

Article

Rethinking Hospitality Performance Through Transformative Experience: A Narrative-Based DEA Framework for Experiential Evaluation in a Climate-Constrained Context

Maciej Kozłowski ^{1,*}  and Jerzy Korzeniewski ² 

¹ Department of Institutional Economics and Microeconomics, Faculty of Economics and Sociology, University of Lodz, Rewolucji Street 1905r. no 41, 90-255 Lodz, Poland

² Department of Demography, Faculty of Economics and Sociology, University of Lodz, Rewolucji Street 1905r. no 41, 90-255 Lodz, Poland; jerzy.korzeniewski@uni.lodz.pl

* Correspondence: maciej.kozlowski@uni.lodz.pl; Tel.: +48-660-665-218

Abstract

This study critically examines hospitality performance evaluation practices by proposing a narrative-based quantitative framework grounded in online reviews. While tourism experiences, particularly in transformative contexts, are understood as subjective and meaning-oriented, empirical evaluation remains reliant on standardized, growth-oriented indicators such as star ratings and satisfaction scores. To address this disconnect, on-line reviews are conceptualized not as post hoc satisfaction measures but as narrative expressions of evaluation. Drawing on sentiment analysis of narratives from eleven hotels, evaluations are operationalized through indicators and incorporated into a DEA framework. Efficiency is reframed as experiential conversion capacity—the ability of hospitality providers to transform material and organizational conditions into experiences perceived as meaningful by guests. Two aggregation configurations and a super-efficiency extension are applied to examine robustness and differentiation. The findings reveal divergence between narrative-based evaluations and star classifications, suggesting that rating systems fail to capture dimensions of meaning and affective resonance. Notably, a lower-category hotel emerges as the strongest in experiential positioning, challenging assumptions linking quality, classification, and value. From a sustainability perspective, the study contributes to critiques of tourism evaluation and supports post-growth, sufficiency-oriented approaches. Methodologically, it demonstrates how techniques can be repurposed as interpretive tools when grounded in narrative data.

Keywords: hospitality performance; transformative tourism; online consumer reviews; sentiment analysis; data envelopment analysis (DEA); experiential evaluation; sustainability; customer experience



Academic Editor: Chung-Jen Wang

Received: 15 May 2026

Revised: 23 June 2026

Accepted: 1 July 2026

Published: 6 July 2026

Copyright: © 2026 by the authors.

Licensee MDPI, Basel, Switzerland.

This article is an open access article distributed under the terms and conditions of the [Creative Commons Attribution \(CC BY\)](https://creativecommons.org/licenses/by/4.0/) license.

1. Introduction

Despite growing calls to rethink tourism in the face of an accelerating climate emergency, hospitality performance continues to be assessed primarily through standardized metrics that privilege efficiency, growth, and formal quality signals. Such indicators, including star classifications and aggregate ratings, shape not only how tourism providers are evaluated but also which forms of experience are legitimized as valuable. Under conditions of increasing ecological constraint, this growth-oriented and comparative logic renders certain experiential narratives visible and legitimate, while marginalizing others that are

less easily standardized or scaled. However, it remains unclear whether these evaluation systems adequately capture how tourism experiences are actually lived, narrated, and made meaningful by travelers themselves. At the same time, broader sustainability debates have intensified critical reflection on whether tourism evaluation systems centered on growth, benchmarking, and competitive positioning remain adequate under conditions of ecological constraint and post-growth transition.

This tension is particularly visible within the emerging field of transformative tourism, which emphasizes self-reflection, meaning-making, and experiential change rather than consumption or escape. While transformative tourism scholarship has advanced rich conceptual accounts of personal and relational transformation, empirical evaluation practices remain largely anchored in instrumental and technocratic indicators. As a result, experiences that tourists describe as emotionally meaningful, personally significant, or reflective may remain only partially visible within dominant performance frameworks, raising broader questions about which dimensions of tourism experience contemporary evaluation systems are able to recognize.

Against this background, scholarship increasingly highlights the importance of narrative forms of evaluation. User-generated online reviews are not merely repositories of customer satisfaction data but also spaces in which travelers retrospectively interpret and narratively organize their experiences. Such narratives frequently combine emotional, symbolic, and contextual dimensions of evaluation that extend beyond discrete service assessments. Consequently, online reviews may provide access not only to attribute-level evaluations but also to broader forms of experiential judgment grounded in lived experience and retrospective sensemaking.

At the same time, quantitative tourism evaluation remains dominated by operational and efficiency-oriented approaches. Data Envelopment Analysis (DEA), for example, has traditionally been used to examine relative efficiency in tourism and hospitality contexts. More recent methodological discussions, however, suggest that DEA may also function as a comparative interpretive framework rather than solely as a managerial optimization tool. When combined with narrative-derived evaluative indicators, quantitative techniques may therefore support the examination of how experiential value is comparatively articulated across hospitality providers without reducing tourism experience to standardized performance metrics.

Against this background, the central research problem addressed in this study concerns the growing disconnect between standardized hospitality evaluation systems and lived, narratively constructed tourism experiences. The study investigates whether experiential evaluations derived from guest narratives diverge from formal hotel star classifications, and whether Data Envelopment Analysis (DEA) can be repurposed as a comparative interpretive framework for identifying such divergences.

Accordingly, the study pursues three interrelated objectives. First, it develops a narrative-based quantitative framework that treats online reviews as expressions of experiential evaluation rather than conventional satisfaction signals. Second, it empirically explores how narrative-derived experiential evaluations vary across eleven urban hotels characterized by different formal classifications. Third, it examines the extent to which DEA may support the comparative interpretation of experiential conversion capacity, understood here as the ability to translate hospitality conditions into experiences positively evaluated and perceived as meaningful by guests.

Importantly, the study does not assume that positive sentiment directly indicates transformative experience in a strong psychological or existential sense. Rather, transformative tourism serves primarily as a conceptual lens for interpreting how experiential meaning may emerge beyond standardized satisfaction metrics. Accordingly, sentiment

analysis captures evaluative orientation and holistic experiential judgment rather than transformation itself.

Empirically, the analysis is based on user-generated online reviews collected for eleven urban hotels and processed through sentiment-based narrative indicators integrated into DEA configurations. The study is positioned primarily as a methodological and conceptual contribution illustrated through an exploratory empirical case rather than as a statistically generalizable hospitality performance study.

The contribution of the paper is fourfold. Theoretically, it contributes to transformative tourism scholarship by foregrounding narrative meaning and lived experience as evaluative dimensions of tourism performance. Methodologically, it demonstrates how sentiment analysis and DEA may be combined within a narrative-oriented comparative framework. Empirically, it reveals divergences between formal hotel classifications and guest-derived experiential evaluations. Practically, it offers an alternative perspective for critically reflecting on tourism evaluation under conditions of sustainability transition and ecological constraint.

2. Literature Review

2.1. *Transformative Tourism and the Centrality of Meaning*

Transformative tourism has gained prominence as a response to increasingly instrumental and consumption-driven models of travel, foregrounding personal change, self-reflection, and meaning-making over hedonic satisfaction or momentary pleasure. Early conceptualizations of transformative tourism positioned travel as a potential catalyst for identity work and existential reflection rather than leisure consumption [1,2]. More recent syntheses emphasize that transformation is not a standardized outcome but a deeply subjective and relational process [3].

Across theoretical perspectives, scholars converge on the argument that transformative processes unfold through travelers' interpretations of their experiences rather than through externally observable effects. Transformation is mediated through narrative sensemaking, emotional resonance, and post-travel reflection, emphasizing subjectivity over uniform benchmarks or external criteria of success [4]. This emphasis on subjectivity challenges evaluation approaches relying on standardized measures of outcomes.

Recent work further complicates early assumptions that transformative tourism occurs primarily in extraordinary or explicitly therapeutic settings. Empirical studies show that subtle, incremental, and everyday forms of transformation emerge in ordinary tourism contexts, including urban stays and short-term hospitality experiences [5,6]. Such micro-transformations are often expressed through feelings of coherence, resonance, or reassessment rather than dramatic life changes [7].

Taken together, this literature positions lived experience and narrative interpretation as central analytical dimensions of transformative tourism. If transformation is articulated and makes sense through personal narratives, then empirical access to transformation requires methods capable of engaging with how tourists describe, evaluate, and remember their experiences.

2.2. *Evaluating Tourism Experience: The Limits of Standardized Performance Metrics*

Despite the conceptual shift toward meaning and subjectivity in tourism theory, hospitality evaluation practices remain dominated by standardized performance metrics. Star classifications, customer satisfaction indices, and operational efficiency measures continue to structure both academic research and managerial practice [8,9]. These metrics privilege comparability and control, offering simplified representations of quality that facilitate benchmarking while obscuring experiential complexity.

Critical tourism scholarship highlights that such indicators are poorly suited to reveal important dimensions of experience that travelers themselves find meaningful. Managerial evaluation tools tend to reduce experience to discrete and measurable attributes, neglecting affective coherence, personal relevance, and reflective depth, thereby marginalizing experiential qualities that resist formalization [10,11].

Importantly, evaluation systems operate not merely as neutral measurement instruments but as epistemic devices that shape what counts as valuable tourism experience. Quantification practices actively construct social realities by privileging certain forms of knowledge while rendering others invisible [12]. Applied to tourism, this implies that standardized metrics do not simply represent experience but participate in defining legitimate forms of experiential value.

In the context of climate emergency and sustainability debates, reliance on efficiency-driven indicators raises further concerns. Metric-focused approaches risk reinforcing performative sustainability and growth-oriented imaginaries, even as they claim to promote responsible tourism development [13,14]. Consequently, modest, sufficiency-oriented, or experientially meaningful forms of tourism may be systematically undervalued precisely because they escape conventional performance frameworks.

2.3. Online Reviews as Narrative Expressions of Lived Experience

User-generated online reviews have become a prominent data source in tourism research, providing large-scale access to tourists' unsolicited accounts of experience [15,16]. Traditionally, reviews have been operationalized as indicators of satisfaction, service quality, or reputation, often aggregated into numerical ratings for predictive or comparative purposes [17]. An alternative stream of research conceptualizes online reviews as narrative practices through which travelers retrospectively construct meaning. Reviews function as sensemaking devices that embed emotions, expectations, and personal values within coherent experiential accounts, reflecting how tourists interpret and remember their experiences rather than merely evaluating discrete service attributes [18,19].

Sentiment expressed in reviews should therefore be understood as a holistic experiential judgment rather than a simple affective reaction. Emotional valence often encapsulates broader assessments of authenticity, resonance, and symbolic value that extend beyond isolated service features [20]. Narrative data thus provides a unique empirical window into lived experience at scale, combining interpretive depth with analytical reach.

Treating online reviews as narrative expressions rather than performance signals aligns closely with transformative tourism theory. It enables empirical engagement with meaning-making processes while preserving the breadth and comparability associated with quantitative research designs.

2.4. Quantitative Approaches to Narrative-Based Experience Evaluation

Recent methodological advances indicate that quantitative approaches can be deployed in ways that do not strip narrative data of experiential meaning. When theoretically grounded, computational tools may support critical inquiry rather than managerial optimization [21]. Sentiment analysis, in particular, allows systematic examination of evaluative tone across large corpora of narrative data derived from user-generated reviews [22].

While sentiment analysis is often criticized for oversimplification, several studies suggest that it can serve as a proxy for overall experiential evaluation when interpreted cautiously [23]. Rather than measuring discrete emotions, sentiment scores reflect aggregated judgments shaped by context, expectations, and personal relevance.

Data Envelopment Analysis (DEA), traditionally applied to operational efficiency, has been widely used in tourism and hospitality research [24,25]. More recent contributions

indicate that DEA can be conceptually reinterpreted beyond cost-efficiency logic, functioning instead as a comparative relational framework for examining how different outcomes emerge relative to one another [26].

When narrative-derived sentiment indicators are integrated into DEA models, quantitative methods can operate as interpretive tools rather than optimization devices. Such configurations enable the identification of discrepancies between formal classifications and lived experiential value, revealing structural tensions at the heart of contemporary tourism performance evaluation [27].

2.5. Conceptual Distinction Between Satisfaction, Experience, and Transformation

Importantly, the present study does not assume that positive sentiment or favorable evaluations necessarily indicate transformative tourism experiences in a strong psychological or existential sense. A distinction must be made between attribute-specific satisfaction, holistic experiential evaluation, memorable experience, and transformative experience involving durable self-reflection or identity-related change. While online reviews may contain traces of meaning-making, emotional resonance, or reflective interpretation, sentiment analysis primarily captures evaluative orientation rather than transformation itself.

Accordingly, the proposed framework should be understood as examining narrative-based experiential evaluation that may, in certain cases, align with dimensions emphasized in transformative tourism scholarship, rather than as a direct measurement of personal transformation. The transformative tourism perspective therefore serves primarily as a conceptual lens for interpreting experiential meaning rather than as an empirically verified outcome of the analyzed hotel stays.

2.6. Research Positioning and Conceptual Gap

Taken together, existing literature reveals a persistent disconnect between transformative tourism theory—centered on meaning, subjectivity, and lived experience—and dominant evaluation practices grounded in standardized performance metrics. Although narrative data and quantitative tools have been examined separately, they are often mobilized in ways that reproduce managerial paradigms rather than critically interrogating them [28].

There remains limited empirical work that explicitly integrates narrative-derived indicators into comparative evaluation frameworks to reassess hospitality performance from a transformative tourism perspective. In particular, few studies reconceptualize efficiency as an experiential relation grounded in tourists' own evaluations rather than in institutional benchmarks or formal classification systems.

Against this background, the present study addresses this gap by proposing a narrative-based quantitative framework that reframes efficiency as experiential conversion capacity. By integrating sentiment analysis of online reviews into a DEA-based comparison of hospitality providers, the study contributes to ongoing debates on how tourism experience can be empirically evaluated in ways aligned with sustainability-oriented and post-growth imaginaries.

3. Methodology: A Narrative-Based Quantitative Framework

3.1. Research Design and Dataset

This study adopts a narrative-based quantitative research design that integrates sentiment analysis of user-generated online reviews with Data Envelopment Analysis (DEA) as a comparative analytical approach. Rather than operationalizing guest feedback as conventional customer satisfaction, the approach conceptualizes online reviews as narrative expressions of experiential judgment, reflecting how travelers interpret, remember, and

assign meaning to their hospitality experiences. This design aligns with transformative tourism scholarship by prioritizing lived experience and retrospective sensemaking over standardized performance indicators.

The empirical dataset consists of user-generated online reviews collected for hotels and accommodation facilities located in Poland. Reviews were obtained from Google Maps using the Outscraper extraction tool in August 2024 (<https://outscraper.com>). Outscraper was used solely as a technical extraction interface for publicly accessible review content. The dataset covered reviews published between 2017 and 2022, based on the review_datetime metadata available in the exported database.

From the full dataset, a subset of eleven hotels was selected for further analysis. To ensure transparency, descriptive statistics of the analyzed sample, including positive and negative sentiment assignments across the DEA narrative dimensions (General, Premises, Staff, Finance, and Resting), are provided in Appendix A.3. Neutral sentiment was not explicitly analyzed, as the study focuses on clearly polarized sentiment categories. Sentiment classification was based on review ratings, where ratings of 1–2 were classified as negative and 4–5 as positive. Given the attribute-based structure of the sentiment analysis, the reported statistics reflect the frequency of sentiment assignments underlying the DEA narrative dimensions rather than the total number of reviews analyzed at the document level.

The extraction process relied on publicly accessible business review data indexed within Google Maps. Searches were conducted using hotel-related location queries and accommodation categories available through the platform interface. The resulting dataset included review text, review scores, review dates, and selected metadata associated with hospitality establishments.

The analyzed database included hotel properties from multiple Polish regions and contained information on hotel characteristics, review distributions, textual user feedback, accommodation types, and selected operational attributes. The analyzed corpus covered hotels, serviced apartments, guesthouses, resorts, villas, hostels, and other lodging establishments operating in urban, coastal, mountain, and regional tourism destinations across Poland.

The number of reviews retrieved for each property was determined by the volume of publicly available reviews indexed by the platform and returned through the extraction process. Consequently, review counts varied substantially across establishments and were not artificially balanced or standardized. The extraction procedure was intended to capture naturally occurring review distributions rather than to construct a statistically representative sample. Review counts ranged from single-review objects to facilities with more than 12,000 reviews. Most properties achieved average ratings between 3.8 and 4.7 on the Google review scale. While review counts are reported to characterize the dataset at the data collection stage, the subsequent analysis is conducted at the attribute level, based on narrative-derived sentiment assignments.

Only textual reviews written in Polish and English were included in the analysis, as these languages constitute the dominant share of the dataset. No machine translation was performed.

Following extraction, the dataset was manually screened to remove duplicate entries, reviews consisting exclusively of numerical ratings, and extremely short or otherwise non-informative entries. Neutral reviews were retained in the dataset and incorporated into the narrative classification procedure described below.

The hotels were selected purposively to include properties representing different formal star classifications and diverse tourism settings within Poland, while maintaining overall comparability as hospitality accommodation providers. Hotel categories and op-

erational characteristics were verified using publicly available classification information, booking platforms, hotel websites, and Google Maps business profiles.

No personal or identifying user data was collected, stored, or processed beyond publicly accessible review content. The analysis focused exclusively on aggregated narrative material used for research purposes. Due to platform and ethical considerations, the raw review corpus is not publicly redistributed; however, aggregated analytical data, methodological procedures, and review-classification guidelines can be made available upon reasonable request.

3.2. Attribute Definitions and Grouping

The identified attributes are not treated as objective service features, but as narrative anchors through which guests articulate experiential meaning. The initial review-level analysis identified seventeen narrative terms associated with hospitality experience evaluation: location, building, room, bathroom, parking, lifts, breakfast, entertainment, prices, payment-related issues, payment methods, staff, cleanliness, quietness, heating, air conditioning, and internet quality.

For analytical purposes, several closely related terms were subsequently merged into broader evaluative attributes to reduce fragmentation, improve model stability, and enhance the discriminatory capacity of the DEA framework within the relatively small sample of hotels analyzed. In particular, room and bathroom references were combined into a single “rooms and bathrooms” category, while prices, payment-related issues, and payment methods were aggregated into a broader financial factors category. The term “entertainments” refers to recreational amenities and entertainment-related services mentioned by guests within hotel reviews.

The resulting attributes were reorganized into four conceptually coherent groups. The first group, premises and facilities, includes location, rooms and bathrooms, the building, parking, and lifts, capturing the physical environment and spatial accessibility. The second group, staff and services, covers staff behavior, breakfast quality, and entertainment-related amenities. The third group aggregates financial factors associated with pricing and payment-related experiences. The fourth group, living and resting conditions, includes quietness, heating, cleanliness, air conditioning, and internet quality. This aggregation reduces dimensionality and mitigates a known limitation of DEA, whereby models with too many outputs often lose discriminatory power.

3.3. Treatment of General, Specific, and Mixed Reviews

General reviews express overall sentiment toward a hotel experience without referencing specific attributes, yet they constitute a substantial and meaningful portion of user-generated content. Excluding such reviews would result in a systematic loss of experiential information, particularly with respect to holistic emotional judgments central to narrative sensemaking. Specific reviews, in contrast, explicitly reference individual attributes and articulate targeted experiential evaluations. Mixed reviews combine both forms, expressing an overall sentiment while concurrently highlighting particular experiential dimensions.

To accommodate these different narrative structures, the study applies a weighting procedure that allows general sentiment to be proportionally incorporated into attribute-level evaluation without privileging explicitly mentioned attributes at the expense of holistic experience. In this context, positive general statements are assigned a weight of +0.65 and negative ones a weight of −0.65, which are distributed across all attributes. In mixed reviews, attributes explicitly referenced by the reviewer receive a full weight of ± 1 , while non-mentioned attributes retain the general sentiment weight of ± 0.65 .

The weighting scheme is informed by Cronbach's alpha values exceeding the commonly accepted reliability threshold of $\alpha \geq 0.7$, with 0.65 representing the minimum weight required to maintain internal consistency across aggregated attribute-level indicators [29]. Importantly, the use of Cronbach's alpha in this context does not aim to establish psychometric scale validity but serves as a consistency-preserving heuristic that enables proportional integration of general narrative sentiment into attribute-level indicators [30].

Importantly, the selected weighting parameter should be understood as a heuristic and exploratory calibration rather than as a statistically optimal value. The procedure does not assume that general evaluative statements correspond equally to all individual attributes. Instead, proportional sentiment distribution serves as an interpretive mechanism intended to preserve holistic experiential orientation within the aggregated narrative framework.

To assess the robustness of the weighting parameter, sensitivity analysis across a range of weights was conducted, indicating that the resulting rankings remained stable across the tested values.

The classification of reviews as general, specific, or mixed followed a rule-based procedure grounded in the presence or absence of explicit attribute references within the review text. Reviews expressing only overall evaluative orientation without identifiable attribute-level references were treated as general reviews, whereas reviews containing explicit references to particular hospitality attributes were classified as specific or mixed depending on whether both global and attribute-specific evaluations co-occurred. The algorithm processes reviews at the sentence and descriptor level, allowing different attributes within the same review to receive different evaluative assignments. Where conflicting evaluations of the same attribute appeared within a single review, the algorithm assigned sentiment contextually at the descriptor level rather than imposing a single document-wide polarity. Conditional constructions and negation expressions were treated separately within the rule-based procedure to reduce incorrect sentiment propagation across attributes.

This approach ensures that general reviews are neither ignored nor overemphasized, while preserving the additional informational value conveyed by attribute-specific evaluations. As a result, the treatment of general, specific, and mixed reviews maintains narrative coherence and supports subsequent comparative analysis without reducing experiential expression to purely attribute-based signals.

3.4. Experiential Evaluation Measures (See: Appendix A.1)

Having defined how sentiment loads are assigned to attributes, the next step is to transform these values into quantitative satisfaction measures compatible with DEA. To generate DEA-compatible outputs, quantitative satisfaction measures were defined based on sentiment data. Two alternatives were considered: the percentage of positive reviews among all reviews and the ratio of positive to negative reviews. Although intuitive, the percentage measure compresses variability due to its bounded nature, limiting the model's discriminatory power. In contrast, the ratio of positive to negative reviews provides a more sensitive representation of sentiment by explicitly capturing the balance between favorable and unfavorable opinions, thereby highlighting differences that may be obscured by percentage-based metrics. This ratio also mitigates the risk of artificially inflating satisfaction scores in cases where a hotel receives a high number of reviews, but only a small proportion of them express dissatisfaction. Moreover, because the ratio expands the scale of variation between units, it enhances the discriminatory capacity of the DEA framework by facilitating differentiation between hotels without producing flat or uninformative outcomes.

However, ratio-based measures may be sensitive to denominator effects, particularly when the number of negative reviews is small, and should therefore be interpreted with

caution within the exploratory framework adopted in this study. Despite these limitations, the ratio-based measure was used as the primary satisfaction metric in the DEA configurations. In addition, a smoothed ratio variant using Laplace correction $((\text{pos} + 1)/(\text{neg} + 1))$ was tested as a robustness check to address potential denominator sensitivity.

3.5. DEA Model Specification

Data Envelopment Analysis (DEA) is employed as the comparative analytical framework used to examine how narrative-based experiential evaluations differ across hotels. In line with the theoretical positioning of the study, DEA is not applied as an optimization or performance improvement tool but as a relational method that enables the identification of relative experiential positioning among comparable hospitality providers.

Drawing on DEA-based comparative logic, the analysis adopts an output-oriented framework, as the objective is to compare hotels based on the degree to which positive experiential narratives emerge from comparable contextual conditions rather than to minimize resource use. The conceptual rationale for DEA, together with its formal foundations, is provided in Appendix A.2 to avoid unnecessary repetition in the main text.

The empirical analysis is based on the CCR (Charnes–Cooper–Rhodes) model [25], which assumes constant returns to scale and enables the comparison of decision-making units (DMUs) based on multiple outputs. In the present context, each hotel is treated as a DMU, while narrative-derived experiential evaluation measures (defined in Section 3.4) constitute the model outputs.

No explicit input variables are specified, as the study does not conceptualize hospitality experience as a conventional production process. Instead, DEA is employed here as a comparative and relational framework to examine differences in narrative-derived experiential evaluations across broadly comparable hospitality settings. Accordingly, the aim is not to model production efficiency but to provide a structured, comparative representation of narrative-based experiential evaluations across hotels.

The present configuration does not measure efficiency or conversion in a strict input–output sense. Rather, the DEA framework functions as an interpretive device for examining variation in experiential evaluations across comparable hospitality contexts. In this context, the notion of “experiential conversion capacity” should be understood heuristically and interpretively rather than as a direct measurement of observable resource transformation processes.

The DEA model is formulated as follows. The experiential evaluation score for the hotel k , denoted as LCS_k , is obtained by solving the linear programming problem given in Equation (1):

$$\text{LCS}_k = \max \sum_{r=1}^s u_{rk} \text{Sc}_{rk} \quad (1)$$

subject to the constraints:

$$\sum_{r=1}^s u_{rk} \text{Sc}_{rj} \leq 1 \text{ for all } j = 1, \dots, n,$$

$$u_{rk} \geq 0 \text{ for all } r,$$

where Sc_{rj} denotes the experiential evaluation score associated with output r for hotel j , s is the number of output categories, and $n = 11$ is the number of hotels under analysis. This formulation corresponds to the reduced form of the output-oriented CCR model.

In applications involving relatively small sets of DMUs, it is common for several units to attain identical frontier scores. To further differentiate among these hotels, the Andersen–Petersen super-efficiency model is employed [26]. This extension allows efficient

DMUs to be ranked by temporarily excluding the evaluated unit from the reference set, thereby enabling finer-grained comparative interpretation of experiential positioning.

All DEA computations were implemented using the Solver module in Microsoft Excel, following standard DEA spreadsheet formulations. Because the study relies on standard DEA formulations, the proposed configurations may also be reproduced using widely available DEA packages implemented in statistical environments such as R (e.g., rDEA, deaR). The purpose of the DEA analysis in this study is not to prescribe optimal configurations or improvement paths, but to reveal relative patterns in how experiential value, as expressed through guest narratives, emerges across hotels. Accordingly, the resulting DEA scores should be interpreted as indicators of relative experiential positioning rather than as measures of operational efficiency in the conventional economic sense.

A key methodological consideration in DEA concerns the trade-off between the number of DMUs and the dimensionality of outputs. When an excessive number of output variables is included relative to the number of units, DEA models tend to lose discriminatory power. For this reason, alternative output configurations were examined based on attribute aggregation strategies described in Section 3.2. Specifically, the DEA was conducted under two configurations: (i) a three-group structure (premises, staff, living conditions) and (ii) a four-group structure (premises, staff, financial factors, living conditions). This sensitivity analysis enables assessment of the robustness of experiential comparisons to different levels of narrative aggregation. These configurations therefore do not eliminate the possibility of frontier saturation but provide partial differentiation while preserving the exploratory comparative logic of the framework.

Finally, supplementary analyses were conducted to contextualize DEA-derived rankings. Spearman rank correlations were used to examine associations between narrative-based experiential evaluation and formal hotel star ratings, while frontier units identified through the Andersen–Petersen model were examined separately to explore patterns among hotels exhibiting relatively strong narrative-derived experiential positioning.

3.6. Operational Structure of the Proposed Framework

This subsection outlines the operational structure through which the proposed narrative-based quantitative framework is implemented, clarifying how unstructured narrative data are systematically transformed into comparative experiential indicators without presenting a fully developed decision support system. The framework serves as a conceptual guide for operationalizing experiential conversion capacity, linking guest narratives to evaluative outcomes in a transparent and interpretable manner.

The framework consists of four analytically connected stages. The first stage concerns the acquisition and preparation of narrative data. Online hotel reviews are collected from publicly available platforms and preprocessed to establish coherent textual units. To reduce potential bias associated with language heterogeneity, the analysis was restricted to reviews in Polish and English, which constitute the dominant share of the dataset. While this approach improves consistency within the rule-based sentiment extraction procedure, residual bias related to cross-linguistic variation cannot be entirely excluded. Reviews are subsequently classified as general, specific, or mixed, reflecting different forms of experiential articulation. This classification does not impose an external evaluative structure but preserves how guests themselves narrate and prioritize their experiences.

The second stage focuses on narrative processing and sentiment extraction. Attribute-related expressions are identified within reviews and associated with sentiment polarity, enabling the reconstruction of experiential evaluations at the attribute level. Importantly, general reviews, although lacking explicit attribute references, are retained as expressions

of holistic experiential judgment. Their sentiment is proportionally incorporated into attribute-level indicators to preserve narrative coherence.

In the third stage, narrative-derived sentiment indicators are aggregated and transformed into quantitative measures suitable for comparative evaluation. At this point, Data Envelopment Analysis is employed not as an optimization device but as a relational analytical framework. DEA enables comparison across hotels by examining how narrative-based experiential outcomes emerge relative to one another, thereby identifying patterns of experiential conversion rather than absolute performance levels.

The final stage concerns interpretive use rather than automated decision-making. The resulting comparative indicators support reflective interpretation of experiential positioning across hotels, highlighting divergences between formal classifications and guest-evaluated experience. This stage does not prescribe managerial action but provides an analytical basis for understanding how experiential meaning is unevenly translated from underlying hospitality resources.

To support construct validity, two complementary validation procedures were conducted.

First, sentiment assignments were compared against rating-derived sentiment proxies (1–2 negative, 3 neutral, 4–5 positive), yielding an overall percentage agreement (PA) of 0.74.

Second, a manually coded subset of reviews ($n = 50$) was evaluated, producing Cohen's kappa of 0.66, indicating acceptable consistency for exploratory analysis.

These results provide empirical support for the reliability of the constructed narrative indicators.

In the manual coding procedure, each review was assigned a single dominant sentiment category (positive, neutral, or negative) based on its overall tone, and a primary attribute category reflecting the main focus of the review (e.g., service, food, room, or general experience). This simplification was used solely for validation purposes to ensure comparability with the rule-based procedure. The coding thus aimed to capture the dominant narrative signal expressed in each review.

Taken together, this operational structure demonstrates how narrative data can be systematically mobilized within a quantitative framework without reducing experience to managerial efficiency metrics. By foregrounding interpretation over optimization, the framework remains aligned with sustainability-oriented and transformative tourism perspectives, supporting critical reflection on how tourism experience is evaluated rather than merely improved for managerial purposes. Accordingly, the methodological design prioritizes interpretive transparency and comparative consistency over predictive optimization, reflecting the study's objective of examining how experiential value is narratively articulated rather than accurately predicted.

Several analytical limitations should be acknowledged in interpreting the results. Although narrative-derived sentiment indicators provide insight into evaluative orientation and holistic experiential judgment, the framework does not attempt to directly measure transformative processes, as it focuses on the comparative evaluation of narratively articulated experiential judgments.

Additional technical details concerning sentiment extraction procedures and formal DEA formulations are provided in Appendices A.1 and A.2.

4. Results

This section reports the empirical results of the text mining-based DEA evaluation carried out for eleven hotels. Appendix A.3 provides descriptive statistics including positive and negative sentiment assignments for each hotel analyzed across the narrative dimensions used in the DEA framework (General, Premises, Staff, Finance, and Resting). Two attribute grouping configurations were analyzed, followed by an Andersen–Petersen

super-efficiency assessment and a correlation analysis examining the relationship between sentiment-based performance measures and formal star classifications.

4.1. LCS Distributions Across DEA Setups

LCS distributions across DEA setups are interpreted here as comparative patterns of experiential evaluation rather than direct indicators of performance efficiency. The two DEA configurations produce distinct distributions of experiential evaluation scores. In Setup 1, which uses three attribute groups, LCS values range from 0.105 to 1.000, indicating substantial heterogeneity in how hotel experiences are positively and negatively articulated in guest narratives. In Setup 2, which includes the financial attribute group as a fourth output dimension, several hotels attain relative frontier positions within the DEA framework. The distribution becomes more uniform, and a greater number of hotels occupy the frontier, illustrating the effect of increased dimensionality on the model's discriminatory capacity. Despite these shifts, the overall pattern of relative experiential positioning remains broadly consistent across both setups. The descriptive statistics presented in Table 1 further illustrate the distribution of LCS values across the analyzed hotels and DEA configurations.

Table 1. Values of LCS for 11 hotels across both attribute configurations and the Andersen–Petersen model.

LCS	1_2★	1_3★	2_3★	3_3★	4_3★	5_3★	6_3★	1_4★	2_4★	3_4★	4_4★
setup1	0.640	0.283	0.105	1	0.950	0.492	0.608	0.670	0.371	1	0.990
setup2	1	0.414	0.887	0.507	0.84	0.663	1	1	0.909	1	1
A-P model	3.75	0.414	0.887	0.507	0.84	0.663	1.10	1.05	0.909	1.33	1.29

★ denotes the official hotel star classification. Source: own calculations.

Figure 1 visualizes these differences, showing that Setup 1 yields the largest variation, whereas Setup 2 produces more compact profiles. Across both configurations, narrative-derived experiential evaluation measures reveal patterns of experiential positioning that do not align fully with formal star classifications.

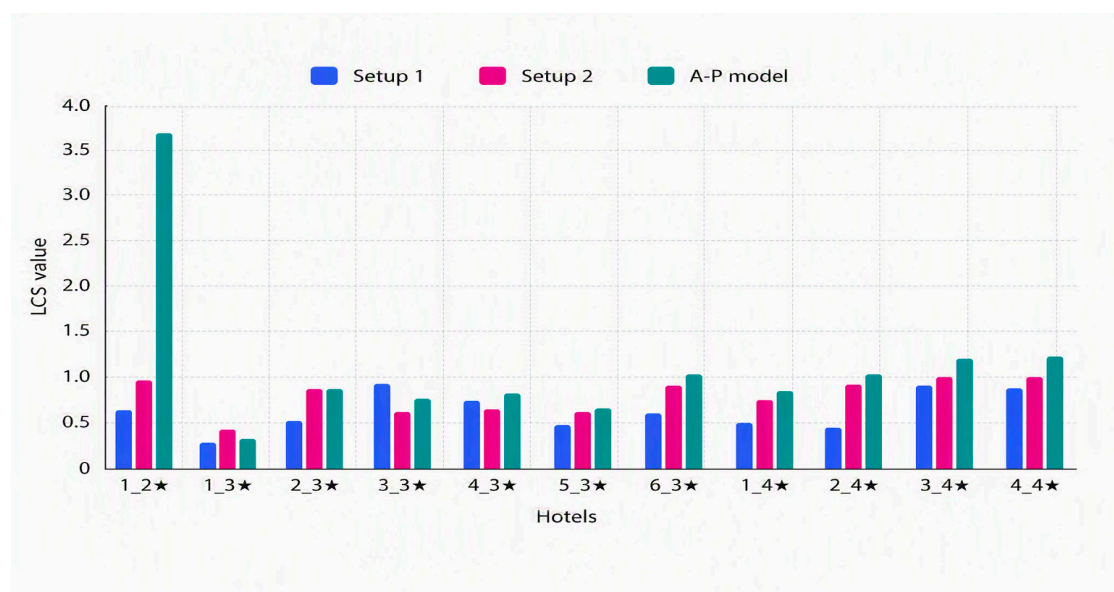


Figure 1. LCS values for 11 hotels across two DEA setups and the Andersen–Petersen model. Hotel labels include the official star classification (★). Source: own calculations.

Figure 1 highlights structural differences between the two setups. Setup 1 creates a more irregular distribution with clearer separation between hotels, whereas Setup 2

produces a more uniform pattern. Andersen–Petersen super-efficiency values extend the scale beyond 1.0, enabling deeper differentiation among frontier hotels. The figure also visually confirms inconsistencies between formal star classifications and narrative-derived experiential rankings.

4.2. Robustness Checks

To assess the stability and robustness of the results obtained, several additional analyses were conducted.

First, the sensitivity of hotel rankings to the general sentiment weight parameter was examined across a range of values ($w \in [0.3, 1.0]$). The results indicated high ranking stability throughout the tested interval, suggesting that the comparative positioning of hotels is not driven by a specific choice of the weighting parameter.

Second, alternative sentiment formulations were evaluated by comparing the baseline positive-to-negative ratio with a smoothed ratio specification using Laplace correction $((\text{pos} + 1)/(\text{neg} + 1))$. The resulting rankings and DEA scores remained directionally consistent across both specifications, indicating that the findings are not sensitive to potential denominator effects in the sentiment measures.

Third, robustness was assessed through comparison of alternative DEA aggregation structures. Both the three-group and four-group configurations produced broadly similar patterns of experiential positioning despite differences in discriminatory power, suggesting that the main conclusions are not dependent on a particular aggregation scheme.

Finally, additional exploratory sensitivity analyses were conducted to assess the robustness of DEA scores and rankings. The results suggested that the main ranking patterns remained broadly stable across alternative specifications and did not affect the substantive conclusions, providing additional support for the consistency of the proposed framework.

Overall, the robustness analyses suggest that the reported findings are structurally stable across alternative weighting assumptions, sentiment specifications, and DEA configurations. Although the exploratory character of the study and the limited number of analyzed units call for cautious interpretation, the consistency of results across robustness checks supports the reliability of the proposed narrative-based evaluation framework.

4.3. Results for Setup 1 (Three Attribute Groups)

Setup 1 evaluates hotels based on narrative-derived experiential evaluations aggregated across three domains: premises, staff, and living conditions. LCS values range from 0.105 to 1.000, indicating pronounced variation in how hotel experiences are articulated and evaluated by guests. Only one hotel attains a frontier position within the DEA framework.

Several mid-category hotels exhibit relatively low LCS values despite belonging to higher formal classifications, suggesting that narrative evaluations of experience do not necessarily correspond to expected quality levels implied by star ratings. The distribution reveals clear contrasts between hotels with more positively and more negatively articulated experiential profiles and provides insight into heterogeneity across experiential domains.

Overall, Setup 1 highlights substantial differentiation in experiential evaluations across hotels and offers a nuanced view of how experiences are narratively constructed within each attribute group. These results form the basis for subsequent comparison with the four-group configuration presented in the next subsection.

4.4. Results for Setup 2 (Four Attribute Groups)

Setup 2 introduces financial factors as a separate output group, increasing model dimensionality and altering the distribution of LCS values. The resulting experiential evaluation profiles are more compressed, with a larger number of hotels occupying frontier positions within the DEA framework.

Key observations include:

- Five hotels attain $LCS = 1.000$, indicating reduced discriminatory capacity under higher dimensionality.
- The relative ordering of experiential evaluations remains broadly consistent with Setup 1, despite the greater number of frontier units.
- The inclusion of financial narrative anchors yields more stable but more compressed experiential evaluation patterns.

4.5. Andersen–Petersen Super-Efficiency Results

The Andersen–Petersen model was applied to Setup 2 to further differentiate hotels occupying frontier positions in the standard CCR formulation. By allowing super-efficiency values to exceed 1.0, this extension enables finer comparative interpretation and reveals variation within the subset of hotels exhibiting similar frontier status. The results indicate substantial heterogeneity among the five frontier hotels identified in Setup 2. The two-star hotel (1_2★) attains the highest super-efficiency value (3.75), indicating the strongest relative experiential valuation within the frontier group. The remaining frontier hotels achieve more moderate super-efficiency values, ranging from 1.05 to 1.33, suggesting more comparable experiential positions.

Interestingly, the Andersen–Petersen ordering broadly aligns with formal star classifications, with the notable exception of the two-star hotel, which emerges as a particularly strong case of experiential conversion. This finding suggests that narrative-derived evaluations capture dimensions of experiential meaning that are insufficiently reflected in traditional star-rating systems. Overall, the super-efficiency analysis provides additional insight into experiential heterogeneity among hotels that appear similar under the standard DEA framework. Given the relatively small number of DMUs and the dimensionality of the model, the Andersen–Petersen super-efficiency results should be interpreted cautiously, particularly in cases of highly differentiated frontier scores.

4.6. Correlation Analysis

Spearman’s rank correlations were computed to examine the consistency between the two DEA configurations and to explore the relationship between narrative-derived experiential evaluations and formal hotel star classifications. The detailed results are reported in Table 2.

Table 2. Spearman’s rank correlation coefficients between experiential evaluation rankings, DEA setups, and formal star categories.

Hotel Category	Category Rank	LCS_1	LCS_1 Rank	LCS_2	LCS_2 Rank
1_2-★	1	0.64	6	1	9
1_3-★	4.5	0.283	2	0.414	1
2_3-★	4.5	0.105	1	0.887	5
3_3-★	4.5	1	10.5	0.507	2
4_3-★	4.5	0.95	8	0.84	4
5_3-★	4.5	0.492	4	0.663	3
6_3-★	4.5	0.608	5	1	9
1_4-★	9.5	0.67	7	1	9
2_4-★	9.5	0.371	3	0.909	6
3_4-★	9.5	1	10.5	1	9
4_4-★	9.5	0.99	9	1	9

Source: own calculations.

The correlation between the two DEA setups is relatively strong ($\rho = 0.79$), indicating that alternative configurations of narrative aggregation yield broadly consistent patterns

of experiential positioning across hotels. In contrast, the associations between narrative-derived experiential rankings and formal hotel star classifications are weak to moderate ($\rho = 0.21$ for Setup 1 and $\rho = 0.46$ for Setup 2). Given the exploratory nature of the study and the non-probabilistic structure of the dataset, no formal statistical inference is intended, and the reported correlations should be interpreted as descriptive comparative indicators rather than as statistically generalizable relationships.

This evidence confirms that experiential evaluations derived from guest narratives do not closely follow official star rating systems. Instead, they suggest that narrative feedback articulates key dimensions of experiential meaning and interpretation that are only partially reflected in formalized hotel classification schemes.

Taken together, the correlation patterns provide exploratory support for the consistency of the text-driven DEA approach while underscoring the structural divergence between narrative-based experiential assessment and institutionalized rating frameworks.

4.7. Summary of Key Findings

The results demonstrate that narrative-derived experiential evaluation measures can be successfully integrated with DEA to differentiate hotels based on how experiences are articulated and evaluated in guest narratives rather than on formal efficiency or satisfaction metrics. These comparisons should not be interpreted as controlled evaluations of hotel performance. Observed differences may also reflect contextual factors such as pricing structures, market positioning, location, and review volume.

The three-group configuration (Setup 1) provides the highest discriminatory capacity, revealing substantial heterogeneity in experiential positioning across hotels and highlighting marked contrasts in narrative-based experiential evaluation that diverge from formal star classifications. The four-group configuration (Setup 2), while maintaining broadly consistent relative ordering, produces more compressed experiential profiles and a larger number of frontier units, illustrating the well-known trade-off between dimensionality and discriminatory power in DEA models. The application of the Andersen–Petersen super-efficiency model further differentiates among frontier hotels, revealing that a two-star hotel exhibits the strongest relative experiential valuation, thereby challenging expectations derived from formal classification systems. Correlation analysis confirms strong consistency between alternative DEA configurations while indicating weak to moderate associations between narrative-derived experiential rankings and official star categories, underscoring a structural divergence between institutionalized rating frameworks and lived, narratively constructed experience. Taken together, these results provide an exploratory empirical illustration of the proposed narrative-based quantitative framework and support further investigation into narrative-oriented comparative evaluation approaches. Observed differences between hotels should be interpreted as exploratory comparative patterns rather than as definitive indicators of stable experiential superiority. The consistency of results across alternative specifications further supports the robustness of the proposed framework within the exploratory context of the study.

5. Discussion

This section interprets the empirical findings in light of existing theoretical perspectives and outlines their methodological and conceptual implications.

The results of this study contribute to ongoing debates in transformative tourism and sustainability research by illustrating how narrative-derived experiential evaluations may diverge from dominant indicator-driven systems of hospitality performance assessment.

Traditional evaluation regimes rely on standardized classifications and efficiency-oriented metrics. However, the results indicate that guest narratives articulate experiential

value in ways that are not adequately captured by formal star ratings or conventional satisfaction frameworks.

These analyses suggest that experiential value articulated in narratives may reflect not only service evaluation but also broader forms of experiential interpretation and personal resonance. This divergence is consistent with theoretical perspectives in transformative tourism emphasizing interpretation, affective coherence, and narrative sense-making as important dimensions of tourism experience rather than externally imposed quality benchmarks.

By integrating sentiment-derived narrative signals with the DEA framework, the study reframes efficiency not as operational optimization. Instead, it introduces the concept of experiential conversion capacity, understood as the ability of hospitality providers to transform material and organizational conditions into positively evaluated experiences.

The empirical results indicate that hotels occupying similar positions within formal classification systems may display markedly different experiential profiles when evaluated through guest narratives, highlighting the epistemic limitations of growth-oriented and standardizing evaluation models. In this sense, the findings are broadly consistent with critiques of managerial quantification that argue such metrics privilege comparability and control at the expense of experiential depth and lived meaning.

Importantly, the emergence of a lower-category hotel as the strongest case of experiential valuation within the super-efficiency analysis should not be interpreted as evidence that lower-category hotels are inherently more sustainable, less resource-intensive, or more compatible with post-growth tourism models. Rather, the finding suggests that narratively articulated experiential evaluations do not necessarily align with formal classification hierarchies or assumptions linking higher levels of service standardization with greater experiential value.

From a sustainability perspective, these findings can be interpreted as conceptually consistent with post-growth and sufficiency-oriented critiques of tourism development that question the dominance of growth, accumulation, and ranking as primary evaluative principles. Within broader debates concerning tourism under conditions of ecological constraint, the observed divergence between formal classifications and narrative-based experiential evaluations raises questions about the adequacy of standardized evaluation systems for capturing experiential dimensions of tourism value.

Narrative-based experiential evaluations foreground how guests perceive and interpret their stays, allowing forms of value that are relational, modest, and context-sensitive to become visible. By contrast, traditional metrics risk marginalizing such experiences because they do not easily translate into standardized indicators or competitive benchmarking tools. The narrative-based framework proposed here thus offers a methodological pathway for reconsidering how tourism experiences may be comparatively evaluated beyond exclusively growth-oriented performance frameworks.

Illustrative examples from the review corpus further suggest that some guests articulated their experiences in ways extending beyond purely functional satisfaction. Selected reviews included formulations such as “the stay allowed us to truly slow down and disconnect,” “the atmosphere felt authentic and emotionally calming,” or “this was one of the few places where we genuinely felt comfortable and present.” These examples are illustrative and are not intended as representative evidence of broader behavioral patterns. Such narrative expressions may be interpreted as reflecting sufficiency-oriented and meaning-focused experiential modes, which are conceptually aligned with post-growth and sustainability-oriented perspectives on tourism. Such narratives appeared particularly frequently in reviews associated with hotels exhibiting relatively strong experiential positioning within

the DEA framework, suggesting that such evaluations may capture relational and affective dimensions of the hospitality experience not fully reflected in formal classification systems.

While such statements cannot be treated as direct evidence of transformative change, they suggest that at least some narrative evaluations contain reflective, affective, or meaning-oriented dimensions extending beyond attribute-specific service assessments. In this sense, the findings are broadly consistent with transformative tourism perspectives emphasizing everyday, incremental, and contextually situated forms of experiential meaning rather than exceptional or luxury-oriented experiences. At the same time, the study does not claim to empirically demonstrate transformative outcomes or sustainability-oriented forms of tourism practice.

The present framework also does not incorporate explicit environmental or operational input variables, such as resource-consumption or energy-intensity indicators, which may constitute an important direction for future sustainability-oriented extensions of the proposed framework.

Methodologically, the study demonstrates that quantitative tools commonly associated with managerial optimization can be repurposed as critical and meaning-oriented instruments when grounded in narrative data and theory. The use of DEA as a relational comparator rather than a prescriptive efficiency model illustrates how quantitative comparison can coexist with sensitivity to lived experience. Importantly, the proposed framework is not intended to measure transformative outcomes or sustainability performance directly, but rather to offer an alternative evaluative perspective focused on narratively articulated experiential value.

Unlike recent ABSA-based approaches that focus on predictive accuracy and fine-grained aspect extraction, the present study emphasizes interpretive aggregation and comparative evaluation. This reflects a deliberate trade-off between model complexity and interpretive transparency within a narrative-oriented analytical framework. Rather than optimizing classification performance, the proposed approach ensures transparency through explicit rule construction and traceable sentiment assignment at the attribute level.

This repositioning contributes to methodological discussions in tourism research by showing that rigor and reflexivity need not be opposed, and that computational approaches can support more interpretive and meaning-oriented forms of tourism evaluation when their assumptions and limitations are made explicit.

Observed experiential differences may also partially reflect contextual factors such as review volume, pricing structures, location-specific dynamics, and market positioning, which were not explicitly controlled for within the present exploratory framework. Additionally, platform-mediated reviews may reflect demographic, visibility-related, or reputational biases, including the influence of uneven participation patterns or strategically managed online reputation practices. The relatively small number of analyzed units (eleven hotels) also limits the generalizability of the findings and should be taken into account when interpreting the results.

Overall, the findings suggest that narrative-driven quantitative evaluation can enrich understandings of transformative tourism by making experiential dimensions empirically accessible at scale without reducing them to managerial abstractions.

In doing so, the study contributes to ongoing theoretical and methodological discussions concerning how tourism performance is conceptualized, measured, and valued in contexts increasingly shaped by ecological constraints, sustainability transitions, and broader debates surrounding post-growth tourism futures. The purpose of the study is not to provide statistically generalizable findings, but to develop and illustrate a conceptually grounded analytical framework. Its transferability lies primarily in the methodological approach rather than in the specific empirical outcomes. Overall, the proposed approach

provides a foundation for further development of narrative-based comparative evaluation frameworks in tourism and hospitality research.

6. Conclusions

This study develops and empirically illustrates a narrative-based quantitative framework for the comparative evaluation of hospitality experiences. By integrating sentiment-derived signals from user-generated online reviews with a DEA-based approach, the study proposes a shift from efficiency understood as operational optimization toward efficiency conceptualized as experiential conversion capacity.

The findings demonstrate that guest narratives provide evaluative signals that are not fully captured by formal classification systems or standardized performance metrics. Hotels occupying similar positions within official star rating systems may exhibit substantially different experiential profiles when assessed through narrative-based evaluation. This highlights the potential value of incorporating user-generated content into comparative analytical frameworks in tourism research.

Methodologically, the study contributes by showing that quantitative techniques commonly associated with managerial optimization can be repurposed as interpretive tools when grounded in narrative data and theory. The proposed framework offers a transparent and replicable approach that balances computational structure with sensitivity to experiential meaning, providing an alternative to purely indicator-driven evaluation models.

At the same time, the study remains exploratory in nature. The relatively small number of analyzed units and the absence of explicit environmental or operational input variables limit the generalizability of the findings. In addition, the analysis is based on platform-mediated reviews, which may reflect participation biases and contextual influences not fully controlled for in the present framework.

Future research may extend the proposed approach by incorporating additional variables related to sustainability performance, resource use, or operational efficiency, as well as by applying the framework to larger and more diverse samples. Further work could also explore the temporal dynamics of narrative evaluations and compare results across platforms or geographic contexts.

Hence, the study provides a conceptually grounded foundation for further development of narrative-based comparative evaluation frameworks in tourism and hospitality research. By making experiential evaluation empirically accessible at scale, the approach contributes to ongoing discussions concerning how tourism performance may be conceptualized and assessed beyond narrowly defined efficiency and growth-oriented paradigms.

7. Future Research Directions

Future research may also explore alternative strategies for aggregating narrative sentiment, including semantic or context-sensitive weighting approaches that better capture ambiguity and mixed experiential judgments. Longitudinal applications of the framework could examine how experiential evaluations evolve over time, revealing seasonal dynamics or shifts in narrative meaning. Extending the approach to other tourism and service contexts may also help clarify how narrative-based experiential value varies across settings characterized by different institutional and cultural conditions. Such research would further advance efforts to align quantitative tourism evaluation with sustainability-oriented and transformative understandings of experience.

Author Contributions: M.K.: Conceptualization, Writing—original draft preparation, Investigation, Resources, Writing—Review & Editing, Supervision. J.K.: Methodology, Writing—original draft preparation, Formal analysis and Investigation, Visualization, Validation. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The original contributions presented in this study are included in the article. Further inquiries can be directed to the corresponding author.

Acknowledgments: Declaration of Generative AI and AI-Assisted Technologies in the Writing Process: During the preparation of this manuscript, the authors used Microsoft Copilot (M365 Copilot) to assist with language polishing and minor wording improvements. After using this tool, the authors reviewed and edited the content as necessary and took full responsibility for the final version of the manuscript.

Conflicts of Interest: The authors declare no conflicts of interest.

Abbreviations

The following abbreviations are used in this manuscript:

A-P	Andersen–Petersen
CCR	Charnes–Cooper–Rhodes
DEA	Data Envelopment Analysis
DMU	Decision-Making Unit
DOI	Digital Object Identifier
eWOM	Electronic Word-of-Mouth
LCS	Level of Customer Satisfaction
N	Sample Size
PA	Percentage Agreement

Appendix A

Appendix A.1. Description of Sentiment Detection Algorithm

In view of what has been done so far in the field of applying sentiment analysis in decision making, we decided to propose a new algorithm for a couple of reasons. For example, the following packages are available in the R language environment: meanr, sentiment, Rsentiment, Sentiment Analysis, and syuzhet, which focus on detecting the document sentiment rather than the entity or object sentiment, and their quality is unsatisfactory. Their ideas consist of obtaining the emotional load of words from external sources and averaging these loads to obtain the final sentiment. This kind of approach will not work in detecting object sentiment as in this case very often the emotions are verb based and verbs do not have any emotional load. Another difficult particular problem is that one has to pay close attention to punctuation marks to decide to which object the given description refers. Therefore, we proposed a new algorithm designed especially for detecting the object sentiment from hotel reviews. The algorithm has the following characteristics:

- is unsupervised;
- no stoplist words removal;
- no lemmatisation;
- has some punctuation mark removal, but not all.

We also exclude the use of external sources until the point at which it is impossible to extract any new information from the text corpus analysed. In our opinion, there are scarcely any words that can be considered meaningless, certainly not as many as one can find in popular stoplists, and this is the reason for abandoning the removal of stoplist words. Closely connected to this issue is the operation of lemmatisation, i.e., transforming all words to their basic grammatical forms. In our opinion lemmatisation can kill some

possibly valuable information and in view of that we exclude lemmatisation. In many traditional approaches to sentiment detection, the removal of all punctuation marks is mandatory. We advocate to keep any punctuation mark which symbolises the end of a sentence, e.g., full stop, exclamation mark and question mark. In the case of detecting object sentiment, it is absolutely crucial to know where the sentence begins and where it ends.

The idea of the algorithm is to start from two short lists with known sentiment, trace all user-defined objects, count positive cases and negative cases, and in the next stage, to find new descriptors and to append them to one of the lists, thus enlarging them. Below are the two initial lists of words with established sentiment.

Initial list of positive words:

adequate; beautiful; best; better; convincing; decent; delicious; ecstatic; effective; effectively; efficient; efficiently; excellent; fantastic; fine; good; happy; interesting; love; magnificent; nice; nicely; okay; outstanding; positive; positively; pretty; professional; professionally; recommend; recommendation; recommended; reliable; right; satisfied; successful; successfully; super; superb; tasty; true; truly; valuable; value; well; wonderful; worth; worthy.

Initial list of negative words:

bad; break; broken; defeat; destroy; destroyed; dirty; disappointed; disappointing; disappointment; disillusion; disrupt; disturbance; disturbed; dull; failure; inadequate; lack; letdown; lousy; miserable; missing; monotone; mundane; plain; poor; poorest; problem; problems; scandal; scandalous; scandalously; shit; shitty; spoil; spoiled; tacky; unclean; unfortunately; uninteresting; unprofessional; unrealistic; unreliable; unsatisfied; unsuccessful; unsuccessfully; until; untrue; unvaluable; unworthy; valueless; waste; worse; worst.

The detailed steps of the algorithm

Run each step for each of the following terms: location, building, room, bathroom, parking, lifts, breakfast, entertainment, prices, fraud, payment, staff, cleanliness, quietness, heating, air conditioning, and internet.

1. If a review contains a general statement and contains the current term, add 1 to either the positive or negative number of cases (according to the sentiment of the general statement) and add ± 0.65 to each of the remaining terms.
2. If a review contains the current term preceded or followed by any term from one of the lists, add 1 to either the positive or negative number of cases (according to the list sentiment).
3. If a review contains the current term with established sentiment and descriptors, add the descriptors to one of the two descriptor lists corresponding to the term.
4. If a review contains the current term without established sentiment and descriptors, add the descriptors to the waiting list of term descriptors.
5. Verify the waiting list of the current term descriptors to remove from the waiting list each descriptor that is in one of the two term descriptor lists (adding it to this list).
6. Revisit each worthy review to verify if it is possible to assign sentiment to the current term due to one of its two descriptor lists.
7. Apply an external source to identify the sentiment of each descriptor on all waiting lists.
8. Revisit all worthy reviews to verify if it is possible to assign sentiment to the current term due to one of its two descriptor lists.
9. All conditional parts of reviews are treated in a different way.

After applying the proposed algorithm, each attribute is characterized with two positive real numbers, one representing the load of positive opinions and the other the load of negative opinions.

Appendix A.2. General Characteristics of Data Envelopment Analysis

Data Envelopment Analysis (DEA) is a non-parametric method designed to evaluate the relative performance of comparable decision-making units (DMUs) operating under similar conditions. The method was originally proposed to measure relative efficiency by comparing observed combinations of inputs and outputs without imposing a predefined functional form on the production process.

The classical DEA formulation is the CCR (Charnes–Cooper–Rhodes) model, which assumes constant returns to scale and constructs an empirical efficiency frontier based on observed best practices [25]. In this framework, a DMU is considered efficient if no other unit or combination of units produces greater output levels using the same or fewer inputs. Efficiency is thus defined relationally, relative to the performance of other units in the sample.

In the present study, DEA is employed in an output-oriented form and deliberately abstracted from conventional production efficiency interpretation. Rather than assessing efficiency as input minimization or output maximization in an economic sense, the method is repurposed as a comparative relational device for examining how narrative-derived experiential outcomes differ across hotels. Consequently, no explicit input variables are specified, and the analysis focuses exclusively on the relative structure of experiential evaluations expressed in guest narratives.

A common limitation of standard DEA models is that multiple DMUs may attain identical frontier efficiency scores, particularly when the number of units is small relative to the number of output dimensions. To address this issue and obtain finer differentiation among frontier units, the Andersen–Petersen super-efficiency extension is applied [26]. This approach evaluates each efficient unit by temporarily excluding it from the reference set, allowing super-efficiency scores to exceed unity and enabling ranking among otherwise equally efficient DMUs.

Within this study, the super-efficiency framework does not serve as a prescriptive benchmarking or optimization tool. Instead, it supports interpretive comparison by highlighting heterogeneity in experiential conversion among hotels that appear similar under the standard DEA formulation. The method thus functions as an analytical instrument for revealing relative experiential positioning rather than for identifying optimal operational configurations.

Appendix A.3. Descriptive Statistics of Sentiment Assignments Across DEA Narrative Dimensions

This appendix provides descriptive information for the eleven hotels included in the empirical analysis. It summarizes the sentiment distributions underlying the narrative dimensions used in the DEA framework.

Table A1 presents the number of negative and positive sentiment assignments for each hotel across five narrative dimensions: General, Premises, Staff, Finance, and Resting. For each hotel, the first row reports negative values, while the second row reports positive values. These data provide a transparent overview of the sentiment structure on which the DEA-based experiential evaluation is built. These dimensions correspond directly to the aggregation structure used to construct DEA outputs in the empirical analysis.

These statistics demonstrate that the DEA indicators are based on substantial numbers of both positive and negative sentiment assignments across all analyzed hotels. The observed variation in sentiment distributions across narrative dimensions provides additional support for the discriminatory capacity of the proposed framework and offers readers greater transparency regarding the empirical basis of the subsequent DEA analysis.

Table A1. Positive and negative sentiment assignments by hotel and narrative dimension. For each hotel, the first row reports negative sentiment assignments, and the second row reports positive sentiment assignments.

	General	Premises	Staff	Finance	Resting
Andels	10	59	30	15	29
	112	335	427	1	84
Boss	32	65	94	15	23
	68	330	447	10	219
Boutique	14	226	60	16	162
	24	386	247	77	10
Ibis	6	33	11	25	16
	38	262	409	37	164
Double Tree	10	44	88	36	15
	86	211	390	31	107
Focus	9	51	56	41	10
	26	349	283	42	225
Citi	10	215	64	12	60
	36	182	200	38	84
Iness	12	62	26	20	26
	53	275	378	43	142
Nobo	5	46	42	27	46
	64	219	391	31	93
Novotel	6	22	71	27	10
	56	287	385	10	87
Qubus	5	32	23	21	41
	51	349	352	31	172

References

1. Reisinger, Y. *Transformational Tourism: Tourist Perspectives*; CABI: Wallingford, UK, 2013.
2. Lean, G.L. Transformative travel: Inspiring sustainability. *Tour. Recreat. Res.* **2009**, *34*, 103–108.
3. Pung, J.M.; Del Chiappa, G. Transformative travel experiences: A conceptual model. *J. Travel Res.* **2020**, *59*, 537–552.
4. Kirillova, K.; Lehto, X.; Cai, L. Existential authenticity and anxiety as outcomes: The tourist in the experience economy. *Int. J. Tour. Res.* **2017**, *19*, 13–26.
5. Kirillova, K.; Lehto, X. Destination aesthetics and aesthetic distance in tourism experience. *J. Travel Tour. Mark.* **2015**, *32*, 1051–1068. [[CrossRef](#)]
6. Noy, C. Tourism as narrative: Producing, consuming, and traveling stories. *Ann. Tour. Res.* **2007**, *34*, 349–368.
7. Zátori, A.; Smith, M.K.; Puczkó, L. Experience involvement, memorability and authenticity. *Tour. Manag.* **2018**, *67*, 111–126. [[CrossRef](#)]
8. Tribe, J. The indiscipline of tourism. *Ann. Tour. Res.* **1997**, *24*, 638–657. [[CrossRef](#)]
9. Hall, C.M. Degrowing tourism: Décroissance, sustainable consumption and steady-state tourism. *Anatolia* **2009**, *20*, 46–61. [[CrossRef](#)]
10. Espeland, W.N.; Stevens, M.L. A sociology of quantification. *Eur. J. Sociol.* **2008**, *49*, 401–436. [[CrossRef](#)]
11. Higgins-Desbiolles, F. Socialising tourism for social and ecological justice. *Tour. Geogr.* **2020**, *22*, 610–623. [[CrossRef](#)]
12. Gössling, S.; Scott, D.; Hall, C.M. Pandemics, tourism and global change: A rapid assessment. *J. Sustain. Tour.* **2020**, *28*, 1–20. [[CrossRef](#)]
13. Xiang, Z.; Gretzel, U. Role of social media in online travel information search. *Tour. Manag.* **2010**, *31*, 179–188. [[CrossRef](#)]
14. Filieri, R.; McLeay, F. eWOM and accommodation choice. *Tour. Manag.* **2014**, *40*, 26–40.

15. Sparks, B.; Browning, V. The impact of online reviews on hotel booking intentions. *Tour. Manag.* **2011**, *32*, 1310–1323. [[CrossRef](#)]
16. Bosangit, C.; Dulnuan, J.; Mena, M. Using travel blogs to examine the meaning of place. *Tour. Manag.* **2015**, *53*, 185–197.
17. Pera, R. Empowering the new traveller: Storytelling as a co-creative behaviour in tourism. *Tour. Manag.* **2014**, *43*, 27–35.
18. Kladou, S.; Mavragani, E. Assessing destination image through narratives. *J. Destin. Mark. Manag.* **2015**, *4*, 137–145. [[CrossRef](#)]
19. Salganik, M.J. *Bit by Bit: Social Research in the Digital Age*; Princeton University Press: Princeton, NJ, USA, 2017.
20. Pang, B.; Lee, L. Opinion mining and sentiment analysis. *Found. Trends Inf. Retr.* **2008**, *2*, 1–135. [[CrossRef](#)]
21. Thelwall, M.; Buckley, K.; Paltoglou, G. Sentiment strength detection. *J. Am. Soc. Inf. Sci. Technol.* **2012**, *63*, 163–173.
22. Ameer, A.; Hamdi, S.; Ben Yahia, S. Sentiment analysis for hotel reviews: A systematic literature review. *ACM Comput. Surv.* **2023**, *56*, 51. [[CrossRef](#)]
23. Erdoğan, D.; Kayakuş, M.; Çelik Çaylak, P.; Ekşili, N.; Moiceanu, G.; Kabas, O.; Ichimov, M.A.M. Developing a deep learning-based sentiment analysis system of hotel customer reviews for sustainable tourism. *Sustainability* **2025**, *17*, 5756. [[CrossRef](#)]
24. Park, J. Combined text mining/DEA method for customer satisfaction. *Expert Syst. Appl.* **2023**, *232*, 120767. [[CrossRef](#)]
25. Charnes, A.; Cooper, W.W.; Rhodes, E. Measuring the efficiency of decision-making units. *Eur. J. Oper. Res.* **1978**, *2*, 429–444. [[CrossRef](#)]
26. Andersen, P.; Petersen, N.C. A procedure for ranking efficient units in DEA. *Manag. Sci.* **1993**, *39*, 1261–1264. [[CrossRef](#)]
27. Emrouznejad, A.; Yang, G.L. Forty years of DEA. *Socio-Econ. Plan. Sci.* **2018**, *61*, 4–8. [[CrossRef](#)]
28. Barros, C.P. Measuring efficiency in the hotel sector. *Int. J. Contemp. Hosp. Manag.* **2005**, *17*, 205–213.
29. Cronbach, L.J. Coefficient alpha and the internal structure of tests. *Psychometrika* **1951**, *16*, 297–334. [[CrossRef](#)]
30. Korzeniewski, J.; Kozłowski, M.; Piotrowska-Piątek, A. Importance–performance analysis based on text mining. *Eur. Res. Stud. J.* **2025**, *28*, 450–470. [[CrossRef](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.