this background, the current study aims to apply time-series analysis to examine guest sentiment in TripAdvisor reviews and its effect on room occupancy in Raffles Bali (Sanjiwani, Pitanatri, & Loanata, 2025).

This study focuses on analyzing guest sentiments based on TripAdvisor reviews and their correlation with room occupancy at Raffles Bali using a time-series approach. It explores customer perceptions of the services received and whether they align with brand expectations (Sanjiwani et al., 2025). The data used comprise reviews from guests who have stayed at the hotel and are analyzed to identify sentiment patterns and their effects on hotel occupancy. The goal of this study was to provide strategic recommendations for improving room occupancy. This study aims to identify guest sentiment and factors influencing room occupancy derived from a time series analysis of TripAdvisor reviews at Raffles Bali (Sahadev, Seiler, & Scarf, 2025).

2. Literature review

2.1 Guest Sentiment in TripAdvisor Reviews

Customer sentiment refers to the expression of opinions or emotions conveyed through online reviews. According to (W. Liu, Liang, Bao, Qin, & Lim, 2022), sentiment analysis categorizes reviews as follows.

1. Positive Sentiment: Reviews with positive connotations like "excellent," "comfortable," and "friendly service."
2. Negative Sentiment: Reviews containing negative words such as "unsatisfactory," "slow service," and "office-like rooms."
3. Neutral Sentiment: Reviews that do not show strong emotions.

2.2 Time Series Analysis

A time series is a sequence of data points measured at uniform intervals (Das & Barman, 2025). Analysis involves understanding and forecasting patterns, such as horizontal, trend, seasonal, and cyclical patterns.

2.3 Customer Reviews

According to Iduozee (Chen, Samaranayake, Cen, Qi, & Lan, 2022), customer reviews inform buyers about products and evaluate their post-purchase experiences. Kamila et al. Melati and (Zhao, Wang, Guo, & Law, 2015) list six criteria for effective online consumer reviews: usefulness, reviewer expertise, timeliness, volume, valence, and comprehensiveness.

2.4 TripAdvisor

TripAdvisor is the world's largest travel website for planning and booking trips, featuring reviews and opinions of millions of travelers. Founded in 2000, it started as a B2B platform but has evolved into a community-driven site. It uses a bubble rating system (1-5) to rate properties based on the quality, quantity, and recency of reviews (Alaimo, Kallinikos, & Valderrama-Venegas, 2020).

2.5 Customer Satisfaction

Customer satisfaction is the feeling of pleasure or disappointment resulting from a comparison of perceived performance with expectations (Aimee, 2019). Indrasari (Khoironi, Syah, & Dongoran, 2018) identified five main factors affecting customer satisfaction: product quality, service quality, emotional appeal, price, and cost.

2.6 Customer Satisfaction Indicators

Indicators for measuring customer satisfaction include expectation alignment, willingness to return, and the likelihood of recommendation (Schiebler, Lee, & Brodbeck, 2025)

2.7 Occupancy Rate

The occupancy rate in hospitality refers to the percentage of occupied rooms compared to the total available rooms over a specific period of time. It is a crucial performance indicator influenced by factors such as seasonality, external events, service quality, pricing strategies and reputation.

3. Research methodology

This study focuses on analyzing guest reviews of Raffles Bali using a time-series approach to understand their influence on room occupancy rates from 2020 to 2024. This study was conducted at Raffles Bali. The research was conducted over two months, from December 2024 to January 2025. This timeframe was chosen to allow comprehensive data collection and in-depth analysis, providing the necessary flexibility to produce reliable and significant findings that contribute meaningfully to the field. The data used in this study comprised secondary sources, specifically guest reviews from the TripAdvisor platform, which were cleaned and processed to become valid and relevant for the primary analysis. This "absolute population" included all review data relevant to the research scope.

The operational variables included TripAdvisor reviews, room occupancy rate, time period, sentiment classification, and observed trends over time. These variables were measured using indicators such as the number of reviews per period, review ratings, sentiment polarity (positive, negative, neutral), occupancy percentages, and fluctuations over time. To collect data, the study employed scraping methods and API-based access, particularly Python programming, because of its powerful libraries for automation and data extraction. TripAdvisor review data were automatically retrieved and structured for analysis with the help of Google Colab to run scripts and manage data over a five-year span, which were then exported into Excel format.

The next step involved data cleaning, which is a crucial process for ensuring the accuracy and consistency of the dataset. This process included identifying and removing errors, duplicate entries, and irrelevant information. Cleaning was conducted using Julius AI, an AI platform specializing in natural language processing (NLP) that supports sentiment classification, topic identification, and deeper insights into customer feedback. Before the main analysis, preprocessing steps were performed to prepare the data for analysis. This included case folding (converting all text to lowercase and removing unnecessary symbols), tokenizing (splitting text into word tokens and converting them to their base forms), filtering (removing unimportant words using Indonesian stop words), and stemming (reducing words to their root forms to standardize the vocabulary used in the analysis). These pre-processed data were then ready for core data analysis, which aimed to uncover trends in sentiment and occupancy using a time series method, setting the foundation for predictive modeling and strategic recommendations for improving hotel performance based on customer feedback.

This study employs a time series method using the Seasonal Autoregressive Integrated Moving Average approach, which is a forecasting method for data with seasonal patterns. SARIMA is an extension of the Autoregressive Integrated Moving Average (ARIMA) model. The time series data used in this study consisted of review data from 2020 to 2025.

4. Result and discussion

4.1 Sentiment Analysis Result

Table 1. Sentiment Analysis

	Sentiments	Total
0	Positive	218
1	Negative	5
2	Neutral	2

Source: reach result, 2025

Table 1 presents the sentiment classification results obtained from the guest reviews of the Raffles Bali Hotel sourced from the TripAdvisor platform. Of the total reviews analyzed, 218 were categorized as positive sentiments, indicating that the majority of guests expressed satisfaction with their stay and the services provided. Meanwhile, five reviews were classified as negative, reflecting dissatisfaction with certain aspects of the hotel's services or facilities. Two reviews were categorized as neutral, meaning that they did not clearly express positive or negative emotions. The dominance of positive sentiment suggests that, overall, the public perception of Raffles Bali is quite favorable, although negative reviews should still be taken seriously as feedback to improve the overall service quality.

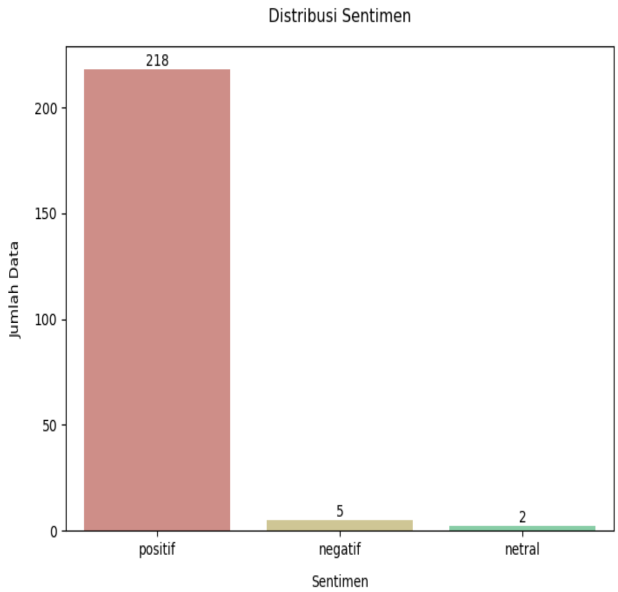

Figure 1. Sentiment Distribution
Source: Reach result, 2025

Monthly sentiment aggregation was conducted to observe patterns in guest perceptions over the course of this study. The analysis showed that positive sentiment dominated almost every month, with average sentiment scores close to the maximum value (1.0) in most months from 2024 to 2025. The number of reviews also fluctuated, peaking in November 2024 with 21 reviews. However, there were months with very few reviews, ranging from one to three reviews. Interestingly, in March 2024, the average sentiment score dropped significantly to 0.5, suggesting an increase in negative reviews and dissatisfaction. However, the sentiment scores improved in the following months. This pattern provides valuable insights for hotel management to identify periods of declining satisfaction and evaluate internal factors that may have contributed, such as service quality or environmental conditions.

4.2 Time Series Sentiment and Occupancy Trends

Monthly sentiment aggregation results from guest reviews collected between July 2024 and March 2025 are presented. The findings show that the average sentiment scores were generally high, with most months recording scores close to the maximum value of 1.0. This indicates that most guests left positive reviews of their stay. November 2024 recorded the highest number of reviews, totaling 21, although the sentiment score for that month decreased slightly to 0.90, suggesting the presence of neutral or negative feedback. This dip in sentiment may be attributed to factors such as reduced service quality, slow staff responses, or environmental disruptions such as construction, as mentioned in the review. In contrast, from December 2024 to March 2025, sentiment scores returned to a perfect 1.0, suggesting that management likely made improvements in service delivery, thus enhancing the guest experience. This trend is a crucial indicator for management to maintain high service standards and respond promptly to online feedback.

Based on the monthly aggregation analysis, it can be concluded that the average sentiment score for Raffles Bali guests shows a positive trend, with most months scoring very high, close to, or even at the maximum value of 1.0. However, there was a significant drop in sentiment scores between March and May 2024, with the lowest score in March 2024, when the score dropped to 0.5. This indicates an increase in negative reviews or guest dissatisfaction during this period. The number of reviews varied considerably each month. November 2024 had the highest number of reviews (21), whereas several other months had only one to three reviews, indicating low guest interaction with TripAdvisor during those times. Looking at the overall trend, the number of reviews increased consistently toward the end of 2024 and into early 2025, coinciding with the rebound in sentiment scores after the dip in the year.

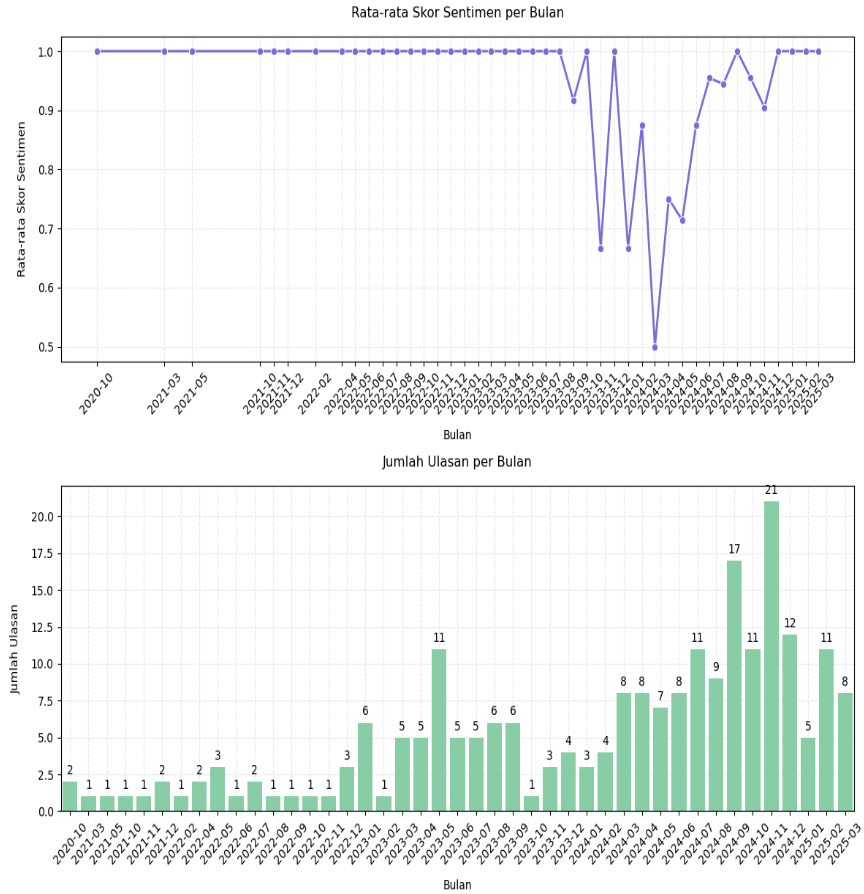

Figure 2. The Monthly Aggregation Analysis
Source: reach result, 2025

Based on the results of the monthly aggregation analysis, it can be concluded that the average sentiment score of guests at the Raffles Bali Hotel showed a positive trend, with most months recording very high scores approaching or even reaching the maximum value of 1.0. However, there was a significant decline in sentiment scores from March to May 2024, with the lowest point occurring in March 2024, when the score dropped to 0.5. This indicates an increase in negative reviews and guest dissatisfaction during this period. In terms of the number of reviews, there was a noticeable variation from month to month. November 2024 recorded the highest number of reviews, totaling 21, whereas several other months had only one to three reviews, suggesting low guest engagement with the TripAdvisor platform during those periods. Overall, it is evident that the number of reviews increased consistently in the final months of 2024 through early 2025, in line with the recovery of sentiment scores to higher levels following the dip from early to mid-2024.

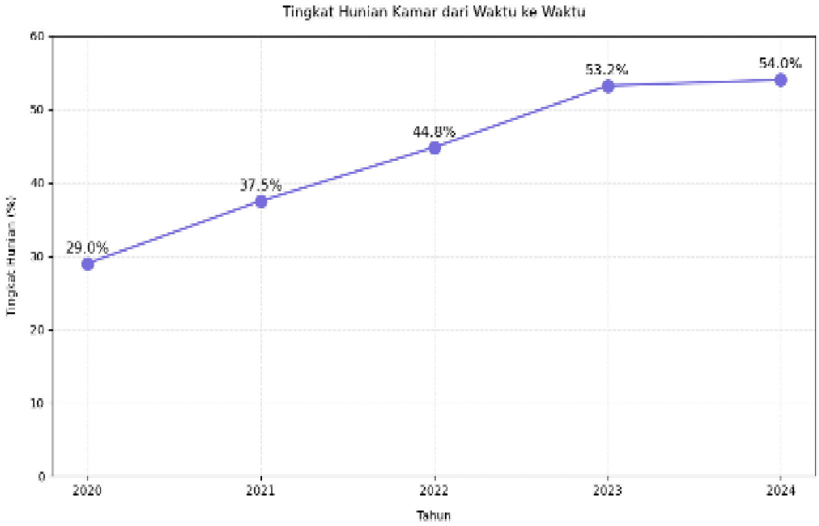

Figure 3. Room Occupancy Rate
Source: Reach result, 2025

Based on the graph of room occupancy rates at Raffles Bali from 2020 to 2024, a consistent upward trend was observed every year. In 2020, the occupancy rate was 29.0%, increasing to 37.5% in 2021, 44.8% in 2022, 53.2% in 2023, and reaching 54.0% in 2024. In connection with the sentiment analysis of guest reviews on TripAdvisor and the previously discussed time-series decomposition chart, this upward trend in room occupancy appears to be closely related to the positive perceptions that guests have of their stay. Most sentiment scores remain high (close to 1.0), particularly after the mid-2024. This suggests that most reviews were positive, which likely played a significant role in building the hotel's reputation and enhancing public trust in the quality of its service. However, there was a noticeable decline in sentiment scores from early to mid-2024, as shown in the sentiment trend chart, although it did not significantly affect overall occupancy levels. This indicates that despite the challenges in maintaining consistent customer satisfaction, Raffles Bali was able to sustain and slightly improve its occupancy rate. It is likely that the hotel responded quickly to service improvements, preventing a significant drop in occupancy.

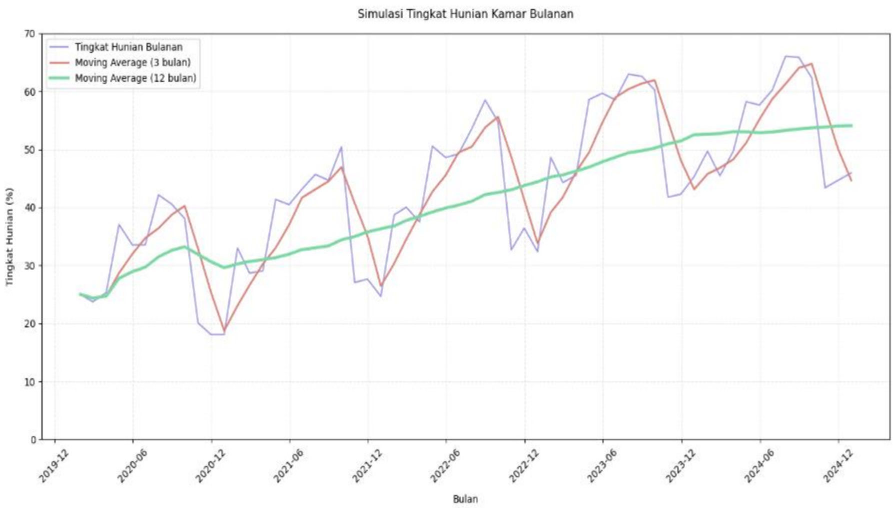

Figure 4. Monthly Occupancy
Source: Reach result, 2025

The graph illustrates the development of the hotel's room occupancy rate from the end of 2019 to the end of 2024, with three lines representing the actual monthly data, a 3-month moving average, and a

12-month moving average. The data show consistent seasonal fluctuations each year, with occupancy rates typically increasing in the middle of the year and declining toward the end of the year. Nevertheless, the long-term trend, as depicted by the 12-month moving average, indicates a steady year-over-year increase. This suggests that hotel room occupancy performance has continuously improved. The rise in positive guest sentiments, as reflected in online reviews, has likely contributed to the increase in occupancy rates, which reflects improvements in service quality and customer satisfaction. In other words, despite seasonal patterns causing short-term fluctuations, the overall trend indicates healthy and sustainable growth in the industry.

The Pearson correlation coefficient between the average sentiment score and room occupancy rate was -0.1639, indicating a weak, negative correlation. This means that when guest sentiment increases, occupancy rates tend to decrease slightly and vice versa. However, this relationship was not strong enough to suggest a meaningful or consistent pattern. Subsequently, the Granger Causality Test was conducted to further explore the causal relationship between sentiment and occupancy over different lag periods. The results are presented in the next section in tabular format for three different lags.

4.3 Granger Causality and Forecasting

Table 2. Granger Causality test presented

Test	Statistic Value	p-value	Df
SSR based F-test	F = 0.9466	0.3371	df_num = 1, df_denom = 36
SSR based chi²-test	$\chi^2 = 1.0254$	0.3112	df = 1
Likelihood ratio test	$\chi^2 = 1.0122$	0.3144	df = 1
Parameter F-test	F = 0.9466	0.3371	df_num = 1, df_denom = 36
SSR based F-test	F = 0.9657	0.3912	df_num = 2, df_denom = 33
SSR based chi²-test	$\chi^2 = 2.2241$	0.3289	df = 2
Likelihood ratio test	$\chi^2 = 2.1614$	0.3393	df = 2
Parameter F-test	F = 0.9657	0.3912	df_num = 2, df_denom = 33
SSR based F-test	F = 0.8130	0.4968	df_num = 3, df_denom = 30
SSR based chi²-test	$\chi^2 = 3.0082$	0.3904	df = 3
Likelihood ratio test	$\chi^2 = 2.8921$	0.4086	df = 3
Parameter F-test	F = 0.8130	0.4968	df_num = 3, df_denom = 30

Source**:** reach result, 2025

Across all lags (lag 1, lag 2, and lag 3), the p-values from the four types of tests (SSR-based F-test, SSR-based chi²-test, likelihood ratio test, and parameter F-test) are all above the significance level of 0.05. At lag 1, the SSR-based F-test showed an F-value of 0.9466 with a p-value of 0.3371, while the SSR-based chi²-test and the likelihood ratio test yielded $\chi^2 = 1.0254$ ($p = 0.3112$) and $\chi^2 = 1.0122$ ($p = 0.3144$), respectively. Similar results were observed at lags 2 and 3, where the p-values remained above 0.05, indicating no statistically significant causal relationship. Therefore, it can be concluded that there is insufficient evidence to suggest a causal relationship between the variables under examination in the short term across all three periods. This implies that changes in one variable cannot be used to predict changes in the other variable.

4.3.1 Sentiment Forecast

Modeling was performed to forecast guest sentiment at the Raffles Bali Hotel for the next 12 months. The forecast results indicate that the predicted sentiment mean values range from 0.976 to 0.983, reflecting a consistently high and positive sentiment trend. These values suggest that guests' perceptions of services and overall experience at the Raffles Bali Hotel are expected to remain favorable over the next year. This forecast reflects the hotel's consistency in delivering quality services and its potential to maintain a positive image in the eyes of its guests. Assuming that both internal and external conditions remain stable, the forecast serves as a strong indication that future guest reviews will likely continue to fall within the positive range.

4.3.2 Occupancy Forecast
Historical occupancy data were integrated with an exogenous variable (sentiment) to model and forecast occupancy rates for the next 12 months using the LSTM model. The following table presents the historical data, which presents the historical monthly occupancy rates of the Raffles Bali Hotel along with the average sentiment scores from January to May 2020, with occupancy levels ranging from 23%–37%. Despite the relatively low occupancy rates during this period, sentiment scores consistently remained at the maximum value of 1.0 in both the actual and imputed values. This suggests that guests continued to provide highly positive reviews of their stay and service quality, indicating that service standards were well maintained from the outset. Forecast room occupancy rates for the next 12 months, covering the period from January to May 2025. The predicted values indicated a significant upward trend, with occupancy rates ranging from 53.22% to 64.87%. The highest prediction was observed in April 2025, reaching 64.87%, which could be associated with the holiday season or the success of the hotel's marketing strategies.

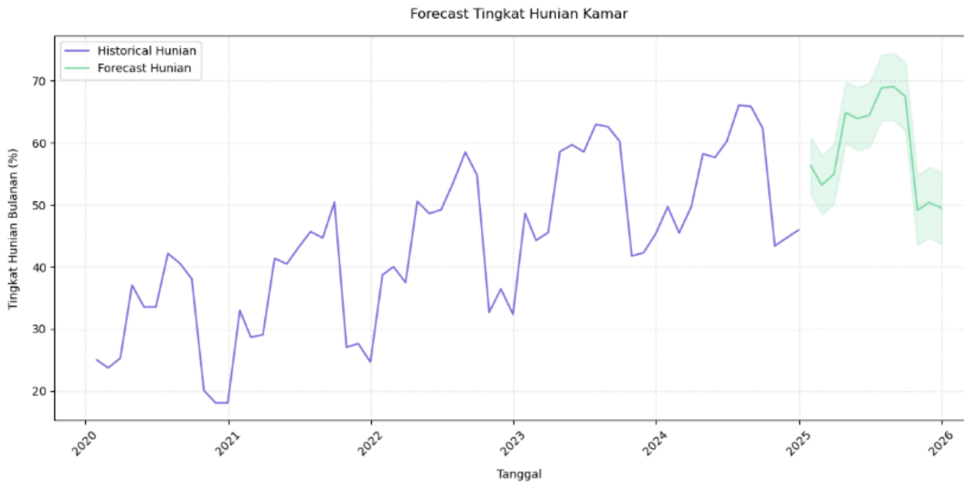

Figure 5. Occupancy Forecast
Source*:* reach result, 2025

The figure that visualizes this forecast also demonstrates the continuity between historical data and future estimates, where the trend graph shows a steady upward movement.
This suggests that the combination of positive sentiment reviews and the projected increase in room occupancy supports the assumption that favorable guest perceptions of the hotel may serve as a key factor in driving occupancy growth in the future.

4.4 Model Interpretation
Model Equation and Interpretation
Sentiment Model Equation (Hamid et al.) SARIMA Model for Sentiment: SARIMA(1,0,1)(1,0,1,12)
$y_t = 0.0000 + 0.9950 * y_{(t-1)} + -0.5847 * \varepsilon_{(t-1)} + 0.0598 * y_{(t-12)} + 14454845019.9674 * \varepsilon_{(t-12)} + \varepsilon_t$

4.4.1 Sentiment Model Interpretation (Hamid et al.)
The SARIMA(1,0,1)(1,0,1,12) model used in this study was designed to forecast guest sentiment over time while accounting for annual seasonality. In this model, a constant value of 0.0000 represents the baseline sentiment level when all other factors are set to zero. The AR(1) coefficient of 0.9950 indicates a strong dependency on the sentiment score from the previous month, suggesting that guest sentiments tend to persist over time. A high AR(1) value reflects the model's reliance on past sentiment trends to predict future sentiment. The MA(1) coefficient of -0.5847 shows the influence of the previous month's error term on the current sentiment score, highlighting the effects of past shocks or unpredictable changes. Meanwhile, the SAR(1) coefficient of 0.0598 indicates the seasonal impact of sentiment scores exactly 12 months prior, showing the presence of repeating annual patterns in guest sentiment.

Additionally, the SMA(1) coefficient of 14454845019.9674 points to the seasonal effect of the error term from 12 months ago, although its magnitude suggests a potential anomaly or a modeling artifact.

This predictive model complements the SARIMAX model used for forecasting room occupancy by considering annual seasonal components. In essence, the SARIMA model reveals that sentiment dynamics are influenced by immediate past values and follow seasonal patterns. Understanding these trends can help hotels anticipate changes in guest perceptions and tailor their strategies accordingly.

$$y_t = 0{,}000 + 0{,}9942 \text{ x }_{yt-1} + (-0{,}7563) \text{ x }_{et-1} + 0{,}9659 \text{ x }_{yt-12} + (-1{,}3146) \text{ x }_{et-12} + 6{,}4545 \text{ x}\chi_t + {}_{et}$$

The SARIMAX(1,0,1)(1,0,1,12) model used in this study incorporates sentiment scores as exogenous variables (x_t) to predict room occupancy rates. A constant value of 0.0000 represents the baseline occupancy level when all other factors are set to zero. An AR(1) coefficient of 0.9942 indicates a strong dependency on the previous month's occupancy rate, implying that the current occupancy is significantly influenced by past performance. Similarly, the MA(1) coefficient of -0.7563 reflects the impact of the error term from the previous month, suggesting that previous random shocks or unmodelled variations affect the current occupancy levels.

In terms of seasonality, the model includes a SAR(1) coefficient of 0.9659, showing a strong seasonal effect based on occupancy 12 months earlier. The SMA(1) coefficient of -1.3416 captures the seasonal impact of the error term from the same period, indicating that random shocks from a year ago played a role in shaping current trends. Notably, the sentiment coefficient of 6.4545 provides compelling evidence that guest sentiment has a direct and positive effect on room occupancy. A positive sentiment score increase is associated with higher occupancy levels, emphasizing the importance of customer perception and satisfaction in driving hotel performance and revenue.

These findings affirm that room occupancy at Raffles Bali is influenced by internal historical patterns and guest perceptions, as expressed in online reviews. Therefore, actively and strategically managing guest sentiments can serve as a valuable tool for enhancing operational outcomes. The models were evaluated using the AIC and BIC values. The SARIMAX model for occupancy achieved lower AIC (231.7657) and BIC (242.7375) values than the SARIMA sentiment model (AIC: 712.2155 and BIC: 719.0519). These results suggest that the occupancy model is more effective in explaining data variation and provides a strong basis for occupancy improvement strategies based on sentiment analysis

4.5 Discussion

The sentiment classification results from Raffles Bali guest reviews collected via TripAdvisor showed that the overall guest perception of the hotel was very positive. Of the 225 reviews analyzed, 218 were classified as positive, five as negative, and two as neutral. This finding suggests that the majority of guests were satisfied with the services, facilities, and overall experience at the Raffles Bali. The dominance of positive sentiments serves as strong evidence that the hotel has largely succeeded in meeting guests' expectations of it. However, it is important to note that sentiment scores dropped significantly during certain periods. One such drop occurred in March 2024 when the average sentiment score was only 0.5. This decline was associated with reviews expressing dissatisfaction with inconsistent butler services, noise from construction activities, and slow responses to complaints from hotel management. This highlights that, although positive reviews prevail, the presence of negative reviews at specific times can affect public perception and potentially reduce prospective guest interest. Therefore, efforts to maintain service quality and respond promptly to criticism are essential for preserving hotels' positive images.

Time-series analysis combined with guest sentiment data allows the identification of several key factors that significantly affect room occupancy at Raffles Bali. First, guest review sentiment emerged as a major variable in the SARIMAX predictive model, with a coefficient value of 6.4545. This indicates that any increase in the sentiment score has a significant positive impact on occupancy rates. In this context, an online reputation is a strategic asset for attracting potential guests. Second, seasonality was an important component of the SARIMA model. Historical data show a seasonal pattern in which occupancy tends to rise mid-year, corresponding to the holiday or high season, and decline toward the end of the year, coinciding with the low season. Third, the influence of unexpected events or residual trends is evident, contributing to fluctuations in both sentiment scores and occupancy rates. For example, in early 2024, a spike in residuals indicated the occurrence of an unusual event, such as disturbances caused by renovation work, which not only decreased sentiment scores but also posed a risk to the occupancy. Finally, the quality of service and alignment between guest expectations and the actual experience received were found to impact satisfaction, particularly due to the hotel's branding as part of Accor's "Ultra Luxury Collection." When services failed to meet expectations, as noted in several negative reviews, guests were more likely to leave unfavorable feedback despite luxurious facilities.

The relationship between guest sentiment and room occupancy was further examined using a quantitative approach with the SARIMAX model and Granger Causality testing. Although the Granger Causality test results indicated no statistically significant short-term causal relationship between sentiment and occupancy (as shown by p-values > 0.05 across all lags), the SARIMAX predictive model demonstrated that sentiment still plays an important role as an indicator that can influence occupancy in the long term. This is evidenced by the significant sentiment coefficient in the SARIMAX model, as well as the occupancy forecasts for the next 12 months, which show an upward trend in parallel with stable sentiment scores near the maximum range (between 0.976–0.983). Therefore, although a direct causal relationship was not found in the short term, the long-term association between sentiment and occupancy remains relevant and important for management decision-making.

The findings of this study align with and extend those of previous research. Smith et al. (2021) found that positive customer reviews contribute to the development of a positive hotel image and increased customer trust in the brand. This study reinforces this finding by incorporating a time-series dimension that reveals sentiment fluctuations over time and their connection with hotel occupancy. Furthermore, Wang and Lee (2022) used forecasting methods from online reviews to predict occupancy and showed that review trends have strategic implications for hotel management. This study takes this further by integrating sentiment as an exogenous variable in the SARIMAX model, which has proven effective in predicting future occupancy growth. Rahman et al. (2023) concluded that positive sentiment significantly affects hotel revenue. In this context, the Raffles Bali study also found a correlation between positive guest reviews and increased occupancy, which directly contributed to the room revenue. Finally, Wijaya's (2022) research on sentiment analysis based on social media using the TF-IDF and sentiment classification approach supports the technical methods used in this study, albeit with a different focus on TripAdvisor reviews and on time-series integration. These studies provide a strong theoretical and empirical foundation that supports the current research and highlights the importance of actively and continuously managing guest sentiment.

5. Conclusion

Based on the research findings regarding guest sentiment at Raffles Bali derived from TripAdvisor reviews and analyzed using a time-series approach, it can be concluded that most guest reviews carry a positive sentiment, reflecting a favorable perception of the hotel's service quality and overall guest experience. By utilizing the SARIMA model to analyze sentiment trends over time and the SARIMAX model to correlate sentiment with room occupancy rates, it was found that customer sentiment had a significant positive impact on occupancy levels. An increase in the sentiment score by one point can substantially enhance room occupancy, indicating that customer opinions on online platforms are important indicators of hotel operational performance. Although the Granger Causality test did not reveal a direct short-term causal relationship, the historical sentiment pattern remained useful for long-term forecasting. Therefore, regular sentiment analysis integrated with occupancy forecasting models can serve as an effective strategy for sustainably improving room occupancy at Raffles Bali. Raffles Bali management should actively monitor customer reviews on TripAdvisor as a key indicator of service quality and brand perception. Prompt responses to negative reviews can help maintain a hotel's reputation.

Marketing and customer service teams can leverage the results of this analysis to develop promotional strategies aligned with sentiment trends, such as launching positive campaigns after peak holiday seasons. The regular integration of sentiment and time-series analyses should be incorporated into decision-making systems. This allows occupancy improvement strategies to be based on real-time data and forecast. Future studies should incorporate additional variables, such as room pricing, tourism seasons, local events, and promotional campaigns, to create a more comprehensive predictive model for hotel occupancy. To improve model accuracy, advanced machine learning techniques, such as LSTM or Facebook Prophet, should be considered, as these are more suitable for handling nonlinear and high-dimensional data.

5.1 Limitation

This study was conducted within a limited scope, focusing only on the specific legal provisions and selected case studies related to [topik penelitian kamu]. The research relied primarily on secondary legal sources and did not include extensive fieldwork or empirical data collection. The analysis was confined to the Indonesian legal framework, without in-depth comparison to other jurisdictions that may offer alternative perspectives. Time constraints and limited access to certain legal documents may have also restricted the depth of analysis. These limitations may influence the generalizability and comprehensiveness of the findings.

5.2 Suggestion

Future research should broaden the scope to include comparative legal analysis with other relevant jurisdictions, enabling a more holistic understanding of the topic. Incorporating empirical research methods, such as interviews with legal practitioners, policymakers, and affected communities, could provide richer insights. Further studies could also examine the evolving legal landscape in light of technological developments, policy reforms, or emerging jurisprudence. Expanding the analysis to include socio-economic and cultural dimensions would strengthen the interdisciplinary relevance of the research and support more practical recommendations for legal reform.

References

Aimee, R. M. (2019). A thorough literature review of customer satisfaction definition, factors affecting customer satisfaction and measuring customer satisfaction. *International Journal of Advanced Research, 7*(9), 828-843. doi:http://dx.doi.org/10.21474/IJAR01/9733

Alaimo, C., Kallinikos, J., & Valderrama-Venegas, E. (2020). Platform evolution: A study of tripAdvisor. doi:http://dx.doi.org/10.24251/HICSS.2020.672

Artini, N. W. P., Antara, M., Susrusa, I. K. B., & Ambarawati, I. (2020). Impact of tourism on development in Bali Province. *International Journal of Life Sciences, 4*(2), 19-31.

Barbera, G., Araujo, L., & Fernandes, S. (2023). The value of web data scraping: An application to TripAdvisor. *Big Data and Cognitive Computing, 7*(3), 121.

Chang, Y.-M., Chen, C.-H., Lai, J.-P., Lin, Y.-L., & Pai, P.-F. (2021). Forecasting hotel room occupancy using long short-term memory networks with sentiment analysis and scores of customer online reviews. *Applied Sciences, 11*(21), 10291. doi:https://doi.org/10.3390/app112110291

Chen, T., Samaranayake, P., Cen, X., Qi, M., & Lan, Y.-C. (2022). The impact of online reviews on consumers' purchasing decisions: Evidence from an eye-tracking study. *Frontiers in Psychology, 13*, 865702. doi:https://doi.org/10.3389/fpsyg.2022.865702

Chin, S. W. L. (2017). *Tourism, communities and quality of life indicators in Bali*: Institute of Asian Studies, Universiti Brunei Darussalam.

Das, P., & Barman, S. (2025). Perspective Chapter: An Overview of Time Series Decomposition and Its Applications. *Applied and Theoretical Econometrics and Financial Crises*, 59.

Hamid, S. F. A., Norman, W. M. N. W., Sariman, M. H., Shahudin, N. N., Saleh, M. M., & Mustaffa, W. M. Z. W. (2017). Parental Support and Pressure in Sport: The Perception of Athletes. *Conference: Konferensi Akademik Universiti Teknologi MARA Cawangan Pahang, 1*, 59-64.

Khoironi, T. A., Syah, H., & Dongoran, P. (2018). Product quality, brand image and pricing to improve satisfaction impact on customer loyalty. *International Review of Management and Marketing, 8*(3), 51.
Kim, J. M., & Han, J. (2023). Examining the negative relationship between length of stay at a hotel and customer satisfaction: evidence from online customer reviews. *International Journal of Contemporary Hospitality Management, 35*(12), 4099-4116. doi:https://doi.org/10.1108/IJCHM-04-2022-0437
Liu, W., Liang, Y., Bao, X., Qin, J., & Lim, M. K. (2022). China's logistics development trends in the post COVID-19 era. *International Journal of Logistics Research and Applications, 25*(6), 965-976.
Liu, X., Lin, J., Jiang, X., Chang, T., & Lin, H. (2024). eWOM information richness and online user review behavior: evidence from TripAdvisor. *Journal of Theoretical and Applied Electronic Commerce Research, 19*(2), 880-898. doi:https://doi.org/10.3390/jtaer19020046
Mattila, A. S., & O'Neill, J. W. (2003). Relationships between hotel room pricing, occupancy, and guest satisfaction: A longitudinal case of a midscale hotel in the United States. *Journal of Hospitality & Tourism Research, 27*(3), 328-341. doi:https://doi.org/10.1177/1096348003252361
Mokgehle, D., & Fitchett, J. M. (2024). Efficacy of automating the analysis of TripAdvisor data in tourism and climate studies. *Tourism Planning & Development, 21*(6), 959-979.
Peterhans, S. (2010). Standards, training, and guests' perceptions in luxury hotels.
Phillips, P., Barnes, S., Zigan, K., & Schegg, R. (2017). Understanding the impact of online reviews on hotel performance: an empirical analysis. *Journal of travel research, 56*(2), 235-249. doi:https://doi.org/10.1177/0047287516636481
Sahadev, S., Seiler, A., & Scarf, P. (2025). The impact of review sentiments on occupancy: Evidence for signalling theory from peer-to-peer accommodation. *Journal of Vacation Marketing, 31*(2), 351-365. doi:https://doi.org/10.1177/13567667231201406
Sanjiwani, A. T., Pitanatri, P. D. S., & Loanata, C. P. (2025). Big data analytics to understand guest sentiment: Time series study of tripadvisor reviews for luxury hotel in Bali. *Journal of Multidisciplinary Academic Business Studies (JoMABS), 3*(2), 165-175. doi:https://doi.org/10.35912/jomabs.v3i2.2994
Schiebler, T., Lee, N., & Brodbeck, F. C. (2025). Expectancy-disconfirmation and consumer satisfaction: A meta-analysis. *Journal of the academy of marketing science*, 1-22. doi:https://doi.org/10.1007/s11747-024-01078-x
Toker, T. (2024). Analyzing TripAdvisor Reviews: Investigating Customer Feedback to Enhance Guest Retention Strategies in Leading Ankara Hotels. *Seyahat ve Otel İşletmeciliği Dergisi, 21*(3), 589-614. doi:https://doi.org/10.24010/soid.1432200
Williady, A., Wardhani, H. N., & Kim, H.-S. (2022). A study on customer satisfaction in Bali's luxury resort utilizing big data through online review. *Administrative Sciences, 12*(4), 137. doi:https://doi.org/10.3390/admsci12040137
Wiranatha, A. S., Suryawardani, I. G. A. O., Purbanto, I. G. R., Yudiastina, I. G. B. A., & Bantacut, T. (2024). A conceptual model of rural tourism in bali to build up global competitiveness. *Systems, 12*(7), 245. doi:https://doi.org/10.3390/systems12070245
Zhang, C., Xu, Z., Gou, X., & Chen, S. (2021). An online reviews-driven method for the prioritization of improvements in hotel services. *Tourism Management, 87*, 104382. doi:https://doi.org/10.1016/j.tourman.2021.104382
Zhao, X., Wang, L., Guo, X., & Law, R. (2015). The influence of online reviews to online hotel booking intentions. *International Journal of Contemporary Hospitality Management, 27*(6), 1343-1364. doi:https://doi.org/10.1108/IJCHM-12-2013-0542