



Tourist perceptions of cultural tourism: A sentiment analysis of TripAdvisor reviews

I Made Widiantera^{1,*}, I Nyoman Yoga Segara², I Gusti Ayu Ratna Pramesti Dasih²

¹ Business Administration Department, Politeknik Negeri Bali, Badung, Bali 80361, Indonesia;

² Hindu Communication Science Doctoral Program, Universitas Hindu Negeri I Gusti Bagus Sugriwa, Denpasar, Bali 80235, Indonesia.

*Correspondence: madewidiantera@pnb.ac.id

Received Date: June 12, 2026

Revised Date: October 11, 2026

Accepted Date: October 11, 2026

ABSTRACT

Background: Penglipuran, a Bali Aga village in Bangli, Bali, is an internationally recognized example of community-based cultural tourism whose daily life, customary law, and ceremonies are organized around Balinese Hindu practice. Empirical analyses of how visitors perceive such a living religious-cultural landscape remain scarce. **Methods:** We analysed TripAdvisor reviews of Penglipuran as a religiously framed, cognitively and affectively mediated phenomenon. Of 512 reviews collected (2014–2025), 454 contained sufficient text for computational analysis (58 short entries were excluded); sentiment was scored with VADER and complemented by TextBlob, reviews were coded into Perception and Practice dimensions, and Perception-dimension reviews were modelled thematically with non-negative matrix factorization (NMF), validated against K-means and latent dirichlet allocation (LDA). **Findings:** Sentiment was strongly positive overall (95.8% positive; mean compound 0.751), and the Perception and Practice dimensions did not differ significantly. Three exploratory themes emerged, overall impression and hospitality, the bamboo forest, and a tension between authenticity and commercialization, which align with the three relations of *Tri Hita Karana*. Post-pandemic reviews were significantly less positive than pre-pandemic reviews, indicating that visitor sentiment was not unaffected by the tourism disruption. **Conclusion:** Penglipuran is received favourably, but its appeal rests on a communal and religious life-world that commercialization can erode; protecting community hospitality and ceremonial life, and developing the bamboo forest as a natural-cultural asset, may help sustain it. Because TripAdvisor users are not representative of all visitors, these implications are indicative rather than definitive. **Novelty/Originality of this article:** The study offers an integrated, reproducible pipeline for reading user-generated content as an empirical trace of religious-cultural tourist experience, bringing computational method (science) to the emotional construal (mind) of a Hindu-Balinese heritage landscape (religion).

KEYWORDS: Bali; computational sentiment analysis; cultural tourism; Penglipuran Village; *Tri Hita Karana*.

1. Introduction

Cultural tourism has emerged as one of the fastest-growing and economically significant segments of global tourism, accounting for approximately 40% of all international tourist arrivals and generating substantial multiplier effects across host economies (Richards, 2018; United Nations World Tourism Organization, 2019). Propelled by increasing traveler interest in the living traditions, heritage, and social practices of host communities, cultural tourism differs fundamentally from mass tourism in its emphasis on

Cite This Article:

Widiantera, I. M., Segara, I. N. Y., & Dasih, I. G. A. R. P. (2026). Tourist perceptions of cultural tourism: A sentiment analysis of TripAdvisor reviews. *Dharmakirti: International Journal of Religion, Mind and Science*, 4(1), 19–35. <https://doi.org/10.61511/ijroms.v4i1.2026.3846>

Copyright: © 2026 by the authors. This article is distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).



the co-creation of meaning between visitors and host communities, rendering the quality and authenticity of the visitor experience a central determinant of destination success (Richards, 2018; Zhang et al., 2021). In the Asia-Pacific region, community-based cultural tourism has gained particular policy salience as governments seek to leverage indigenous heritage assets for sustainable economic development while simultaneously preserving intangible cultural heritage under threat from rapid urbanization and globalization (United Nations World Tourism Organization, 2019). The tension between heritage preservation and tourism commercialization has been identified as one of the most persistent governance challenges in this region, where commercial viability and community welfare must be carefully balanced against cultural sustainability (Wang et al., 2024).

Within this regional landscape, Indonesia occupies a strategic position as one of the world's most culturally diverse tourism destinations, with Bali functioning as the nation's primary cultural tourism hub and international gateway. In 2024, Bali welcomed approximately 6.33 million international visitors and 10.12 million domestic tourists, representing a full recovery to pre-pandemic levels and a 7.9% increase over 2023 figures. Cultural tourism constitutes the primary motivational driver for a significant proportion of these arrivals, with Hindu-Balinese customs, traditional performing arts, temple ceremonies, and village life forming the experiential core of the island's tourism offer (Richards, 2018). Unlike mass resort destinations, Bali's competitive advantage rests substantially on the perceived authenticity and cultural coherence of its heritage landscape attributes that are simultaneously its most valuable tourism asset and its most vulnerable resource (Li et al., 2023).

Among Bali's diverse cultural attractions, Penglipuran Village in Bangli Regency stands as an exceptional case of indigenous heritage preservation. Classified as a *Bali Aga* settlement a pre-Majapahit indigenous community estimated to be over 700 years old, the village is internationally recognized as one of the world's cleanest and best-preserved traditional villages, receiving the 2023 United Nations World Tourism Organization (UNWTO) Best Tourism Village award (Kompas, 2023). Its distinctive characteristics include a uniformly preserved traditional compound architecture (*pekarangan*) arranged symmetrically along a north-south axis, the prohibition of motorized vehicles on its main ceremonial thoroughfare, the continued observance of customary law (*awig-awig*), and a community-managed bamboo conservation forest bordering its northern perimeter. These attributes collectively produce a heritage landscape that is visually coherent, ecologically conscious, and culturally legible to a broad international visitor base, making Penglipuran an exemplary site for the empirical study of cultural tourism perception.

Tourist perception in cultural heritage settings is understood as a multidimensional construct encompassing cognitive, affective, and conative components through which visitors assign meaning to their experiences (Huete-Alcocer et al., 2019; Li et al., 2023). The cognitive dimension refers to the beliefs and knowledge tourists form about a destination's cultural attributes; the affective dimension captures the emotional responses evoked during the visit; and the conative dimension reflects behavioral intentions such as revisit likelihood and recommendation behavior (Huete-Alcocer et al., 2019). Central to cultural tourism perception is the concept of authenticity, which Wang (1999) theorizes in three enduring forms: objective authenticity, the genuine nature of heritage objects; constructive authenticity, the social negotiation of what tourists perceive as authentic based on prior expectations and cultural frames; and existential authenticity, the bodily and emotional feelings of genuineness activated through participation in cultural activities.

In community-based heritage village settings, the balance between these forms of authenticity and the commercial demands of tourism frequently becomes contested. Tourism commercialization positively affects constructive and existential authenticity perceptions and tourist satisfaction in heritage contexts; however, excessive commodification risks generating a perceived authenticity deficit that ultimately undermines the destination's cultural appeal (MacCannell, 1973; Wang et al., 2024). The dialectical relationship between authenticity and alienation in intangible cultural heritage tourism represents a key tension in the field, particularly at sites where back-stage

community life is progressively displaced by front-stage commercial display (Zhang & Lee, 2022).

This study is positioned at the intersection of the three domains that define the journal's scope. Tourist perception of Penglipuran is treated not as a tourism-management variable alone but as a religiously framed, cognitively and affectively mediated phenomenon that can be observed empirically through computational analysis of visitor discourse. Penglipuran is a living Bali Aga community whose spatial order, customary law (*awig-awig*), and daily life are organized around Hindu religious practice; visitors' evaluations are therefore, in part, responses to a sacred and communal order rather than to a purely aesthetic or commercial product.

Perception in heritage settings is a multidimensional cognitive–affective construct encompassing the beliefs tourists form about a destination, the emotions its environment evokes, and the behavioural intentions that follow, its cognitive, affective, and conative components (as cited above). Sentiment, the affective valence expressed in a review is thus a measurable trace of the emotional component of this meaning-making process. By analysing sentiment and its thematic organization, we access how visitors cognitively and emotionally construe the village, not merely whether they enjoyed it.

Penglipuran's distinctiveness is inseparable from its religious life. The Balinese Hindu philosophy of *Tri Hita Karana*, harmony between the human and the divine (*parahyangan*), among people (*pawongan*), and between people and nature (*palemahan*), organizes the village's temples, compound architecture, ceremonies, and the conservation of its bamboo forest (Peters & Wardana, 2013). Read through this lens, the facets visitors repeatedly articulate map onto the three relations of *Tri Hita Karana*: community hospitality and village life (*pawongan*), the bamboo forest and natural setting (*palemahan*), and the ceremonial–architectural order anchored by the temple (*parahyangan*). The recurring tension between cultural authenticity and commercialization is, in these terms, a tension in religious and communal meaning-making, the concern that a sacred, communal life-world is being displaced by a front-stage arranged for consumption (MacCannell, 1973; Wang, 1999), rather than a narrowly economic complaint.

User-generated review platforms make this religiously framed, cognitively mediated experience observable at scale. Computational sentiment analysis (Hutto & Gilbert, 2014) and unsupervised thematic modelling provide a reproducible, quantifiable method for recovering the affective and thematic structure of visitor experience from a decade of unsolicited testimony. The study therefore brings a scientific, data-driven method (science) to bear on the emotional and interpretive construal (mind) of a religious-cultural landscape (religion); it is this triangulation, not the sentiment pipeline alone that constitutes the manuscript's contribution to the journal's interdisciplinary scope. Accordingly, the research questions below examine not only the valence and thematic organization of perception but also how the religious-communal dimensions of the village surface in, and are negotiated through, visitor discourse.

The study of tourist perception has been substantially transformed by the proliferation of online travel review platforms, which generate large-scale, ecologically valid, and demographically diverse corpora of visitor evaluations available for computational analysis (Lu & Stepchenkova, 2015; Taecharungroj & Mathayomchan, 2019). TripAdvisor, as one of the world's most widely studied travel platforms, provides longitudinal, multilingual, and candid review data that enable researchers to examine tourist perceptions at scale and across extended time periods (Filiari et al., 2021). User-generated content has been applied to test and extend memorable tourist experience scales across diverse cultural tourism contexts, confirming the analytical value of online review corpora for destination perception research (Bigne et al., 2020). Such data capture voluntary, unsolicited expressions of tourist experience that structured survey instruments cannot fully replicate, a methodological advantage of particular significance when investigating emotionally salient cultural heritage destinations (Lu & Stepchenkova, 2015).

Computational sentiment analysis using tools such as the Valence Aware Dictionary and sEntiment Reasoner (VADER) has proven effective in extracting nuanced emotional

orientation from the informal, mixed-register texts characteristic of online travel reviews (Hutto & Gilbert, 2014). In tourism research, VADER and related lexicon-based methods have been successfully applied to restaurant experiences (Jia, 2020), cultural heritage attractions (Islam, 2025), and cultural-heritage tourism in the Indonesian context (Alfaritzi et al., 2026). Large-scale sentiment analysis of multilingual review corpora has shown that expressed tourism sentiment varies systematically with tourists' nationality and language, underscoring the value of applying such methods to multilingual data (Wang & Kirilenko, 2021). Furthermore, combining sentiment classification with thematic text clustering can reveal latent perceptual structures within tourist discourse that aggregate sentiment scores alone are unable to capture (Mehraliyev et al., 2022).

Despite this body of methodological and empirical literature, studies of tourist perception specifically at Penglipuran Village remain limited in scope and scale. A foundational assessment of community-based tourism governance at the village has focused primarily on government-community relations rather than large-scale visitor perception analysis (Adi et al., 2017). Key factors in sustainable tourism development at Penglipuran have been examined through prospective analysis, though without integrating computational methods or UGC data (Sutrisna et al., 2020). The experience economy framework, while applicable to similar heritage settings, has not been specifically applied to Penglipuran (Pine & Gilmore, 2011). No published study has yet undertaken a large-scale computational sentiment analysis of tourist perceptions at the village integrating dimensional coding and unsupervised text clustering within a unified analytical framework.

We distinguish two complementary facets of experience: Perception, the evaluative, cognitive-affective impressions visitors form of the destination's cultural identity and atmosphere and Practice, the embodied activities and encounters they describe. The distinction is theoretically motivated: evaluative perception and participatory practice activate different modes of authenticity (constructive versus existential; Wang, 1999), so separating them allows us to test whether evaluative and participatory stances carry different affective signatures.

Accordingly, the study is guided by four research questions: RQ1, What is the overall sentiment orientation and emotional intensity of tourist perceptions of Penglipuran across 2014–2025, and did it change across the pandemic disruption?. RQ2, Do the Perception and Practice dimensions of visitor experience differ in sentiment?. RQ3, What latent themes structure visitors' perceptual discourse, and how do they relate to the religious-communal order of the village (*Tri Hita Karana*)?. RQ4, How is the tension between cultural authenticity and tourism commercialization expressed in visitor discourse, and what does it reveal about the negotiation of a sacred-communal life-world?

Despite a growing literature applying sentiment analysis to tourism user-generated content, no study has analysed visitor perception at Penglipuran at decade scale with an integrated, validated computational pipeline, and none has framed such analysis within a religious–cognitive–scientific perspective. This study addresses that gap and contributes in three ways. Methodologically, it offers a reproducible pipeline, multilingual sentiment scoring, validated dimensional coding, and NMF topic modelling with explicit stability diagnostics for reading heritage-site UGC. Empirically, it provides the first decade-scale account of visitor sentiment at Penglipuran, including a previously unreported post-pandemic decline. Theoretically, it reframes computational perception analysis within the journal's Religion–Mind–Science scope by reading sentiment as an empirical trace of religiously situated experience, interpreted through *Tri Hita Karana*.

2. Methods

This study employed a sequential mixed-methods analytical design combining quantitative computational text analysis with qualitative thematic interpretation of online review data. The analytical framework integrates four sequential phases: (1) systematic data collection, (2) text preprocessing and multilingual translation, (3) computational sentiment scoring and classification, and (4) dimensional coding, unsupervised clustering,

and interpretive synthesis. The use of publicly available user-generated content (UGC) as the primary data source means the study did not involve interaction with human participants; ethical considerations relating to privacy, data minimization, and platform terms are addressed in the Ethical Review Board Statement.

2.1 Data collection and source

Review data were collected from the public TripAdvisor listing for Penglipuran Village on 12 July 2025 using a structured web-scraping procedure implemented in Python 3.10. Collection was carried out in compliance with the platform's terms of service, and only aggregate and derived data, sentiment scores, dimensional labels, and paraphrased excerpts are reported; no raw review text that could identify individuals is redistributed. The final dataset comprised 512 reviews published between January 2014 and July 2025, representing all reviews retrievable from the listing on the collection date rather than a guaranteed census of every review ever posted, since platform display and ranking can limit visibility. For de-duplication, reviewer display names were used together with review dates to remove exact duplicates; these identifiers were subsequently discarded, and the analytical dataset contains no reviewer names, usernames, geographic locations, or other direct identifiers. Of the 512 collected reviews, 454 (88.7%) contained sufficient original text content for full computational processing; the remaining 58 comprised very short entries (fewer than five tokens after cleaning) and were retained in aggregate rating statistics but excluded from inferential analysis.

2.2 Preprocessing and translation

Language identification was applied to each review entry using the langdetect library prior to translation. Non-English reviews were processed through an automated neural machine translation service to produce two derived text fields: `text_en` (English translation) and `text_id` (Indonesian translation). The translation-first approach to multilingual sentiment analysis is methodologically supported by evidence that applying established English-language sentiment tools to translated text consistently outperforms language-specific sentiment approaches across diverse target languages (Alsharhan et al., 2022). All subsequent sentiment scoring was performed exclusively on the `text_en` field.

All non-English reviews were translated into English using the Google Cloud Translation API (accessed July 2025), following language identification with the langdetect library; all subsequent sentiment scoring was performed on the English text. Because machine translation can attenuate or alter sentiment-bearing cues (intensifiers, negation, idiom), the validity of the translate-then-score pipeline is assessed downstream against external criteria (reviewers' star ratings) and an independent lexical tool (TextBlob) in Section 2.3, and the translation-first approach to multilingual sentiment is supported by prior evidence (Alsharhan et al., 2022). Residual translation noise is treated as a limitation (see Limitations). Text normalization was then applied to the English corpus through a standardized pipeline comprising: (1) HTML entity decoding; (2) removal of URLs, email addresses, and platform-specific metadata strings; (3) Unicode normalization to ASCII; (4) whitespace normalization; and (5) removal of residual encoding artifacts introduced during scraping. No stemming or stop-word removal was applied at this stage, as VADER's lexicon-based approach requires preservation of syntactic context including degree modifiers, negations, and punctuation signals that are critical to accurate valence scoring (Hutto & Gilbert, 2014).

2.3 Sentiment analysis and scoring

Each preprocessed English review (`text_en`) was scored using the VADER (Valence Aware Dictionary and sEntiment Reasoner) lexicon- and rule-based sentiment analysis tool (Hutto & Gilbert, 2014). VADER is particularly suited to the informal, expressive, and

punctuation-rich character of online travel reviews because its scoring algorithm accounts for grammatical context including negation modifiers (not, never), degree intensifiers (very, extremely), capitalisation emphasis, and emotionally emphatic punctuation. VADER produces four normalized scores per document: pos (proportion of positive-valence tokens), neg (proportion of negative-valence tokens), neu (proportion of neutral tokens), and compound a normalised weighted composite ranging continuously from -1.0 (maximally negative) to +1.0 (maximally positive). The compound score is computed by summing all token valence ratings, applying contextual adjustments, and normalizing using a non-linear transformation (Hutto & Gilbert, 2014).

Reviews were classified into binary sentiment categories using the compound threshold established in VADER's original specification: entries with compound ≥ 0.05 were classified as positive; those with compound ≤ -0.05 as negative. Emotional intensity was operationalized as a three-tier ordinal variable derived from the absolute magnitude of the compound score: Low ($|\text{compound}| < 0.3$), Medium ($0.3 \leq |\text{compound}| < 0.6$), and High ($|\text{compound}| \geq 0.6$). This intensity classification captures the strength of tourist affect independently of its direction a distinction of particular relevance in cultural tourism research, where strong positive and strongly critical responses carry different managerial implications (Mehraliyev et al., 2022).

Complementary to VADER, TextBlob was applied to produce two supplementary measures: polarity (a signed score from -1.0 to +1.0) and subjectivity (0 to 1.0), providing additional quantification of sentiment orientation and the degree of personal opinion expressed relative to factual description (Loria, 2018). A comparative study of VADER and TextBlob on hotel review corpora confirms that the two tools produce complementary rather than redundant outputs, with VADER more sensitive to emotional intensity and TextBlob more sensitive to subjectivity orientation (Nariman, 2025).

Two checks addressed the validity of applying VADER to translated, originally multilingual review text, using external criteria rather than additional manual annotation. First, as a convergent-validity check, VADER compound scores were compared with TextBlob polarity scores computed on the same texts; the two independent tools were moderately and significantly correlated (Spearman's $\rho = 0.41$, Pearson's $r = 0.45$, $n = 472$). Second, as an external-criterion check, VADER scores were compared with reviewers' own star ratings: compound scores correlated positively and significantly with star ratings (Spearman's $\rho = 0.19$, Pearson's $r = 0.28$, $p < 0.001$, $n = 472$), and VADER's positive/negative classification agreed with a rating-based label (≥ 4 stars positive, ≤ 2 negative) at 92.4% accuracy (precision 0.97, recall 0.95, F1 0.96 for the positive class; $n = 432$). The modest Cohen's κ (0.23) reflects the strong positivity skew of the corpus rather than poor classification. Together, these checks indicate that the translate-then-score pipeline captures sentiment consistent with an independent lexical tool and with reviewers' explicit ratings; the weak-to-moderate magnitudes nonetheless counsel caution in over-interpreting fine-grained score differences (full results in Supplementary Table S2).

2.4 Dimensional coding

Each of the 454 reviews was assigned to one of two mutually exclusive analytical dimensions. The Perception dimension captured cognitive and affective evaluations, comprising tourists' overall impressions of the destination's cultural identity, environmental quality, atmosphere, and authenticity (Huete-Alcocer et al., 2019; Li et al., 2023). The Practice dimension captured behavioral descriptions of specific activities, service encounters, and social interactions experienced during the visit. This binary dimensional framework draws on the established distinction in heritage tourism perception research between evaluative impressions formed about a place and enacted experiences within it (Taecharungroj & Mathayomchan, 2019).

Dimensional assignment was performed using a rule-based keyword classifier built in Python. The classifier was trained on a manually labeled reference set of 80 reviews (40 per dimension) selected through stratified random sampling across rating tiers (1–5 stars). The

Perception dimension was operationalized through keyword clusters associated with assessment language (e.g., beautiful, authentic, staged, atmosphere, heritage, impressive), while the Practice dimension was operationalized through activity and behavioral language (e.g., wore, walked, rented, visited, tried, photographed). The two independent coders' labels on this 80-review reference set yielded an inter-rater reliability of Cohen's $\kappa = 0.78$, within the 0.61–0.80 range conventionally interpreted as substantial agreement (Landis & Koch, 1977).

The 80 manually labelled reviews (40 per dimension) were used to develop and tune the keyword classifier rather than as an independent held-out test set. We therefore report the inter-coder reliability of these gold labels (Cohen's $\kappa = 0.78$, indicating substantial agreement) and clarify that this describes agreement between the two human coders, not the predictive performance of the automated classifier. To support transparency and reproducibility in the absence of a held-out evaluation, the complete keyword dictionary and matching rules are provided as Supplementary Material S1. A formal held-out evaluation of the automated classifier was beyond the scope of this study and is noted as a limitation and a direction for future work. The complete keyword dictionary and the matching rules used for dimensional assignment, including handling of context and negation, and the rule for reviews containing cues for both dimensions (assigned to the dimension with the higher weighted keyword count) are provided as Supplementary Material S1 to ensure transparency and reproducibility.

2.5 Topic modeling of the perception dimension

Reviews assigned to the Perception dimension ($n = 252$, non-neutral) were analyzed using non-negative matrix factorization (NMF) to identify latent thematic structure within tourist perceptual discourse (Lee & Seung, 1999). Review texts were first transformed into TF-IDF feature vectors (Salton & Buckley, 1988) using unigram and bigram features, a minimum document frequency of three, and a maximum of 5,000 features, with English stop words removed. Because a small set of ubiquitous destination terms (e.g., village, Penglipuran, Bali, traditional) occurred in nearly every review and dominated all candidate topics, these terms were added to the stop-word list so that discriminating vocabulary could emerge. NMF was initialized with non-negative double singular value decomposition (nndsvd) (Boutsidis & Gallopoulos, 2008) and a fixed random seed (random_state = 42) for reproducibility, and each review was assigned to the topic of highest membership weight. All procedures were implemented in scikit-learn (Pedregosa et al., 2011).

Table 1. Research design summary

Phase	Procedure	Tool/Method	Output
1	Data collection	Web scraping (Python)	512 reviews (2014–2025)
2	Preprocessing	langdetect, neural MT	text_en, text_id fields
3	Sentiment scoring	VADER + TextBlob	neg, neu, pos, compound, polarity, subjectivity
4	Classification	Threshold rules	Sentiment label; Intensity (Low/Med/High)
5	Dimensional coding	Rule-based classifier	Perception ($n=252$) / Practice ($n=202$)
6	Cluster analysis	TF-IDF + K-means ($k=3$)	3 thematic clusters
7	Synthesis	Descriptive statistics	Tables, interpretive discussion

The three-topic solution was selected after comparing $k = 2-6$ and evaluating alternative algorithms. An initial K-means solution on the same features produced weak separation (silhouette coefficient ≈ 0.01 , with no clear inflection in the within-cluster sum of squares across $k = 2-8$), indicating that the corpus is thematically homogeneous rather than composed of sharply bounded clusters. NMF and latent Dirichlet allocation (Blei et al., 2003) were therefore evaluated as topic-model alternatives. Although latent dirichlet allocation (LDA) attained a marginally higher UMass coherence score, its topics overlapped

substantially, whereas NMF produced the most semantically distinct and interpretable topics; NMF was adopted as the primary model, with K-means and LDA reported as robustness comparisons. Topic labels were assigned by inspecting the highest-weighted terms per topic and reading the most representative reviews per topic, consistent with established practice in user-generated-content thematic analysis (Richards, 2018). Solution stability was assessed with a 100-replicate bootstrap, re-fitting NMF on each resample and aligning topic labels by optimal matching; the resulting agreement was modest (mean adjusted Rand index ≈ 0.15), which, together with a near-zero silhouette coefficient, indicates that the corpus is thematically homogeneous and that the topics are best interpreted as an exploratory thematic structure rather than as sharply separated clusters. Table 1 provides a summary of the research design, including data sources, analytical tools, and outputs for each methodological phase.

3. Results and Discussion

3.1 Rating distribution and overall tourist satisfaction

The aggregate rating distribution across all 512 TripAdvisor reviews is presented in Table 2. The overall mean rating of 4.3 out of 5.0 signals a high level of tourist satisfaction with Penglipuran Village as a cultural tourism destination. A combined 87.7% of reviewers awarded four or five stars, reflecting strong positive endorsement, while critically negative ratings (1–2 stars) accounted for only 3.5% of all entries suggesting that active visitor dissatisfaction is marginal rather than systemic. This result is consistent with satisfaction indices reported for similar heritage tourism sites in Asia, where studies of ancient town and traditional village destinations have found mean ratings in the 4.0–4.5 range (Islam, 2025; Zhang et al., 2021).

Table 2. Rating distribution

Rating	Frequency (n)	Percentage (%)	Cumulative (%)	Satisfaction Category
5 ★	236	46.1	46.1	Excellent
4 ★	213	41.6	87.7	Very Good
3 ★	45	8.8	96.5	Average
2 ★	10	2.0	98.4	Poor
1 ★	8	1.5	100.0	Terrible
Total	512	100.0	—	Mean = 4.3/5.0

(TripAdvisor, 2014–2025)

Note. Satisfaction categories correspond to TripAdvisor star ratings: 5 = Excellent, 4 = Very good, 3 = Average, 2 = Poor, 1 = Terrible

The high proportion of five-star reviews (46.1%) is particularly notable. Research on memorable cultural tourism experiences confirms that heritage settings capable of delivering immersive, aesthetically coherent, and socially meaningful encounters consistently elicit the most enthusiastic visitor evaluations (Bigne et al., 2020). In Penglipuran's case, the uniform visual coherence of the traditional compound streetscape, the absence of motorized vehicles, and the bamboo conservation forest collectively generate a sensory environment that departs sharply from everyday urban experience conditions theorized by Pine & Gilmore (2011) as prerequisites for experience-economic value creation. The absence of a bimodal distribution further indicates that the village does not polarize its visitor audience between enthusiasts and critics a pattern consistent with the predominantly homogeneous physical and cultural landscape of the destination, which leaves limited scope for ambiguity in first-impression evaluation.

The 3-star review cluster (8.8%) warrants separate attention. Qualitative reading of reviews in this tier reveals a pattern of visitors who found the destination visually attractive but were disappointed by the degree of commercial activity within the traditional compounds, or who considered the visit too brief given the entry fee. This sub-group can be

characterized as critically engaged visitors with moderately fulfilled expectations neither hostile nor enthusiastic a visitor profile that merits specific attention in FFIGvisitor communication and experience design strategies (Taecharungroj & Mathayomchan, 2019).

3.2 Sentiment analysis results

The VADER sentiment analysis of 454 reviews yielded the results summarised in Table 3. An overwhelming 95.8% of reviews were classified as positive, with a mean compound score of 0.751 (SD = 0.284). VADER's scoring convention classifies compound scores above 0.05 as positive, with values above 0.5 reflecting distinctly positive sentiment orientation (Hutto & Gilbert, 2014). A mean compound score of 0.751 therefore represents a strongly affirmative overall tourist perception a figure that notably exceeds the mean compound scores of roughly 0.45–0.65 reported for hotel and restaurant review corpora analysed with comparable lexicon methods (Jia, 2020; Nariman, 2025), and is consistent with the strongly positive sentiment reported for other heritage attractions on TripAdvisor (Saoualih et al., 2024), suggesting that cultural village tourism elicits more uniformly and intensely positive emotional responses than accommodation-based experiences. Sentiment classification is not equivalent to satisfaction: a positive VADER score indicates a positive affective tone in the review text, not a measured level of visitor satisfaction.

Table 3. VADER sentiment analysis and TextBlob scoring summary (n = 454)

Metric	Positive Reviews	Negative Reviews	Overall Mean	Std. Deviation
Count (n)	435 (95.8%)	19 (4.2%)	454	—
Compound score	0.801	−0.387	0.751	0.284
Polarity	0.278	0.038	0.268	0.191
Subjectivity	0.593	0.535	0.591	0.211
Positivity (pos)	0.212	0.090	0.198	0.120
Negativity (neg)	0.013	0.095	0.017	0.038

The mean subjectivity score of 0.591 indicates that reviews occupy a moderately opinionated register blending factual descriptions of heritage features (architecture, temple structures, bamboo forest) with personal evaluations of atmosphere and authenticity. This mixed factual-evaluative profile is characteristic of cultural tourism UGC, where visitors engage simultaneously as observers of heritage and as evaluators of personal experience (Lu & Stepchenkova, 2015; Taecharungroj & Mathayomchan, 2019). The relatively narrow difference in subjectivity between positive (0.593) and negative (0.535) reviews is noteworthy, it suggests that critical reviewers are not merely more emotional but are engaging in comparably deliberate evaluative reasoning, typically drawing on contrasts between expected and experienced authenticity.

Emotional intensity analysis revealed that 83.0% of reviews demonstrated high intensity ($|\text{compound}| \geq 0.6$), with only 5.1% exhibiting low intensity. This distribution indicates that Penglipuran consistently provokes strong emotional responses an outcome consistent with the characteristics of destinations that deliver memorable experiential encounters aligned with visitors' identity and aesthetic sensibilities (Bigne et al., 2020). Positive high-intensity reviews frequently deployed emotionally charged language describing the village as tranquil, breathtaking, and unlike anywhere else in Bali. The 19 negative reviews, despite comprising only 4.2% of the corpus, showed markedly elevated negativity scores (mean neg = 0.095 vs. 0.013 for positive reviews) and characteristically centred on perceived mismatches between expectations of an authentic traditional village and the commercial reality encountered a recurring theme explored further in Section 3.6.

3.3 Dimensional analysis: Perception versus practice

The dimensional coding partitioned the 454 reviews into 252 Perception-coded (55.5%) and 202 Practice-coded (44.5%) entries (Table 4). Both dimensions were strongly

positive. A Mann–Whitney U test found no statistically significant difference in compound scores between them ($U = 26,627$, $p = 0.398$); we therefore describe the two dimensions as comparably positive rather than claiming that Practice generates more positive affect. Any theoretical reading of the small numerical gap is presented as interpretation, not as a measured effect.

Table 4. Dimensional analysis: Perception and practice dimensions ($n = 454$)

Dimension	n	% of Total	Positive (%)	Mean Compound	High Intensity (%)
Perception	252	55.5%	94.0%	0.746	83.0%
Practice	202	44.5%	98.0%	0.758	83.2%
Total	454	100%	95.8%	0.751	83.0%

The near-identical sentiment of the two dimensions is consistent with evaluative and participatory stances carrying similar affective tone in this corpus. Because the small numerical gap is not statistically significant (see the test above), we refrain from claiming that participatory “Practice” generates more positive affect than evaluative “Perception.” The slightly higher share of negative reviews within the Perception dimension (6.0% vs. 2.0%) is best read as a descriptive pattern — consistent with the greater cognitive complexity of evaluative appraisal, in which visitors compare pre-arrival images against the commercial reality encountered — rather than as a tested effect. Even so, the Perception dimension still recorded 94.0% positive reviews, confirming that critical visitors remain a clear minority.

3.4 Thematic structure of the perception dimension

NMF topic modeling of the 252 non-neutral Perception-dimension reviews produced three interpretable themes (Table 5). Read together, they describe a visitor population that values the village’s aesthetic coherence and community warmth while remaining attentive to the authenticity implications of commercialization.

Table 5. Thematic structure of the perception dimension

NMF topic	Top terms	n (%)	Mean compound	% positive	Mean rating
Houses, local life & commercialization	houses, people, shops, tourists, worth, live	119 (47.2%)	0.650	89.9%	3.96
Overall impression & hospitality	beautiful, clean, friendly, local, culture	89 (35.3%)	0.827	96.6%	4.43
Bamboo forest/natural heritage	forest, bamboo, walk, pleasant, street	44 (17.5%)	0.839	100.0%	4.25

Note. Based on 252 non-neutral Perception reviews (7 neutral excluded), matching the analytical sample. Sentiment computed with VADER on the English-translated text. 2025 figures are partial (January–July). Source: authors’ analysis

Topic 1 - Houses, local life and commercialization ($n = 119$, 47.2%) was the largest theme and recorded the lowest sentiment (mean compound 0.650; 89.9% positive) and the lowest mean rating (3.96). Its terms (houses, people, shops, tourists, worth, live) combine appreciation of the preserved compound architecture and everyday village life with the corpus’s most critical voices, which centre on the conversion of family compounds into near-uniform souvenir retail. We read this tension as an interpretation of the lexical and sentiment evidence, not a direct measurement through MacCannell (1973) staged authenticity and Wang (1999) distinction between existential and constructive authenticity, whereby a commercialized front-stage can signal that the authentic back-stage of community life has been displaced. That the theme nonetheless remained predominantly

positive indicates that ambivalence about commercialization did not, for most visitors, override an overall favourable assessment.

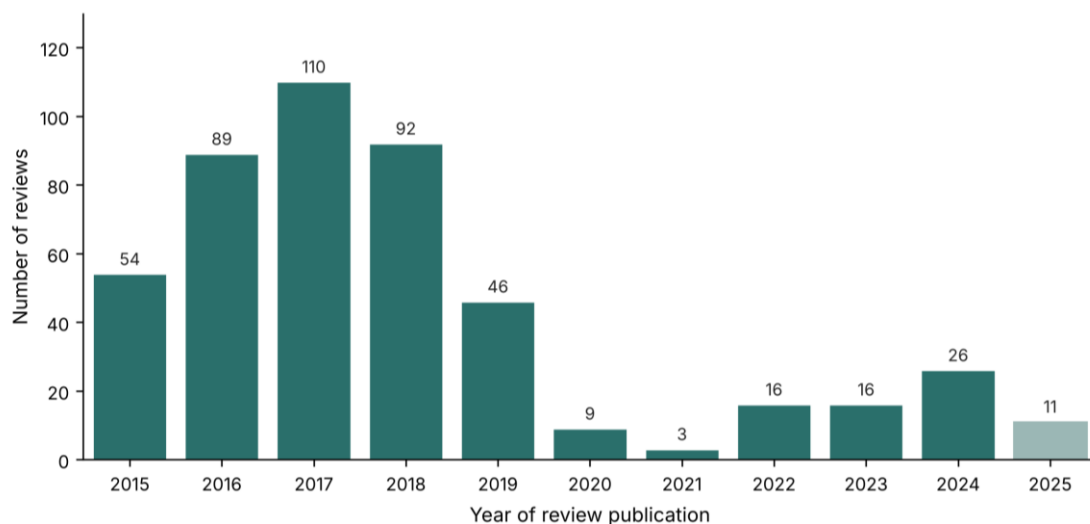
Topic 2 - Overall impression and hospitality ($n = 89$, 35.3%) recorded a high mean compound score (0.827), 96.6% positive reviews, and the highest mean rating (4.43). Leading terms (beautiful, clean, friendly, local, culture) reflect holistic appraisals of the village as a clean, peaceful, and welcoming environment, consistent with what Pine & Gilmore (2011) describe as an esthetic experience. We interpret, rather than measure directly, the recurring references to resident warmth as an expression of the social dimension of the cultural tourism encounter.

Topic 3 - Bamboo forest and natural heritage ($n = 44$, 17.5%) recorded the highest sentiment of the three themes (mean compound 0.839; 100% positive) and a mean rating of 4.25. Centred on forest, bamboo, walk, and pleasant, it captures visitors who extended their itinerary to the community-managed bamboo conservation forest and described it as a refreshing counterpoint to the built-heritage precinct, consistent with Richards's (2018) argument that integrating natural and cultural heritage enhances perceived visit quality.

These topic-level differences in sentiment are reported descriptively; we do not claim they are statistically tested, and the three themes are best understood as an exploratory thematic structure rather than as discretely bounded clusters. A bootstrap stability analysis returned only modest agreement across resamples (mean adjusted Rand index ≈ 0.15), confirming that these themes represent an interpretive reading of a largely homogeneous, predominantly positive corpus; the bamboo-forest theme was the most consistently distinct, and the authenticity-commercialization tension the most consistently recurrent.

3.5 Temporal trends

The corpus spans 2014–2025 (Figure 1). Review volumes rose from a handful of entries in 2014–2015 to a sustained peak across 2016–2018, declined in 2019, and collapsed during the pandemic (single-digit counts in 2020–2021), before a partial recovery in 2022–2025; the 2025 figure covers January–July only. A Mann–Whitney U test comparing pre-pandemic (≤ 2019) and post-pandemic (2022–2025) compound scores found a statistically significant difference ($U = 15,314$, $p = 0.002$), with post-pandemic reviews notably less positive on average (mean compound 0.59, $n = 66$) than pre-pandemic reviews (0.78, $n = 376$).



Note: $n = 472$ reviews with computed sentiment scores. 2025 covers January–July only (partial year).

Fig. 1. Annual TripAdvisor review volumes, Penglipuran Village (2015–2025)

This result should be read cautiously: the post-pandemic sample is small and 2025 is partial, and the decline may reflect changes in visitor composition, heightened sensitivity to crowding and commercialization as mass tourism resumed, or both, rather than a

deterioration in the village's heritage itself. Accordingly, we do not interpret recovering review volumes as evidence that visitor sentiment was unaffected by the disruption; the affective tone of post-recovery reviews is measurably more critical, a pattern that warrants monitoring.

3.6 Authenticity, commercialisation, and hospitality

Beyond the quantitative findings, close reading of review texts reveals three qualitative themes of significance for cultural tourism management at Penglipuran. Authenticity versus commercialization, the tension between authentic cultural preservation and visitor-oriented commercialization is the most consistently articulated critical theme across the corpus, and it is the most consistently articulated critical theme across the corpus, and it is concentrated almost entirely within the houses, local life and commercialization theme (Topic 1) of the Perception dimension. Visitors who arrive with expectations shaped by travel media portrayals of Penglipuran as an unspoiled traditional village frequently report a cognitive gap upon encountering the commercial density of the main thoroughfare, where virtually every family compound entrance presents souvenir displays, food stalls, or clothing rental services. This commercial landscape produces what Zhang et al. (2022) describe as a form of alienation from intangible cultural heritage the sense that the authentic life-world of the community has been displaced by an economic performance for external audiences. The tension reflects a well-documented paradox in heritage tourism: the revenue generated by commercial activity within traditional spaces sustains the economic viability of community-based conservation, yet the density and uniformity of that commercial presence can paradoxically signal to visitors that the cultural authenticity they came to experience has already been compromised (Wang et al., 2024; Zhang et al., 2021). Strategic zoning of commercial activities separating retail zones from ceremonial and residential areas emerges as the most practically actionable management response to this perceptual challenge.

Bamboo forest as a natural heritage complement, the community-managed bamboo conservation forest bordering the village's northern perimeter emerged as a recurrently praised supplementary attraction across all cluster groups. Visitors who extended their itinerary to include the bamboo forest trail described the experience in markedly more superlative terms than those who restricted their visit to the main village street, suggesting that the natural heritage environment provides an experiential counterweight to the commercial saturation of the built heritage precinct. This finding supports the value of integrated nature-culture tourism product development and resonates with Richards (2018) observation that multi-dimensional cultural tourism offerings combining tangible heritage, living culture, and natural landscapes, consistently outperform single-attribute destinations in visitor satisfaction and loyalty metrics. The bamboo forest represents an underutilized differentiating asset that destination managers should develop more strategically as a complementary experiential draw.

Community hospitality as lived religion (*pawongan* and *parahyangan*), a recurrent positive theme across all sentiment categories concerned the warmth of residents, being greeted in the street, invited into compounds, and welcomed to observe segments of Hindu ceremonies, including wedding rites. Read within the study's framework, these encounters are not merely service interactions or photogenic attractions but moments of participation in the village's religious and communal life: visitors are briefly admitted to the *pawongan* (human-human) and *parahyangan* (human-divine) relations that *Tri Hita Karana* places at the centre of Balinese existence. The strongly positive affect attached to these moments (see Topic 2) is consistent with Wang (1999) existential authenticity, in which genuineness is felt through participation rather than observed in objects. This religious-communal dimension, rather than the built fabric alone is what visitors most often single out as distinctive, and it is precisely what commercialization threatens when compounds are reorganized around retail; the authenticity-commercialization tension therefore concerns, at root, the erosion of a sacred-communal order, not only an economic one. We present this

as an interpretation of the review discourse, consistent with the sentiment and thematic evidence, rather than as a direct measurement of religious experience.

4. Conclusions

This study examined tourist perceptions of cultural tourism at Penglipuran Village through computational sentiment analysis and thematic modelling of 454 TripAdvisor reviews spanning 2014–2025, framed as a religiously situated, cognitively and affectively mediated phenomenon (Religion, Mind, and Science). Overall sentiment was strongly positive (95.8% positive; mean VADER compound 0.751; aggregate rating 4.3/5.0), and the Perception and Practice dimensions were comparably positive, with no statistically significant difference between them. Within the Perception dimension, NMF topic modelling surfaced three recurring themes, overall impression and hospitality, the bamboo forest, and the tension between authenticity and commercialization, which map onto the three relations of *Tri Hita Karana* (*pawongan*, *palemahan*, *parahyangan*). These themes are best understood as an exploratory, qualitatively grounded reading of a thematically homogeneous corpus rather than as sharply bounded clusters. Contrary to an assumption of unshaken continuity, post-pandemic reviews were significantly less positive than pre-pandemic reviews; we therefore describe the village's heritage reception as broadly favourable but not immune to the pressures of tourism recovery, and we do not claim that cultural identity was empirically shown to be resilient, a broader construct not directly measured here.

From a management standpoint, protecting the communal and religious life-world that distinguishes Penglipuran, the hospitality (*pawongan*) and the ceremonial-architectural order (*parahyangan*) that visitors value most, may matter more than expanding commercial retail within traditional compounds, while the bamboo forest (*palemahan*) remains an under-used natural-cultural asset. Communicating the village's living-religious character before arrival may also help calibrate visitor expectations and ease the authenticity-commercialization friction concentrated in the most critical theme. Several limitations apply: TripAdvisor over-represents internationally connected travellers and under-represents domestic visitors; automated translation may introduce lexical noise in sentiment scoring; and the thematic solution is exploratory and sample-sensitive. Future research should combine user-generated-content analysis with multilingual visitor surveys and resident perspectives, and should examine the post-pandemic sentiment decline directly, to build a fuller, community-centred account of how religious-cultural meaning is sustained under tourism pressure. In addition, the automated dimensional classifier was developed on the full manually labelled set rather than evaluated on a separate held-out gold set, so its predictive accuracy was not formally quantified; future work should validate it on an independent sample.

Acknowledgement

The authors wish to thank the management of Penglipuran Village and all review contributors whose publicly available data made this study possible.

Author Contribution

Conceptualisation, I.M.W.; Methodology, I.M.W.; Software, I.M.W.; Formal Analysis, I.M.W.; Data Curation, I.M.W.; Writing - Original Draft Preparation, I.M.W., I.N.Y.S., I.G.A.R.P.D.; Writing - Review & Editing, I.M.W.; Visualisation, I.M.W. All authors have read and agreed to the published version of the manuscript.

Funding

This research received no external funding.

Ethical Review Board Statement

This study analysed publicly available user-generated content and did not involve interaction with human participants; under applicable national regulations, formal ethical approval was therefore not required. Public availability does not, however, remove all ethical obligations. We accordingly applied data minimization by removing reviewer identifiers from the analytical dataset, report no information that could identify individual reviewers, present illustrative reviews only in paraphrased and anonymized form rather than as verbatim quotations, and processed all data in accordance with the platform's terms of service. No sensitive personal data were collected or retained.

Informed Consent Statement

Not available. The study used publicly available, anonymized user-generated content and did not involve direct human participants; reviewer identifiers were used only transiently for de-duplication and were not retained.

Data Availability Statement

The original review dataset was collected from TripAdvisor (<https://www.tripadvisor.com>). The processed analytical dataset, including sentiment scores and dimensional classifications, is available from the corresponding author upon reasonable request, subject to the platform's terms of service.

Conflicts of Interest

The authors declare no conflict of interest.

Declaration of Generative AI Use

During the preparation of this manuscript, the authors used Claude (Anthropic) to assist with drafting, structural organization, language refinement, and parts of the computational analysis workflow (sentiment scoring, topic modelling, and statistical testing). All AI-assisted content and analyses were reviewed, verified, and edited by the authors, who take full responsibility for the final content.

Open Access

©2026. The authors. This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original authors and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third-party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit: <http://creativecommons.org/licenses/by/4.0/>

References

- Adi, I. N. R., Utama, M. S., Budhi, M. K. S., & Purbadharmaja, I. B. P. (2017). The role of government in community based tourism and sustainable tourism development at Penglipuran traditional Village-Bali. *IOSR Journal of Humanities and Social Science*, 22(06), 15–20. <https://doi.org/10.9790/0837-2206131520>
- Alfaritzi, I., Abirawa, B., Javid Ahmedthian, S., Wisnubroto, M. S., Farid, F., & Untoro, M. C. (2026). The application of multinomial naive bayes for sentiment analysis of cultural tourism reviews: A case study of Borobudur and Prambanan Temples. *Antivirus: Jurnal Ilmiah Teknik Informatika*, 20(1), 40–47. <https://doi.org/10.35457/dj70zr92>

- Alsharhan, A. M., Almansoori, H. R., Salloum, S., & Shaalan, K. (2022). *Three mars missions from three countries: Multilingual sentiment analysis using VADER* (pp. 371–387). https://doi.org/10.1007/978-3-031-03918-8_32
- Bigne, E., Fuentes-Medina, M. L., & Morini-Marrero, S. (2020). Memorable tourist experiences versus ordinary tourist experiences analysed through user-generated content. *Journal of Hospitality and Tourism Management*, 45, 309–318. <https://doi.org/10.1016/j.jhtm.2020.08.019>
- Blei, D. M., Ng, A. Y., & Jordan, E. (2003). Latent dirichlet allocation. *Journal of Machine Learning Research*, 3. <https://www.jmlr.org/papers/volume3/blei03a/blei03a.pdf>
- Boutsidis, C., & Gallopoulos, E. (2008). SVD based initialization: A head start for nonnegative matrix factorization. *Pattern Recognition*, 41(4), 1350–1362. <https://doi.org/10.1016/j.patcog.2007.09.010>
- Filieri, R., Acikgoz, F., Ndou, V., & Dwivedi, Y. (2021). Is TripAdvisor still relevant? The influence of review credibility, review usefulness, and ease of use on consumers' continuance intention. *International Journal of Contemporary Hospitality Management*, 33(1), 199–223. <https://doi.org/10.1108/IJCHM-05-2020-0402>
- Huete-Alcocer, N., Martinez-Ruiz, M. P., López-Ruiz, V. R., & Izquiedo-Yusta, A. (2019). Archeological tourist destination image formation: Influence of information sources on the cognitive, affective and unique image. *Frontiers in Psychology*, 10. <https://doi.org/10.3389/fpsyg.2019.02382>
- Hutto, C., & Gilbert, E. (2014). VADER: A parsimonious rule-based model for sentiment analysis of social media text. *Proceedings of the International AAAI Conference on Web and Social Media*, 8(1), 216–225. <https://doi.org/10.1609/icwsm.v8i1.14550>
- Islam, M. N. (2025). Visitors' sentiment in archeological heritage sites: A study based on TripAdvisor reviews. *International Journal of Hospitality & Tourism Administration*, 26(1), 90–111. <https://doi.org/10.1080/15256480.2023.2242343>
- Jia, S. (Sixue). (2020). Motivation and satisfaction of Chinese and U.S. tourists in restaurants: A cross-cultural text mining of online reviews. *Tourism Management*, 78, 104071. <https://doi.org/10.1016/j.tourman.2019.104071>
- Kompas. (2023). *Desa Penglipuran Bali raih penghargaan desa wisata terbaik 2023 dari UNWTO*. <https://Travel.Kompas.Com/Read/2023/10/21/153959927/Desa-Penglipuran-Bali-Raih-Penghargaan-Desa-Wisata-Terbaik-2023-Dari-Unwto>.
- Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics*, 33(1), 159. <https://doi.org/10.2307/2529310>
- Lee, D. D., & Seung, H. S. (1999). Learning the parts of objects by non-negative matrix factorization. *Nature*, 401(6755), 788–791. <https://doi.org/10.1038/44565>
- Li, Y., He, Z., Li, Y., Huang, T., & Liu, Z. (2023). Keep it real: Assessing destination image congruence and its impact on tourist experience evaluations. *Tourism Management*, 97, 104736. <https://doi.org/10.1016/j.tourman.2023.104736>
- Loria, S. (2018). TextBlob: Simplified text processing. *Python Software Foundation*. <https://Textblob.Readthedocs.Io/En/Dev/>.
- Lu, W., & Stepchenkova, S. (2015). User-generated content as a research mode in tourism and hospitality applications: Topics, methods, and software. *Journal of Hospitality Marketing & Management*, 24(2), 119–154. <https://doi.org/10.1080/19368623.2014.907758>
- MacCannell, D. (1973). Staged authenticity: Arrangements of social space in tourist settings. *American Journal of Sociology*, 79(3), 589–603. <https://doi.org/10.1086/225585>
- Mehraliyev, F., Chan, I. C. C., & Kirilenko, A. P. (2022). Sentiment analysis in hospitality and tourism: a thematic and methodological review. *International Journal of Contemporary Hospitality Management*, 34(1), 46–77. <https://doi.org/10.1108/IJCHM-02-2021-0132>
- Nariman, D. (2025). Sentiment analysis of hotel reviews using lexicon-based methods: A comparative study of VADER and TextBlob. *International Conference on Broadband and Wireless Computing, Communication and Applications* (pp. 263–274). https://doi.org/10.1007/978-3-031-76452-3_25

- Pedregosa, F., Michel, V., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Vanderplas, J., Cournapeau, D., Varoquaux, G., Gramfort, A., Thirion, B., Dubourg, V., Passos, A., Brucher, M., Perrot, M., Duchesnay, & Édouard. (2011). Scikit-learn: Machine Learning in Python. *Journal of Machine Learning Research*, 12. <http://scikit-learn.sourceforge.net>.
- Peters, J. Hendrik, & Wardana, Wisnu. (2013). *Tri Hita Karana: The Spirit of Bali*. Kepustakaan Populer Gramedia.
- Pine, B. J., & Gilmore, J. H. (2011). *The Experience Economy*. Harvard Business Review Press.
- Richards, G. (2018). Cultural tourism: A review of recent research and trends. *Journal of Hospitality and Tourism Management*, 36, 12–21. <https://doi.org/10.1016/j.jhtm.2018.03.005>
- Salton, G., & Buckley, C. (1988). Term-weighting approaches in automatic text retrieval. *Information Processing & Management*, 24(5), 513–523. [https://doi.org/10.1016/0306-4573\(88\)90021-0](https://doi.org/10.1016/0306-4573(88)90021-0)
- Sutrisna, I. B. M., Saskara, I. A. N., Utama, M. S., & Setyari, N. P. W. (2020). Prospective analysis of sustainable tourism development in Penglipuran Village as a tourist village role model in Bali. *Technium Social Sciences Journal*, 13(1), 184–198. <https://techniumscience.com/index.php/socialsciences/article/view/1791>
- Taecharungroj, V., & Mathayomchan, B. (2019). Analysing TripAdvisor reviews of tourist attractions in Phuket, Thailand. *Tourism Management*, 75, 550–568. <https://doi.org/10.1016/j.tourman.2019.06.020>
- United Nations World Tourism Organization. (2019). *UNWTO Tourism Definitions / Définitions du tourisme de l'OMT / Definiciones de turismo de la OMT*. World Tourism Organization (UNWTO). <https://doi.org/10.18111/9789284420858>
- Wang, C., Wang, Y., Edelman, J. R., & Zhou, J. (2024). Tourism commercialisation and the frontstage-backstage metaphor in intangible cultural heritage tourism. *Tourist Studies*, 24(3), 246–265. <https://doi.org/10.1177/14687976241251514>
- Wang, L., & Kirilenko, A. P. (2021). Do tourists from different countries interpret travel experience with the same feeling? Sentiment analysis of TripAdvisor reviews. *Information and Communication Technologies in Tourism 2021* (pp. 294–301). Springer International Publishing. https://doi.org/10.1007/978-3-030-65785-7_27
- Wang, N. (1999). Rethinking authenticity in tourism experience. *Annals of Tourism Research*, 26(2), 349–370. [https://doi.org/10.1016/S0160-7383\(98\)00103-0](https://doi.org/10.1016/S0160-7383(98)00103-0)
- Zhang, T., Yin, P., & Peng, Y. (2021). Effect of commercialization on tourists' perceived authenticity and satisfaction in the cultural heritage tourism context: Case study of langzhong ancient city. *Sustainability*, 13(12), 6847. <https://doi.org/10.3390/su13126847>
- Zhang, Y., & Lee, T. J. (2022). Alienation and authenticity in intangible cultural heritage tourism production. *International Journal of Tourism Research*, 24(1), 18–32. <https://doi.org/10.1002/jtr.2478>

Biographies of Authors

I Made Widiantara, is a permanent lecturer in the Business Administration Study Program, Department of Business Administration, at Politeknik Negeri Bali, Denpasar, Indonesia. His research interests encompass business communication, cultural tourism, public relations, and digital media. Research outputs have been published in national and international peer-reviewed journals and contributed book chapters.

- Email: madewidiantara@pnb.ac.id
- ORCID: 0000-0002-8423-6955
- Web of Science ResearcherID: ABH-1976-2020
- Scopus Author ID: 57212381122
- Homepage: <https://sinta.kemdiktisaintek.go.id/authors/profile/6034612>

I Nyoman Yoga Segara is a Professor of Anthropology at Universitas Hindu Negeri I Gusti Bagus Sugriwa, Denpasar, Indonesia. His scholarly interests include religious anthropology, Hindu philosophy, and cultural studies. Research outputs have been published extensively in national and international peer-reviewed journals.

- Email: yogasegara@uhnsugriwa.ac.id
- ORCID: 0000-0002-2617-483X
- Web of Science ResearcherID: N/A
- Scopus Author ID: N/A
- Homepage: <https://sinta.kemdiktisaintek.go.id/authors/profile/6197444>

I Gusti Ayu Ratna Pramesti Dasih, is a Lecturer in Communication Studies at Universitas Hindu Negeri (UHN) I Gusti Bagus Sugriwa Denpasar, Indonesia. Her research interests encompass intercultural communication, digital communication strategies, and sustainable tourism marketing in the digital era.

- Email: igustiayuratnapramesti@gmail.com
- ORCID: 0009-0002-9370-093X
- Web of Science ResearcherID: N/A
- Scopus Author ID: N/A
- Homepage: <https://sinta.kemdiktisaintek.go.id/authors/profile/6197675>