Innovative Organizational Design
https://doi.org/10.63808/IOD.v2i3.475 Volume 2 , Issue 3

Research Article

What Drives Luxury Hotel Guest Satisfaction? Evidence from Online Reviews Using Text Mining

Wen Lan[1,*]
[1]Faculty of Hospitality Management, Swiss Hotel Management School/University of Derby, Leysin 1854, Switzerland
[*]**Correspondence:** Wen Lan, yokolan@gmail.com

Abstract

Online reviews have become a strategic data asset for hotel operations, yet which attributes actually drive guest satisfaction in the luxury segment remains insufficiently quantified. This study examines the drivers of luxury hotel guest satisfaction through text mining of 30,442 English-language TripAdvisor reviews covering 55 upscale and luxury hotels in Shanghai. Latent Dirichlet Allocation identified seven interpretable satisfaction themes, sentiment analysis quantified the affective tone of each review, and an ordinary least squares regression linked theme salience and sentiment to the star rating, with the model explaining 53.1% of the variance in ratings. Personalized, loyalty-oriented service emerges as the strongest positive correlate of the overall evaluation, followed by iconic views and overall service excellence, a ranking that diverges from what topic prevalence alone would suggest. Physical facilities contribute more modestly, and the front-desk encounter constitutes the clearest satisfaction shortfall among substantive service attributes. The findings connect data-driven themes to established service-quality dimensions, support a two-factor reading of luxury attributes, and direct managerial investment toward the informational infrastructure of personalization and the digital redesign of the arrival touchpoint. The study offers a replicable, review-based instrument for locating where luxury guest satisfaction is won and lost.

Keywords

Hotel management, Service quality, Digital operations, Guest experience, Industry optimization

1. Introduction

Online reviews have developed from a side effect of reputation into active data resources for hotel management. Research scholars now increasingly consider them direct reports on guests' actual experiences with hotel services (Zarezadeh et al., 2022). For the high-end sector, the stakes are particularly large. Properties are in a race for the attention of guests through rich experiences that guests share on social media. In high-growth areas such as Shanghai, with many international first-class hotels nearby, the ability to extract operational information from unstructured review text has become a new competitive factor. This trend has recently expanded to the evaluation of smart and digitally connected hotels using the same review data (Song et al., 2024).

Research on mining this resource has developed rapidly, and Section 2 reviews it in detail; three deficiencies stand out. Most of the existing studies are descriptive; they list the themes in the evaluation and how positively they are regarded, but they rarely identify which themes have had an impact on the overall judgment. The language of drivers that these studies adopt often exceeds what the data can support. Furthermore, the luxury segment is often grouped with mid-scale properties. Since tangible standards are uniformly high in this segment, the basics leave little room for differentiation, and what truly matters becomes indistinct. A design that identifies the themes guests mention, assesses the emotional response people have to those themes, and then tests their joint association with the numerical score is relatively rare, especially for luxury hotels in a Chinese metropolitan area.

This paper takes up these deficiencies. Based on 30,442 English-language TripAdvisor reviews of 55 upscale and luxury hotels in Shanghai, topic model extraction was performed to obtain latent satisfaction themes, and the affective tone of each review was quantified. The association between theme salience, sentiment and star rating was then estimated to approximate the drivers of luxury guest satisfaction. What separates this design from the closest prior work is that theme salience and sentiment enter a single estimation with the rating as the outcome, so themes are ranked by their association with the score rather than by how often they appear in the text. The contributions are of three kinds: the study shifts the focus of review mining from description to driver identification, provides segment-specific evidence for luxury hotels instead of pooled samples, and translates the resulting ranking of drivers into implications for the digital and personalized service operations in which this segment is increasingly competing. Therefore, this study connects the empirical themes derived from the review text with existing theories of service quality. The aim is a replicable tool for hotel managers to find highly effective sources of guest delight and operational touchpoints that silently reduce satisfaction.

2. Literature Review

2.1. From description to estimation in review mining

The mining of online reviews for hotel management has a clear point of origin. Xiang et al. (2015)

applied text analytics to a large corpus of hotel reviews and showed that the guest experience can be decomposed into interpretable dimensions whose presence is associated with satisfaction ratings. The toolkit has since widened considerably; Bi et al. (2024) trace its expansion from word counts to topic models and deep learning architectures across the tourism and hospitality journals. Reviews of the sentiment literature point in a consistent direction. Mehraliyev et al. (2022) analyzed 70 studies and found that the customer rating is the most frequently investigated dependent variable while moderating and mediating mechanisms are rarely tested; measurement, in other words, has outpaced explanation. Two further findings sharpen the case for treating text and score as related but distinct signals. Y. J. Kim and H. S. Kim (2022) showed that experience features extracted from review text correlate statistically with satisfaction scores. Lai et al. (2021) found the relationship to be asymmetric, with negative sentiment weighing more heavily on ratings than positive sentiment and the strength of the link varying with review characteristics, so sentiment cannot be read as a simple proxy for the numerical score. Text mining has even begun to produce measurement instruments of its own; Kwon (2023) derived a co-created value scale for hospitality service directly from review text. What differs across these studies is not the availability of methods but whether the analysis stops at describing themes or proceeds to estimating their weight in the overall evaluation.

2.2. Luxury attributes and the structure of satisfaction

How text-derived dimensions map onto satisfaction still draws on two older frameworks. Parasuraman et al. (1988) distilled service quality into tangibles, reliability, responsiveness, assurance and empathy, and this five-dimension vocabulary continues to organize the interpretation of review themes. B. Kim et al. (2016), analyzing satisfiers and dissatisfiers in social-media hotel reviews, gave the two-factor account its review-based form: physical attributes behave mainly as expected baselines whose absence hurts, whereas staff-related qualities generate delight. Whether these regularities hold in the luxury segment is a question the field has isolated only recently. A bibliometric review by Jain et al. (2023) organized luxury hospitality research into six clusters and placed digital interactions, online reviews and complaint handling at the head of the list, which locates review-based evidence at the center of the segment's research agenda. Empirical work has followed. Wu et al. (2023) mined luxury hotels' reviews for consumers' affective needs through a five-stage framework built on Kansei engineering, showing that affect carries information beyond functional attributes in this segment. Lv et al. (2024) found that AI-powered personalized recommendations raise customers' life satisfaction in the pre-purchase phase of luxury tourism, a sign that personalization is becoming the axis on which the segment competes. These studies document what luxury guests value; they rarely weigh the attributes against one another on a common outcome.

2.3. Temporal dynamics and evidence from China

Evidence from the Chinese market adds a time dimension to both strands. Sun et al. (2022) documented that the attributes shaping hotel satisfaction in Beijing and Shanghai shifted with the arrival of COVID-19. Y. Song et al. (2022) reached a parallel conclusion with Latent Dirichlet Allocation and sentiment analysis, finding that the factors influencing satisfaction and their emotional strength differed before and after the outbreak. Xu et al. (2022) compared pre- and post-pandemic reviews directly and showed that the structure explaining satisfaction was not stable across the two periods, and Yu et al. (2022), analyzing 6,556 pandemic-related reviews, added that the crisis response strategy a hotel adopted moderated guest satisfaction. Together these studies carry a methodological warning for any dataset whose window spans 2020 to 2022: attribute weights estimated over such a window are averages across regimes, and their interpretation should say so.

2.4. Research gap

The segment-specific work closest to the present design is Pang et al. (2025), who applied topic mining and aspect-based sentiment analysis to five-star hotel reviews and produced a fine-grained account of guest concerns; the analysis stops short of connecting theme salience with the overall rating inside a single estimation system. That missing step is the one toward which the three strands above jointly point. Review-mining methods are mature but mostly descriptive; luxury research names the attributes that matter without ranking their weight on a common outcome; and the Chinese evidence shows that those weights move with time. Extracting the themes, quantifying the affect attached to them, and estimating their combined association with the rating within one luxury market would answer all three concerns at once, and the sections that follow implement exactly that design.

3. Data and Methodology

3.1. Data collection and luxury hotel sampling criteria

This paper draws on the Shanghai subset of a publicly available TripAdvisor hotel review dataset. Shanghai concentrates an unusually dense cluster of international luxury brands inside one competitive market, which keeps market structure, source-market composition and price band broadly comparable across the sampled properties. Because the research question concerns the luxury segment, selection applied two screens. At the property level, a hotel entered the sample only if it operates under an established luxury or upscale brand, such as The Peninsula, Four Seasons, Waldorf Astoria, Mandarin Oriental, The Ritz-Carlton, Park Hyatt, Shangri-La, Kempinski, The Langham, and JW Marriott; properties bearing economy or midscale brand identifiers were excluded. Brand membership was preferred to platform star ratings because stars and price bands are recorded inconsistently on review platforms, whereas the brand a property operates under offers an objective and reproducible standard for inclusion. At the review level, only English-language reviews were retained, and reviews with fewer than five words or implausible stay dates were excluded. Ratings recorded on a ten-point scale in the source data were normalized to the conventional five-point scale. The final sample contains 30,442 reviews of 55 upscale and luxury hotels, with stay dates from 2010 through 2023; this window reflects the coverage of the source dataset and includes the pandemic years, a point taken up in the limitations.

3.2. Data preprocessing

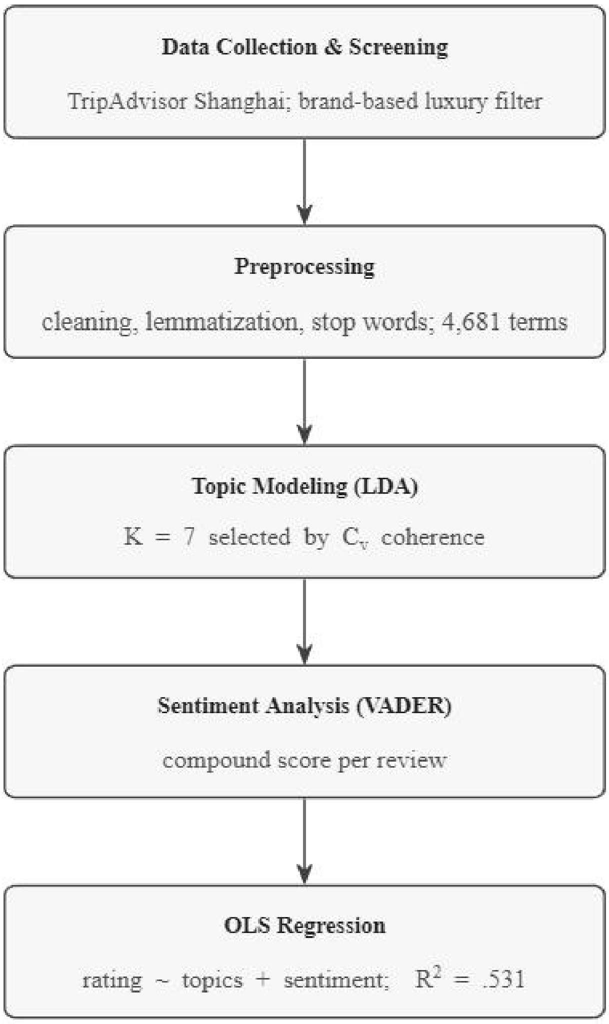
All reviews were lowercased, non-alphabetic characters were removed, and the texts were then tokenized and lemmatized. Standard English stop words were removed, and a domain-specific stop list containing uninformative high-frequency terms such as hotel, room, and stay, as well as hotel and brand names, was added to exclude them from the topic structure. Tokens that appear in fewer than twenty reviews or in more than half of the corpus were excluded from the dictionary, and a vocabulary of 4,681 terms suitable for topic modeling was obtained.

3.3. Analytical methods

The three analysis stages were carried out in sequence. Latent Dirichlet Allocation was used to extract the hidden themes that guests talk about, and recently, in hospitality research, its ability to recover service-quality dimensions from unstructured hotel reviews has been demonstrated (Kalnaovakul & Promsivapallop, 2023). Then, the VADER lexicon was used to assign sentiment scores to all the reviews; as it is suitable for short, informal language on the Internet, it returns a compound polarity score between -1 and 1. To approximate the drivers of satisfaction, ordinary least squares regression was used to estimate a model with a five-point rating as the dependent variable and the predictors being each review's topic-probability distribution along with its sentiment score. Because the topic proportions sum to one within a review, the dining theme (T7) was omitted from the predictor set and treated as the reference category; each remaining coefficient therefore indicates the expected change in rating when review content shifts from the dining baseline toward that theme. This design connects what guests write about and how positively they write about it to the evaluation they give; it identifies associations rather than strict causal effects, and this boundary has been revisited in the limitations of this study. Figure 1 shows the analysis path of the data collection and regression.

Figure 1

Analytical pipeline: from data collection to regression



3.4. Model validation and robustness

The number of topics was not fixed in advance but determined empirically. Candidate models with four to seven topics were estimated, and for each, the Cv coherence score was calculated; the coherence increased steadily with the number of topics and reached a maximum at seven topics (0.401), so this model was selected to balance statistical fit and interpretability. The sentiment measure was validated against the behavioral benchmark available in the data: the compound score positively correlates with the star rating (Pearson $r = .57$), and a three-way classification derived from sentiment agrees with a parallel classification derived from ratings at a level of Cohen's $\kappa = .22$; although this correspondence is modest, it is not negligible given that written tone and numerical scoring capture related but different aspects of evaluation. Descriptive statistics of the final sample, such as the distribution of ratings and review lengths, are shown at the beginning of the results.

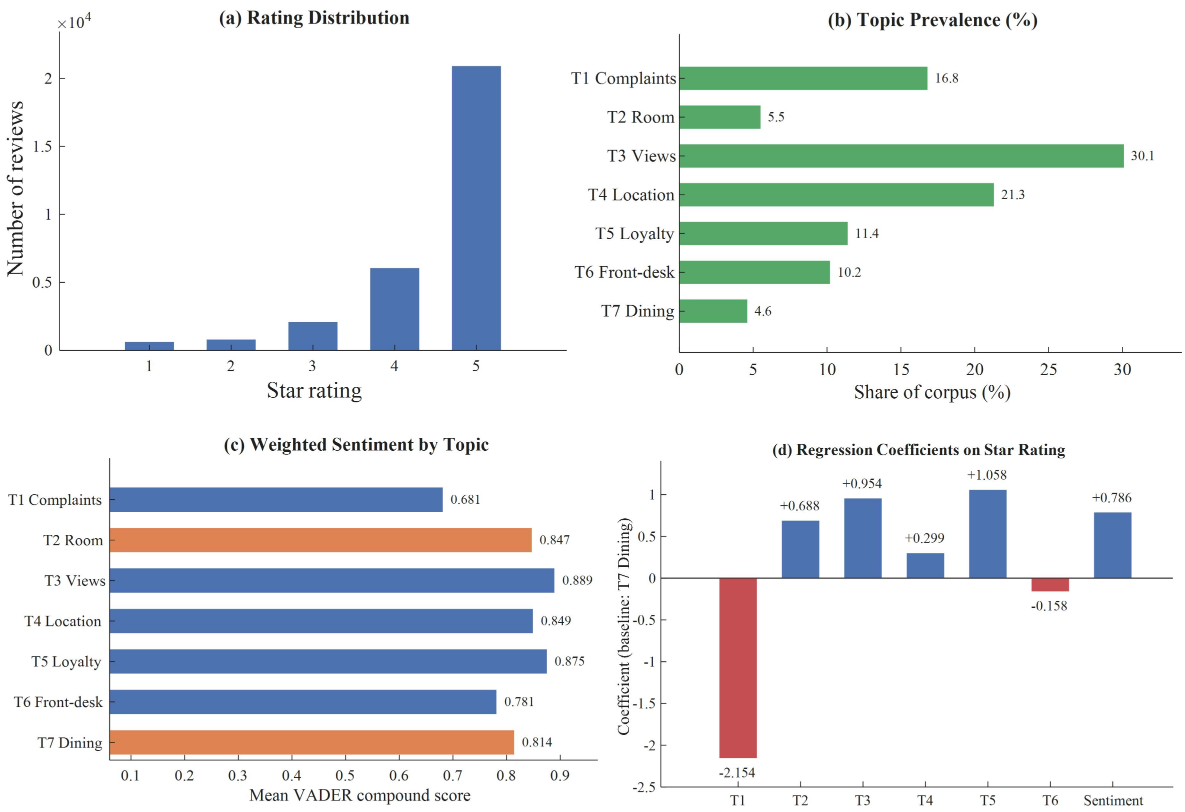
4. Results

4.1. Descriptive analysis of reviews

The final corpus includes 30,442 English-language reviews of the 55 upper-scale and luxury hotels in Shanghai, and the average length of these reviews is 117.8 words. The distribution of the ratings is heavily left-skewed, which is typical for the luxury sector; 68.7 % (n = 20,915) of the samples gave a five-star rating, 19.9 % (n = 6,044) gave a four-star rating, 6.8 % gave a three-star rating, 2.6 % gave a two-star rating, and 2.0 % gave a one-star rating. The mean score is 4.51 (SD = 0.89). The mean compound score is 0.827, most of the comments are positive, and only 5.1 % are negative. This pronounced ceiling effect means that most guests award the highest available score, so the rating scale compresses at its upper end and separates guests only weakly. Competition in the high-end segment plays out inside that narrow band. Differentiation, then, has to be recovered from what reviewers write rather than from the score itself. Figure 2 shows the rating distribution and the topic-level results described below.

Figure 2

Main results: ratings, topics, sentiment, and regression estimates



4.2. Key satisfaction themes identified

According to the coherence-based selection in the method, the seven-topic solution was selected and yielded interpretable themes. Table 1 shows the representative keywords, frequency, weighted sentiment and mean rating of the reviews for each theme. The most popular theme is “views and all-round service excellence” (T3, accounting for 30.1% of the corpus), which suggests that Bund and skyline views are central to the luxury image of Shanghai, followed by “location and convenience”

(T4, at 21.3%). A single complaints theme (T1, 16.8%) groups the evaluative and grievance-oriented language of dissatisfied guests, rather than specifying any single service attribute, and is handled accordingly in the analysis. Loyalty and personalized service (T5, 11.4%) and front-desk and concierge encounters (T6, 10.2%) fall in the middle range of prevalence; room and bathroom facilities (T2, 5.5%) and dining (T7, 4.6%) are mentioned less frequently. The rating profiles of the themes are quite different; those mainly covered by T5 in the reviews have achieved a high mean score in the sample (4.90), while those mainly covered by T1 have scored as low as 3.45 and are nearly a full point lower on a five-point scale.

Table 1

Satisfaction themes identified by LDA: keywords, prevalence, sentiment, and dominant-review rating

Theme	Representative keywords	Prev. (%)	Weighted sentiment	Mean rating (dominant)	N (dominant)
T1 Complaints / unmet expectations	one, would, however, could, still, service, time, star	16.8	0.681	3.45	5,159
T2 Room & bathroom facilities	bathroom, shower, bed, bath, water, toilet, floor	5.5	0.847	4.67	730
T3 Views & overall excellence	view, bund, excellent, best, amazing, great, staff	30.1	0.889	4.86	12,411
T4 Location & convenience	location, breakfast, area, business, restaurant	21.3	0.849	4.55	7,154
T5 Loyalty & personalized service	always, lounge, team, feel, member, year, service	11.4	0.875	4.90	2,609
T6 Front-desk & concierge	check, desk, front, concierge, manager, member	10.2	0.781	4.33	1,894
T7 Dining & food and beverage	food, breakfast, buffet, dinner, tea, wine, chinese	4.6	0.814	4.40	485

4.3. Sentiment analysis results

The weighted sentiment differs clearly across the seven themes and matches the pattern of ratings. The most positive affective tones are attached to views and overall excellence (0.889) and to loyalty and personalized service (0.875); room facilities (0.847) and location (0.849) are also relatively high. Two topics fall clearly below this range. The complaints theme has the lowest weighted sentiment (0.681) and aligns with its contents. Front-desk and concierge interactions are also relatively less positive, with a weighted sentiment score of 0.781 and a dominant review rating of 4.33; specifically, the softest link in an otherwise excellent guest journey occurs at the time of arrival, check-in and concierge assistance. As this pattern is related to a particular service process and not a general dissatisfaction, it will directly impact operations and will be shown in the following regression analysis.

4.4. Regression: drivers of the overall rating

Table 2 shows the ordinary least squares estimates that relate theme salience and review sentiment to the five-point rating, with the dining theme as the reference category. The model is statistically significant, and the extent to which ratings are explained by the model is about 53.1 % ($R^2 = .531$, adjusted $R^2 = .531$, $F = 4927.7$, $p < .001$, $N = 30{,}442$). All predictors are statistically significant. Among the positive drivers, loyalty and personalized service have the largest coefficients ($b = 1.058$, $p < .001$), followed by views and overall excellence ($b = 0.954$, $p < .001$) and room and bathroom facilities ($b = 0.688$, $p < .001$); location and convenience are also positive but less so ($b = 0.299$, $p < .001$). Review sentiment has a relatively strong independent effect ($b = 0.786$, $p < .001$), and thus predicts the rating more effectively than what guests choose to write about. The two coefficients are negative numbers. Complaints are linked to a considerable drop in the rating ($b = -2.154$, $p < .001$), and thus, the language of complaints is associated with lower evaluations. Notably, the front-desk and concierge theme also has a negative coefficient ($b = -0.158$, $p = .005$): compared with reviews focused on dining, greater attention to the arrival and check-in experience is associated with a lower rating. Taken together, the estimates show that relational and experiential attributes, personalized recognition and iconic sense of place among them, have the highest weight in the model; physical facilities contribute positively but less so, and the front desk accounts for the largest satisfaction shortfall among substantive service attributes.

Table 2

OLS regression predicting the five-point rating ($N = 30{,}442$; $R^2 = .531$). Reference category: T7 Dining

Predictor	Coef.	SE	t	p
Constant	3.725	0.048	78.30	< .001
T1 Complaints / unmet expectations	−2.154	0.055	−39.07	< .001
T2 Room & bathroom facilities	0.688	0.065	10.63	< .001
T3 Views & overall excellence	0.954	0.051	18.78	< .001
T4 Location & convenience	0.299	0.050	5.94	< .001
T5 Loyalty & personalized service	1.058	0.055	19.17	< .001

T6 Front-desk & concierge	−0.158	0.057	−2.78	.005
Review sentiment (VADER)	0.786	0.011	73.45	< .001

5. Discussion

5.1. Drivers of luxury satisfaction

According to the pattern of the estimate, satisfaction in the luxury sector is more concerned with relationships and experiences than with tangible goods. The most positive factor is loyalty and personalized service; that is to say, guests are recognized, known well, and given customized care, and the reviews dominated by this theme have also received the highest mean score among all samples. Among the words in the SERVQUAL scale, 'empathy' is closely related to this theme, and its prominence indicates that when tangible standards are all at a high level, the individualized dimension of service becomes the main area of assessment for luxury hotels. The negative coefficient of the front-desk and concierge encounters reads as a responsiveness deficiency at the arrival stage, by the same logic. Room and bathroom facilities are positively associated, but they have a weaker impact than other related attributes; in line with the two-factor theory of hotel attributes, physical tangibles operate as hygiene factors and are thus expected rather than actively desired, whereas staff-related qualities function as genuine motivators. The complaints theme requires a different interpretation: as it is a collection of evaluative expressions rather than an isolated characteristic, it has a large negative coefficient and thus only indicates dissatisfaction where it occurs; the actual content of dissatisfaction lies in the attribute-level contrasts mentioned above.

5.2. Comparison with prior research

The estimates speak directly to the three strands reviewed in Section 2. Within the methods strand, the review of seventy sentiment studies found ratings to be the most common outcome yet mechanisms rarely tested; the present model addresses that imbalance by estimating theme weights and sentiment jointly rather than reporting them side by side. The result also confirms, in a luxury setting, that sentiment is not a simple proxy for the score: the sentiment coefficient ($b = 0.786$) remains strong with all seven themes controlled, in line with the Section 2.1 finding that written tone and numerical score are related but distinct signals. Within the luxury strand, the ranking sharpens what earlier segment work left open. Studies of five-star reviews identified fine-grained concerns but did not weigh them on a common outcome; here the weighing shows that the most discussed theme, views and overall excellence (30.1% of the corpus), is not the strongest driver, whereas the less discussed personalized-service theme (11.4%) is. Prevalence and importance diverge. Frequency counts alone cannot reveal that difference. The pattern that service encounters outweigh physical attributes is consistent with cross-cultural evidence (Zhang et al., 2024) that guest segments weight experience attributes differently across markets and languages; the present ranking therefore describes English-writing guests in one metropolitan luxury market and travels no further than that.

5.3. Theoretical implications

Three implications follow for theory. The mapping from data-driven themes to service-quality dimensions carries weights rather than labels alone: the empathy-adjacent theme of personalized service is estimated at $b = 1.058$ against $b = 0.688$ for the tangibles-adjacent theme of room facilities. In a segment where tangible standards converge at a high level, the five classic dimensions are therefore not equally important, and their ordering can be recovered from unprompted guest text. The estimates also give the two-factor account invoked above a quantitative form. Hygiene and motivator roles separate inside a single model: physical facilities stay positive but modest, relational and experiential themes carry the largest positive weights, and the front-desk theme, at $b = -0.158$, is the one substantive attribute that behaves as a dissatisfier. The design itself carries an implication of its own. Combining topic salience and sentiment in one estimation turns review mining from a description of what guests discuss into a measurement of how much each attribute matters, and that conversion travels to other markets where reviews and ratings coexist.

5.4. Practical implications

For practice, the results identify two complementary levers. Because personalized recognition is the strongest driver, luxury operators should invest in the informational infrastructure that makes empathy repeatable, including unified guest profiles and preference histories that allow returning guests to be recognized consistently rather than incidentally. Because the arrival encounter is the clearest shortfall, it is the natural target for digital service redesign, and evidence that thoughtfully implemented smart-hotel technology enhances guest experience supports directing mobile check-in, digital keys, and pre-arrival communication precisely at this touchpoint (Yang et al., 2021). Automating the transactional layer of arrival can simultaneously remove the friction that depresses ratings and release staff time for the high-touch interactions that this study associates with the highest ratings.

6. Conclusion

This paper set out to locate the drivers of luxury guest satisfaction in online reviews, and the estimation delivers a ranking rather than a list. Personalized, loyalty-driven service carries the largest positive weight, iconic views and overall excellence follow, physical facilities contribute modestly, and the front-desk encounter is the one substantive attribute associated with lower ratings; together with review sentiment, the seven themes explain 53.1% of the variation in scores. The novelty lies in the estimation design. Placing theme salience and sentiment in a single model separates how often guests discuss an attribute from how much that attribute moves their evaluation, and the two orderings disagree. For theory, the weights give the service-quality dimensions and the two-factor account an empirical ordering recovered from unprompted guest text.

For executives and hospitality groups, four directions follow. Guest recognition should be treated as an information asset: because personalized service is the strongest driver (b = 1.058), a guest record shared across properties, carrying preferences and stay history, turns empathy from an individual talent into a repeatable process. The arrival sequence is the immediate target for digital redesign, given the front-desk coefficient of −0.158. Investment cases should be weighed against driver coefficients rather than mention counts, since the most discussed attribute is not the most consequential one. The theme instrument itself is operational: scoring each property's review stream against the same seven themes each quarter gives a group-level view of where each property gains and loses satisfaction.

Several limitations qualify these conclusions. The estimates are associations; no causal ordering between what guests experienced and how they rated it can be established from cross-sectional review data. The 2010 to 2023 window reflects the coverage of the source dataset, so the post-pandemic recovery lies outside it, and because the window pools the pandemic years, the estimated weights are averages across market regimes in the sense flagged in Section 2.3. The sample is bounded in other ways as well: English-language reviews, one metropolitan market, and brand-affiliated properties only, so independent luxury hotels are absent. Sentiment was measured at the review level, which attaches one affective score to a text that may praise one attribute and fault another.

These limitations mark the next steps. Extending the corpus into 2024 and 2025 would show whether the pandemic-era averages have shifted in recovery. Comparisons across cities and languages would test how far the Shanghai ranking travels. Aspect-level sentiment methods would attach affect to individual attributes instead of whole reviews, and designs with panel structure or natural experiments could move the association toward cause. Price, seasonality and traveler type are natural controls to add along the way.

Acknowledgments

The author declares that no specific acknowledgments are required.

Author contributions

Conceptualization, Wen Lan; methodology, Wen Lan; software, Wen Lan; validation, Wen Lan; formal analysis, Wen Lan; investigation, Wen Lan; resources, Wen Lan; data curation, Wen Lan; writing—original draft preparation, Wen Lan; writing—review and editing, Wen Lan; visualization, Wen Lan; supervision, Wen Lan; project administration, Wen Lan; funding acquisition, Wen Lan. All authors have read and agreed to the published version of the manuscript.

Funding

This research received no external funding.

Data availability

The data analysed in this study were obtained from a publicly available TripAdvisor hotel review dataset. The processed data are available from the corresponding author on reasonable request.

Declarations

Conflict of interest

The author declares that there is no conflict of interest.

Ethics approval

Not applicable.

Consent to participate

Not applicable.

Consent for publication

Not applicable.

Additional information

Received: 14 July 2026

Accepted: 5 August 2026

Published online: 8 September 2026

Publisher's Note

Wisdom Academic Press Ltd. remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

References

[1] Bi, J. W., Zhu, X. E., & Han, T. Y. (2024). Text analysis in tourism and hospitality: A comprehensive review. *Journal of Travel Research, 63*(8), 1847–1869. https://doi.org/10.1177/00472875241247318

[2] Jain, V., Wirtz, J., Salunke, P., Nunkoo, R., & Sharma, A. (2023). Luxury hospitality: A systematic literature review and research agenda. *International Journal of Hospitality Management*, 115, 103597. https://doi.org/10.1016/j.ijhm.2023.103597

[3] Kalnaovakul, K., & Promsivapallop, P. (2023). Hotel service quality dimensions and attributes: An analysis of online hotel customer reviews. *Tourism and Hospitality Research, 23*(3), 420–440. https://doi.org/10.1177/14673584221145819

[4] Kim, B., Kim, S., & Heo, C. Y. (2016). Analysis of satisfiers and dissatisfiers in online hotel reviews on social media. *International Journal of Contemporary Hospitality Management, 28*(9), 1915–1936. https://doi.org/10.1108/IJCHM-04-2015-0177

[5] Kim, Y. J., & Kim, H. S. (2022). The impact of hotel customer experience on customer satisfaction through online reviews. *Sustainability, 14*(2), 848. https://doi.org/10.3390/su14020848

[6] Kwon, W. (2023). A co-created value scale for the hospitality service: Applying a text mining approach. *International Journal of Contemporary Hospitality Management, 35*(7), 2352–2373. https://doi.org/10.1108/IJCHM-04-2022-0458

[7] Lai, X., Wang, F., & Wang, X. (2021). Asymmetric relationship between customer sentiment and online hotel ratings: The moderating effects of review characteristics. *International Journal of Contemporary Hospitality Management, 33*(6), 2137–2156. https://doi.org/10.1108/IJCHM-07-2020-0708

[8] Lv, L., Chen, S., Liu, G. G., & Benckendorff, P. (2024). Enhancing customers' life satisfaction through AI-powered personalized luxury recommendations in luxury tourism marketing. *International Journal of Hospitality Management*, 123, 103914. https://doi.org/10.1016/j.ijhm.2024.103914

[9] 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

[10] Pang, T., Liu, J., Han, L., Liu, H., & Yan, D. (2025). A deep learning-based analysis of customer concerns and satisfaction: Enhancing sustainable practices in luxury hotels. *Sustainability, 17*(8), 3603. https://doi.org/10.3390/su17083603

[11] Parasuraman, A., Zeithaml, V. A., & Berry, L. L. (1988). SERVQUAL: A multiple-item scale for measuring consumer perceptions of service quality. *Journal of Retailing, 64*(1), 12–40.

[12] Song, B., Xia, H., Law, R., Muskat, B., & Li, G. (2024). Discovery of smart hotels' competitiveness based on online reviews. *International Journal of Hospitality Management*, 123, 103926. https://doi.org/10.1016/j.ijhm.2024.103926

[13] Song, Y., Liu, K., Guo, L., Yang, Z., & Jin, M. (2022). Does hotel customer satisfaction change during the COVID-19? A perspective from online reviews. *Journal of Hospitality and Tourism Management*, 51, 132–138. https://doi.org/10.1016/j.jhtm.2022.02.027

[14] Sun, S., Jiang, F., Feng, G., Wang, S., & Zhang, C. (2022). The impact of COVID-19 on hotel customer satisfaction: Evidence from Beijing and Shanghai in China. *International Journal of Contemporary Hospitality Management, 34*(1), 382–406. https://doi.org/10.1108/IJCHM-03-2021-0356

[15] Wu, J., Yang, T., Zhou, Z., & Zhao, N. (2023). Consumers' affective needs matter: Open innovation through mining luxury hotels' online reviews. *International Journal of Hospitality Management*, 114, 103556. https://doi.org/10.1016/j.ijhm.2023.103556

[16] Xiang, Z., Schwartz, Z., Gerdes, J. H., Jr., & Uysal, M. (2015). What can big data and text analytics tell us about hotel guest experience and satisfaction? *International Journal of Hospitality Management*, 44, 120–130. https://doi.org/10.1016/j.ijhm.2014.10.013

[17] Xu, J., Wang, X., Zhang, J., Huang, S., & Lu, X. (2022). Explaining customer satisfaction via hotel reviews: A comparison between pre- and post-COVID-19 reviews. *Journal of Hospitality and Tourism Management*, 53, 208–213. https://doi.org/10.1016/j.jhtm.2022.11.003

[18] Yang, H., Song, H., Cheung, C., & Guan, J. (2021). How to enhance hotel guests' acceptance and experience of smart hotel technology: An examination of visiting intentions. *International Journal of Hospitality Management*, 97, 103000. https://doi.org/10.1016/j.ijhm.2021.103000

[19] Yu, M., Cheng, M., Yang, L., & Yu, Z. (2022). Hotel guest satisfaction during COVID-19 outbreak: The moderating role of crisis response strategy. *Tourism Management*, 93, 104618. https://doi.org/10.1016/j.tourman.2022.104618

[20] Zarezadeh, Z. Z., Rastegar, R., & Xiang, Z. (2022). Big data analytics and hotel guest experience: A critical analysis of the literature. *International Journal of Contemporary Hospitality Management, 34*(6), 2320–2336. https://doi.org/10.1108/IJCHM-10-2021-1293

[21] Zhang, X., Yu, J., & Kim, H. S. (2024). Exploring customers experience and satisfaction with theme hotels: A cross-cultural analysis of hotel reviews. *Sage Open, 14*(2), 21582440241255827. https://doi.org/10.1177/21582440241255827