

Data Driven Amenity Strategies: Evidence from Online Reviews and Hotel Ratings in Bali

Prastha Adyatma^{1*}, Luh Putu Kartini², Ni Luh Dita Priliani³
Politeknik Pariwisata Bali, Indonesia

Corresponding Author: Prastha Adyatma prasthaadyatma@gmail.com

ARTICLE INFO

Keywords: Data Driven, Amenity Strategies, Online Reviews, Hotel Ratings

Received : 08, November

Revised : 29, November

Accepted: 28, December

©2025 Adyatma, Kartini, Priliani: This is an open-access article distributed under the terms of the [Creative Commons Attribution 4.0 International](https://creativecommons.org/licenses/by/4.0/).



ABSTRACT

This study translates large-scale online reviews into actionable guidance for hotel amenity investment in Bali. We quantify how specific amenities relate to guest ratings using three complementary lenses: (i) bivariate rating uplift (Welch's t, effect sizes, confidence intervals) to estimate standalone associations, (ii) a coverage-impact matrix that maps market prevalence against rating uplift to categorize amenities into INVEST, MAINTAIN, HYGIENE, or NICETOHAVE, and (iii) interpretable machine learning (Random Forest with modelagnostic importance and optional SHAP) to validate which amenity bundles best separate highrated properties (≥ 4.7). Additionally compare resort vs. city segments to reflect contextdependent preferences. Results indicate a consistent "core utilities + family readiness + convenience" bundle free Wi Fi, kid friendly facilities, laundry service, air conditioning, parking, and onsite restaurant as the strongest predictors of high ratings, with outdoor pool, free breakfast, spa, beach access, and airport shuttle providing further differentiation. The coverage impact matrix highlights where to allocate capital: for example, high impact/low coverage features fall into INVEST, while widely available, reliabilitycritical utilities align with MAINTAIN. discuss managerial implications for amenity roadmapping and budgeting, note the study's associational nature and outline robustness steps to enhance generalizability.

INTRODUCTION

Hotel choice is increasingly filtered by electronic word-of-mouth (eWOM), where large volumes of ratings and reviews compress complex experiences into simple signals managers can mine for action. Evidence shows that analyzing online reviews reveals which experience attributes most shape satisfaction and demand, and that review networks themselves exhibit structures that influence visibility and persuasion hence the need to translate review data into actionable hotel attributes (Kim et al., 2022; Pestana et al., 2024).

However, amenity investments are costly and their effects are asymmetric: some features behave like “hygiene factors” (expected but not differentiating), while others deliver genuine gains in satisfaction and ratings. Research using sentiment/KanoIPA frameworks on hotel reviews demonstrates bidirectional (nonlinear) attribute performance supporting a data driven approach that measures both coverage (how common an amenity is) and impact (its rating uplift) to prioritize spending. This motivates our Coverage vs Impact matrix and bivariate uplift table (Chen et al., 2022).

Preferences are also context-dependent: resorts face different guest expectations than city hotels, and spatial dynamics shape what counts as a differentiator versus a baseline. Studies on resort/hotel network attention and coastal/inland topic differences empirically validate segmentspecific prioritization hence our parallel breakdowns for resort vs. city hotels using the same metrics (Sun et al., 2024; Contu et al., 2024).

Finally, to go beyond simple averages and capture interactions among amenities, interpretable machine learning offers a predictive lens: feature importance summaries indicate which amenity bundles most strongly separate “highrated” hotels (≥ 4.7). Recent open access work on ML/sentiment pipelines for hotel reviews shows how such models can support managerial decisions and sustainable service design justifying our Top Features by Model Importance output to complement the coverage/impact and uplift views (Erdoğan et al., 2025).

Research Gap

First, while a large body of hospitality research mines online reviews to surface drivers of satisfaction, most studies prioritize textcentric sentiment and topic extraction over structured, actionable amenity prioritization. Methods based on Kano/IPA or related asymmetric frameworks show that attributes influence satisfaction nonlinearly, yet few works translate these effects into a simple coverage–impact decision matrix that directly guides amenity investment using observed amenity presence and rating uplift. In short, know “what guests talk about,” but still lack practical tools to rank which amenities to add, maintain, or deprioritize using market coverage and rating impact in tandem. (Chen et al., 2022; Kim et al., 2022).

Second, segment heterogeneity is underaddressed: resorts and city hotels face different guest expectations and seasonality structures, but many analyses pool properties, diluting segment specific signals. Recent studies show distinct spatiotemporal patterns in resort attention and demand, and a Bali focused study

analyzes luxury resorts only leaving a gap in parallel, comparable resort vs city amenity priorities within the same destination. Consequently, managers lack segment specific guidance on which amenities truly differentiate in each context. (Sun et al., 2024; Williady et al., 2022).

Third, machine learning has been used to classify sentiment or discover topics, but its integration as a transparent prioritization lens for amenities remains limited. Specifically, studies seldom pair a predictive, interpretable model with coverage impact and bivariate uplift views to triangulate robust priorities. This leaves a gap for a combined framework that (i) quantifies coverage and impact, (ii) verifies amenity influence via predictive separation of high performers, and (iii) preserves managerial interpretability. (Pestana et al., 2024; Li et al., 2025).

LITERATURE REVIEW

Electronic word of mouth (eWOM) has become central to hotel choice, with ratings and reviews acting as compressed signals of service quality and experience. Bibliometric and empirical syntheses show that online review research has matured into a coherent field, emphasizing how review networks shape visibility and how specific experience attributes drive satisfaction making eWOM a practical data source for managerial decisionmaking. (Pestana et al., 2024).

A major stream uses text mining and asymmetric evaluation models (e.g., Kano/IPA) to demonstrate that hotel attributes exert nonlinear, bidirectional effects on satisfaction: some features behave as “hygiene” (expected, low differentiation) while others trigger delight and rating gains. These findings support translating attribute effects into actionable prioritization tools rather than reporting sentiment alone. (Chen et al., 2022).

Beyond text, amenity prevalence itself can be informative: combining how **common** an amenity is in the market (coverage) with how much it shifts average ratings (impact/uplift) offers a simple decision matrix for capital allocation (invest, maintain, or deprioritize). Prior work on attribute asymmetries and improvement strategies implies such a matrix can operationalize review insights for property level budgeting and product design. (Chen et al., 2022; Pestana et al., 2024).

Segment context matters. Resorts and city hotels face different guest expectations and spatiotemporal attention patterns; what differentiates in a resort may be baseline in an urban property. Recent open access studies on resort hotel network attention and market dynamics underscore the need to analyze priorities by segment rather than pooling all properties, particularly in destinations with diverse sub markets like Bali. (Sun et al., 2024).

Bali specific research using large scale online reviews confirms the feasibility and value of destination focused analytics, showing how post COVID satisfaction drivers can be extracted from Google Travel reviews for luxury resorts. Complementary work in Bali’s spa hotel niche connects psycholinguistic signals to guest experience, reinforcing the destination as a fertile context for applied review analytics. (Williady et al., 2022; Aisha & Hsu, 2025).

Machine learning (ML) extends these insights by capturing interactions among amenities and producing feature importance rankings that help explain what separates high rated hotels. Deep learning pipelines for hotel reviews improve sentiment inference and managerial targeting; related strands address fake review detection to protect the integrity of review signals that underlie such models. Together, these approaches justify pairing coverage–impact matrices and bivariate uplifts with interpretable ML feature importance. (Erdoğan et al., 2025; Gupta et al., 2024; Chang & Chen, 2020).

METHODOLOGY

Quantitative, observational study using secondary data from online listings/reviews of hotels in Bali. The unit of analysis is the property (one row per listing). Estimate associations (not causal effects) between amenity availability and rating outcomes, then triangulate with a predictive model for interpretability and prioritization.

Bivariate Rating Uplift (Difference in Means). Estimate each amenity's direct effect on ratings by comparing the average star score of hotels that have the amenity versus those that do not; to handle unequal variances and unbalanced group sizes typical in observational data, rely on the Welch t test, which is recommended over the classic Student test in such settings and is commonly applied in hospitality comparisons. (West, 2021; Delacre et al., 2017; Moreno Perdigón et al., 2021).

Coverage Impact Matrix (Amenity Prioritization). Convert uplift results into a manager friendly coverage impact map: the Xaxis shows coverage (market prevalence of an amenity) and the Y axis shows impact (rating uplift), yielding four action buckets INVEST, MAINTAIN, HYGIENE, and NICE TO HAVE that align with importance performance style prioritization used in hospitality and with asymmetric (Kano/IPA) findings that amenity effects are not linear (Chen et al., 2022; Çobanoğlu et al., 2010).

Predictive Feature Importance for High Rated Classification. To validate which amenities jointly differentiate high rated hotels, train a Random Forest classifier and report model agnostic feature importance and when needed, SHAP (SHapley Additive exPlanations) for interpretable contributions; this approach is standard for explaining tree ensembles and is widely used with hotelreview ML pipelines. (Breiman, 2001; Strobl et al., 2007; Lundberg & Lee, 2017; Wen et al., 2023; Erdoğan et al., 2025).

RESEARCH RESULT AND DISCUSSION

This section presents three complementary views of how amenities relate to hotel performance in Bali. First, report model based feature importance for predicting high rated properties (≥ 4.7), highlighting which amenities most strongly separate top performers (Figure 1). Second, show bivariate ratings uplift the mean difference in star ratings between hotels versus without each amenity to provide an intuitive, directionally clear effect size (Figure 2). Third, translate those effects into a Coverage Impact matrix that maps market coverage (prevalence on the Xaxis) against impact (rating uplift on the Yaxis), yielding four

action buckets INVEST, MAINTAIN, HYGIENE, and NICETHAVE (Figure 3). Together, these views allow us to (i) identify amenities that consistently matter in multivariate settings, (ii) verify their standalone associations with ratings, and (iii) prioritize investment by balancing impact with how common each amenity already is in the market. Unless stated otherwise, effects are associative and results are interpreted alongside segment context and data quality checks to ensure practical, manager ready guidance.

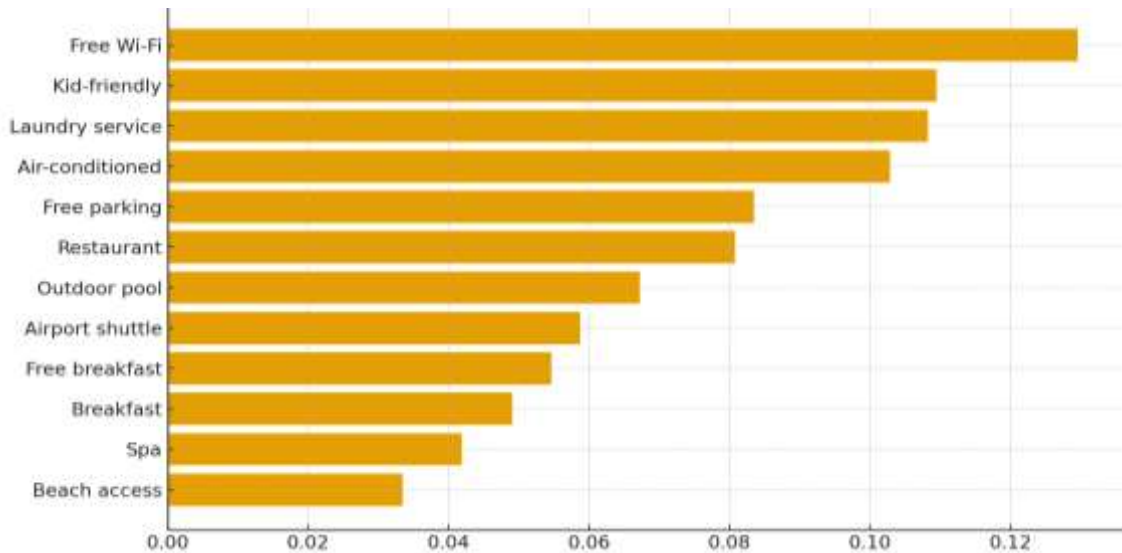


Figure 1: Top Features By Model Importance
Source: Machine Learning Data Analysis

Figure 1 shows the model based feature importance from a Random Forest classifier predicting whether a hotel is highrated (≥ 4.7), indicating which amenities most strongly separate top performers after accounting for interactions among features. The ranking highlights a “core utility + family readiness + convenience” bundle: Free WiFi dominates, followed by Kidfriendly and Laundry service, then Airconditioned, Free parking, and Restaurant, with Outdoor pool, Airport shuttle, (Free) breakfast, Spa, and Beach access also contributing. Managerially, this implies safeguarding reliability for essential utilities (WiFi, AC, parking), reinforcing family/longstay readiness (kidfriendly, laundry), and enhancing onsite completeness (restaurant, pool), while value adds (airport shuttle, free breakfast) can nudge properties into the highrated tier. Because feature importance reflects predictive contribution not causality credit can be split when amenities cooccur (resorts with both pools and spas), so permutation importance or SHAP can be used to validate stability and interpret directionality at both global and perhotel levels (Breiman, 2001; Strobl et al., 2007; Lundberg & Lee, 2017).

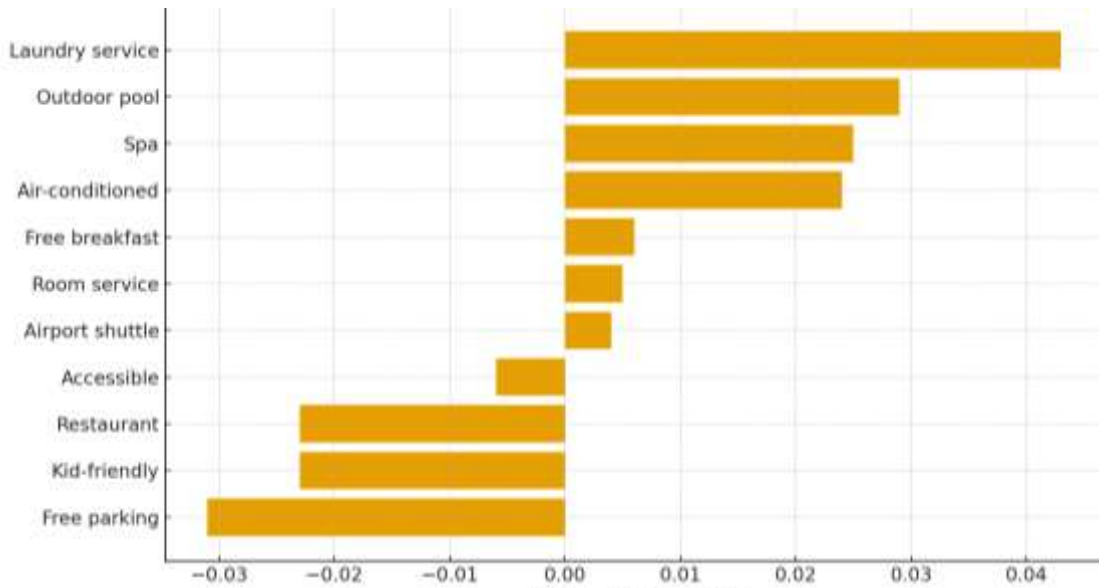


Figure 2: Top Amenities by Rating Uplift

Source: Machine Learning Analysis

Figure 2 reports the bivariate rating uplift for each amenity mean difference in star ratings between hotels with the amenity and those without it so positive bars indicate higher average ratings when the amenity is present. In our results, the largest positive uplifts are for Laundry service, Outdoor pool, Spa, Air conditioned, and Free breakfast, suggesting these features are most consistently associated with better guest evaluations, smaller positives appear for Airport shuttle and Room service. Conversely, a few amenities show near zero or negative uplifts (e.g., Free parking, Kidfriendly, Restaurant, Accessible) signals that these may function more as hygiene factors (expected but weakly differentiating) in this market. Managerially, Figure 2 offers a simple, directional screen to spot quick wins (high uplift) and baseline utilities (low/negative uplift) before deeper modeling. Because this is an observational, two group comparison, Welch's t test (unequal variances) with false discovery rate control for multiple amenities and interpreting signs in light of known asymmetric attribute effects (Kano/IPA), where the same feature can behave like "delighter" in some contexts but only a hygiene factor in others (Delacre et al., 2017; Chen et al., 2022).

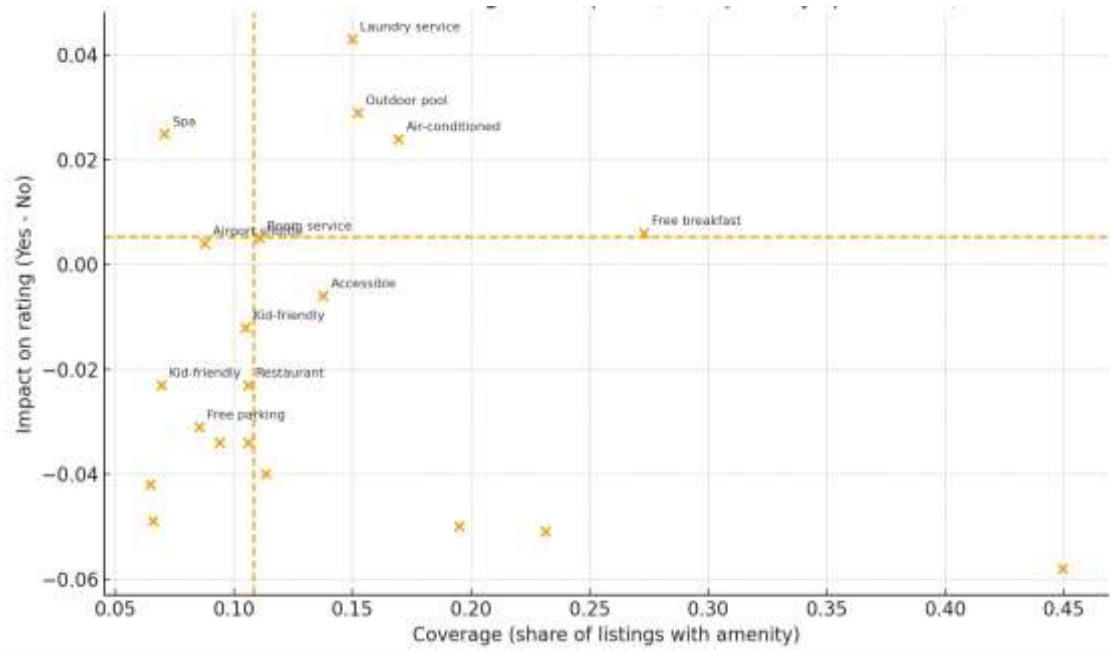


Figure 3: Amenities: Coverage vs Impact
 Source: Machine Learning Analysis

Figure 3 visualizes the Coverage Impact matrix that prioritizes amenities by plotting market coverage (share of listings that offer the amenity) on the X-axis and impact (bivariate rating uplift, Yes No) on the Y axis, with dashed lines at the median coverage and upper quartile (Q3) impact to form four action quadrants. Amenities in the INVEST zone (high impact, low coverage) are underrepresented yet strongly associated with higher ratings here, Spa sits clearly in this quadrant, signaling a differentiator worth adding where feasible. MAINTAIN items (high impact, high coverage) already contribute meaningfully and should be protected for consistency and quality e.g., Laundry service, Outdoor pool, Air conditioned, and Free breakfast. HYGIENE (low impact, high coverage) features, such as Accessible, are common but weakly differentiating and should be treated as baseline standards rather than investment drivers. Finally, NICETOHAVE (low impact, low coverage) items e.g., Free parking, Kidfriendly, Restaurant in this sample show little uplift and lower prevalence, implying lower priority unless required by positioning or segment needs. This matrix translates asymmetric attribute effects from review analytics into a practical, important performance style decision tool for capital allocation and product design. (Chen et al., 2022; Çobanoğlu et al., 2010).

CONCLUSIONS AND RECOMMENDATIONS

Across three complementary lenses, the data point to a clear pattern. The modelbased importance view (Figure 1) shows a “core utility + family readiness + convenience” bundle driving high ratings: Free WiFi, Kidfriendly, Laundry service, Airconditioned, Free parking, Restaurant, with Outdoor pool, Airport shuttle, (Free) breakfast, Spa, and Beach access also contributing. The bivariate rating uplift view (Figure 2) highlights the strongest positive average gains for Laundry service, Outdoor pool, Spa, Airconditioned, and Free breakfast, while

amenities like Free parking, Kidfriendly, Restaurant, and Accessible show nearzero/negative uplifts behaving more like hygiene factors in this market. Translating these signals into the Coverage Impact matrix (Figure 3) yields practical priorities: INVEST in under supplied, high impact differentiators (notably Spa); MAINTAIN high impact, widely adopted features (Laundry service, Outdoor pool, Airconditioned, Free breakfast); treat common, low impact items as HYGIENE (Accessible) and keep NICE TO HAVE items (Free parking, Kid friendly, Restaurant) as lower priority unless required by positioning or segment needs, protect reliability on core utilities, strengthen family/longstay readiness, and add selective experience enhancers especially those sitting in the INVEST quadrant to most efficiently move the needle on ratings.

ADVANCED RESEARCH

The advanced research agenda emerging from this study highlights the need for a unified, multi-lens computational framework capable of transforming large-scale online review ecosystems into precise amenity investment intelligence for hotel markets with strong spatiotemporal heterogeneity such as Bali. Future research should integrate causal inference techniques—such as doubly robust estimators, propensity score weighting, or causal forests—to move beyond associative uplift and approximate the true marginal effects of specific amenities under observational constraints. Additionally, granular temporal modeling of review dynamics, including seasonality shifts, post-event sentiment shocks (e.g., pandemics or destination crises), and cross-platform rating discrepancies, can reveal how amenity importance evolves over time. Machine learning pipelines should be expanded with multimodal inputs (text, images, and metadata) to capture latent experience attributes not represented in structured amenity listings, while interpretable AI approaches (global-local SHAP, surrogate models) can refine managerial actionability. Segment-aware spatial models—leveraging geospatial embeddings or region-specific random effects—could further differentiate urban, coastal, and resort submarkets, enabling hyper-localized amenity roadmapping. Collectively, these advancements will push hospitality analytics toward a more predictive, situationally adaptive, and causally credible decision-support system.

REFERENCES

- Aisha, N., & Hsu, C.C. (2025). Reputation and guest experience in Bali's spa hotels: A LIWCbased analysis of online reviews. *Tourism and Hospitality*, 6(4), 180. <https://doi.org/10.3390/tourhosp6040180> [MDPI](#)
- Breiman, L. (2001). Random forests. *Machine Learning*, 45(1), 5–32. <https://www.stat.berkeley.edu/~breiman/randomforest2001.pdf> [stat.berkeley.edu](https://www.stat.berkeley.edu)
- Chang, Y.C., & Chen, C.W. (2020). Using deep learning and visual analytics to explore hotel reviews and management responses. *Tourism Management*, 80, 104129. <https://doi.org/10.1016/j.tourman.2020.104129> [ScienceDirect](#)

- Chen, Y., Zhong, Y., Yu, S., Xiao, Y., & Chen, S. (2022). Exploring bidirectional performance of hotel attributes through online reviews based on sentiment analysis and KanoIPA model. *Applied Sciences*, 12(2), 692. <https://www.mdpi.com/20763417/12/2/692> MDPI
- Çobanoğlu, C., Berezina, K., Kasavana, M. L., & Erdem, M. (2010). Importance Performance Analysis of Guest Entertainment Technology Amenities. *FIU Hospitality Review*. <https://digitalcommons.fiu.edu/cgi/viewcontent.cgi?article=1489&context=hospitalityreview> digitalcommons.fiu.edu
- Contu, G., Ceddia, M. G., & Naitza, S. (2024). Online reviews explain differences in coastal and inland hotel topics. *PLOS ONE*, 19(7), e0303844. <https://doi.org/10.1371/journal.pone.0303844> PMC
- Delacre, M., Lakens, D., & Leys, C. (2017). Why psychologists should by default use Welch's t test instead of Student's t test. *International Review of Social Psychology*, 30(1), 92-101. <https://ripsirsp.com/articles/10.5334/irsp.82> ripsirsp.com
- Erdoğan, D., Kayakuş, M., Çelik Çaylak, P., Ekşili, N., Moiceanu, G., Kabas, O., & Ichimov, M. A. M. (2025). Developing a deep learning-based sentiment analysis system of hotel customer reviews for sustainable tourism. *Sustainability*, 17(13), 5756. <https://www.mdpi.com/20711050/17/13/5756> MDPI
- Gupta, D., Rane, G., & Rane, S. (2024). Deep learning-based truthful and deceptive hotel reviews detection. *Sustainability*, 16(11), 4514. <https://doi.org/10.3390/su16114514> MDPI
- Kim, Y.J., Kim, B., & Leem, B. (2022). The impact of hotel customer experience on customer satisfaction through online reviews. *Sustainability*, 14(2), 848. <https://doi.org/10.3390/su14020848> MDPI
- Li, J., Lee, B., & Kim, J. (2025). Analyzing the asymmetric effects of COVID19 on hotel selection attributes and customer satisfaction through AIPA. *Sustainability*, 17(19), 8546. <https://doi.org/10.3390/su17198546> MDPI
- Lundberg, S. M., & Lee, S.I. (2017). A unified approach to interpreting model predictions (SHAP). *NeurIPS 2017*. <https://arxiv.org/abs/1705.07874> arXiv
- MorenoPerdigón, M. C., MartínSantana, J. D., & BeerliPalacio, A. (2021). Guest satisfaction in independent and chain-affiliated hotels: A comparative analysis using t tests. *International Journal of Hospitality Management*, 97, 102992. (Abstract) <https://www.sciencedirect.com/science/article/abs/pii/S0278431920303649> ScienceDirect
- Pestana, M. H., Gageiro, M., Santos, J. A. C., & Santos, M. C. (2024). Network structure of online customer reviews and online hotel reviews: A systematic literature review. *Information*, 15(6), 334. <https://doi.org/10.3390/info15060334> MDPI
- Strobl, C., Boulesteix, A.L., Zeileis, A., & Hothorn, T. (2007). Bias in random forest variable importance measures: Illustrations, sources and a solution. *BMC*

- Bioinformatics, 8, 25.
<https://pmc.ncbi.nlm.nih.gov/articles/PMC1796903/PMC>
- Sun, H., Zhang, Y., & Guo, W. (2024). Spatiotemporal characteristics and influencing factors of network attention to resort hotels in China. *Heliyon*, 10(15), e35314. <https://doi.org/10.1016/j.heliyon.2024.e35314> [Cell](#)
- Wen, Y., Li, J., & Zhang, L. (2023). Sentiment analysis of hotel online reviews using the BERT model. *PLOS ONE*, 18(3), e0282487. <https://pmc.ncbi.nlm.nih.gov/articles/PMC10004568/> [PMC](#)
- West, R. M. (2021). Best practice in statistics: Use the Welch ttest when testing the difference between two groups. *Annals of Clinical Biochemistry*, 58(4), 267–269. <https://pubmed.ncbi.nlm.nih.gov/33562996/>
- Williady, A., Wardhani, H. N., & Kim, H.S. (2022). A study on customer satisfaction in Bali's luxury resort utilizing big data through online reviews. *Administrative Sciences*, 12(4), 137. <https://doi.org/10.3390/admsci12040137>