



# Investigating Key Attributes in Experience and Satisfaction of Hotel Customer Using Online Review Data: A Case Study of Five-Star Hotels in Colombo, Sri Lanka

Sewwandi Dulakshika Jayasekara<sup>1</sup>, Priska Tamang<sup>1</sup>, Ani Chioma Favour<sup>1</sup>, Jue Wang<sup>2\*</sup>

<sup>1</sup> Department of Global Business, Kyungsoong University, Busan, Republic of Korea

<sup>2</sup> School of Global Studies, Kyungsoong University, Busan, Republic of Korea

\* Corresponding author: wangjue@ks.ac.kr

## ARTICLE INFO

*Keywords:*

big data;  
customer satisfaction;  
luxury hotels;  
online reviews;  
text mining

## ABSTRACT

The rapid growth of online travel platforms has generated a substantial amount of user-generated content, which offers valuable insights into customer experiences in the hospitality industry. This research examines the key factors of customer satisfaction in five-star hotels using online review data. It focuses on four five-star hotels in Colombo, Sri Lanka, based on a dataset of 66,314 Google reviews. To analyze and identify the most prominent themes in customer feedback, a big data analytics approach was employed, combining text mining techniques with keyword frequency analysis and semantic network analysis. The findings show that service quality, staff friendliness, cleanliness, and overall hotel environment are frequently mentioned elements that influence customer satisfaction. The network analysis further demonstrates a strong relationship between service-related attributes and positive guest experiences, highlighting the central role of service interactions. The study contributes to hospitality and tourism literature by demonstrating the usefulness of big data analytics in extracting meaningful insights from online reviews to gain a better understanding of customer behavior. The findings suggest practical implications for hotel managers seeking to improve service quality, enhance customer satisfaction, and strengthen their competitive advantage through data-driven decision-making.

## 1. Introduction

The global hospitality industry has been fundamentally changed by digitalization, with technology altering how travelers search for hotels, evaluate quality, and share customer experiences (Buhalis & Law, 2008). Tourism acts as a major economic driver worldwide (World Tourism Organization, 2023), and customer-generated online content has become a strategic resource for enhancing service quality and maintaining competitiveness (Xiang et al., 2017).

Among various types of user-generated content, online reviews are especially influential. Travelers consider that reviews are more trustworthy than traditional marketing and advertising. They rely on reviews to reduce uncertainty before booking (Cheng et al., 2019). Electronic word of mouth also strongly shapes their decisions and expectations (Bronner & de Hoog, 2011). The large amounts of reviews have created big data opportunities: text mining and keyword network analysis allow researchers to extract authentic, ex-

perience-based insights from unstructured customer feedback (Ordenes et al., 2014).

Customer satisfaction is a key concept in hospitality research, because it has a strong impact on loyalty and financial performance (Osman & Sentosa, 2013). In hotels, satisfaction is impacted by multiple attributes like staff behaviour, cleanliness, facilities, food quality and also on emotional experiences (Cheng et al., 2019). Most satisfaction related studies focus on well-established destinations, leaving emerging tourism markets under research (Fortanier & van Wijk, 2010).

Sri Lanka is a relevant emerging context. The tourism industry is an important contributor to the economy. Colombo hosts several international five-star hotels that serve business and leisure travellers (World Tourism Organization, 2023). Despite this, no large-scale study has analyzed customer satisfaction drivers for these hotels using big data analytics from online reviews.

To address this gap, this research adopts an integrated framework. It combines expectancy disconfirmation (Abrate et al., 2021) and customer delight (Ladhari, 2009). Luxury hotel satisfaction arises from both functional (service, cleanliness, food) and emotional triggers (surprise, aesthetic pleasure). This study uses 66,314 Google Reviews of four five-star hotels in Colombo, and applies keyword frequency analysis and semantic network analysis to identify the most prominent themes and also to attribute co-occurrence patterns. The study has theoretical contribution (extending satisfaction frameworks to an emerging market), methodological contribution (providing a replicable big data pipeline for hospitality research), and practical contribution (offering data driven recommendations for hotel managers in Colombo).

## 2. Literature Review

### 2.1. Customer Satisfaction in Hospitality

Customer satisfaction is a main concept in hospitality research because it has a direct impact on loyalty, word-of-mouth, and also financial performance (Marcos & Coelho, 2022). The most widely used theories for explaining satisfaction in the service sector is the expectancy-disconfirmation paradigm (Schiebler et al., 2025): satisfaction occurs when perceived service performance meets or exceeds prior expectations. Satisfaction is multidimensional, shaped by functional factors (e.g., room cleanliness, staff responsiveness) and emotional experiences (Sukhu et al., 2019).

Hotel services are intangible, customers tend to rely heavily on interpersonal interactions and physical surroundings to evaluate quality of the hotel (Ali et al., 2014). Torres (2014) reviewed decades of research and identified six main drivers that impact on satisfaction: service quality, staff professionalism, physical facilities, room comfort/cleanliness, food & beverage quality, and emotional engagement.

### 2.2. Online Reviews as a Data Source for Understanding Satisfaction

The rapid growth of online review platforms like Google Reviews and TripAdvisor has changed how travelers search for hotels and share their experiences. Unlike traditional surveys, online reviews are written freely and real experience-based narratives that capture authentic and honest customer opinions (Sukhu et al., 2019). Reviews are considered as more trustworthy than corporate marketing because they come from previous customers who have real experiences and act as an electronic word-of-mouth (eWOM) (Sukhu et al., 2019). Trust in online reviews plays a key role in their influence. Nisar et al. (2020) found that perceived credibility and usefulness of reviews strongly shape travelers' decisions. Sparks & Browning (2011) showed that eWOM is especially powerful for experience-based services. Social media platforms help to spread reviews quickly to a wide audience, while making real-time reputation management essential for

hotels (Zeng & Gerritsen, 2014). For researchers, online reviews provide a large-scale, non-reactive data source. Xiang et al. (2017) demonstrated that text mining of reviews can reveal service features that matter most to customers, often supporting or challenging survey-based findings. Tolkach et al. (2024) argued that review-based research offers higher ecological validity and it captures unprompted feedback in real-life contexts.

### 2.3. Hotel Customer Satisfaction Dimensions

Empirical studies based on online review data have consistently highlighted a core set of customer satisfaction elements, though the importance of these elements can vary depending on hotel class and cultural context. Service quality and staff behavior are the most commonly mentioned attributes in both luxury and budget hotels (Zeng & Gerritsen, 2014). Positive reviews often include words like "friendly," "professional," and "attentive" to describe staff and also negative reviews frequently mention slow service, rude employees, or unhelpful reception staff members (Wirtz & Jerger, 2016). Physical environment and cleanliness both were important after COVID-19. Glaveli et al. (2023) found that "cleanliness" is the strongest predictor of overall satisfaction in five-star hotels, even more important than location. Ding et al. (2025) found that guests now mention hygiene-related terms like "sanitized" and "well-maintained" much more often than before 2020.

Food and beverage experience is another key factor. Suwanamas et al. (2015) found that dining quality strongly impacts overall satisfaction in full-service luxury hotels. Reviews often mention breakfast variety, freshness of buffet, and atmosphere of hotel (Suwanamas et al., 2015). Emotional and experiential factors such as "amazing view," "memorable stay," or "wonderful atmosphere" indicate positive customer perceptions. Berman (2005) suggested that actual delight occurred when employees exceed customer expectations in unexpected ways. These emotional aspects are very difficult to capture and measure through structured surveys but are clearly visible in detailed and narrative reviews. Location and convenience are also significant factors, though their salience varies. For business travelers in city centers such as Colombo, proximity to airports, convention centres, and shopping malls are highly valued (Wirtz & Jerger, 2016).

### 2.4. Online Data Mining in Hospitality Research

The large amounts of online reviews have made manual analysis impractical and also very difficult. As a result, text mining methods have become widely used, including keyword frequency analysis, sentiment analysis, and topic modeling (Kwon et al., 2021; Lu et al., 2021). Sentiment analysis quantifies to measure positivity and negativity of options, while topic modeling such as latent Dirichlet allocation uncovers latent themes within a large set of text (Kwon et al., 2021).

Semantic (keyword) network analysis is a less commonly used but powerful method that focuses on visualizing co-occurrence relationships among words (Suwanamas et al., 2015). Unlike simple frequency lists, this method reveals relationships between different terms. For example, if words like “staff” and “friendly” consistently appear together, the network shows a strong connection between them, while indicating an associative link in customers’ mentality (Drieger, 2013). This approach is useful for understanding how customers integrate different service aspects into an overall impression.

Several research gaps still remain. First, most studies have focused on well-established countries like Europe, North America and emerging tourism markets like Sri Lanka have very low attention (Zeng & Gerritsen, 2014). Second, there is a limited amount of research on five-star hotels in South Asia. Third, few studies have applied semantic network analysis to Google Reviews and mostly use platforms like TripAdvisor or Expedia. Fourth, the relationship between functional (cleanliness, service) and emotional (delight, amazement) attributes in luxury hotel reviews has not been thoroughly studied.

### 2.5. Theoretical Framework and Research Gap

This study adopts an integrated framework that combines the expectancy-disconfirmation model (Carragher-Wolverton & Hirschheim, 2023) with the experiential view of customer delight (Torres et al., 2014) to understand customer satisfaction. The study suggests that satisfaction in luxury hotels arises from both functional attributes (service, cleanliness, food quality) and emotional factors (positive surprise, aesthetic pleasure).

Expectancy-Disconfirmation Theory (EDT) explains customer satisfaction as the result of comparing actual performance with prior expectations. Satisfaction occurs when expectations are met or exceeded, while dissatisfaction arises when performance falls below expectations.

In the hotel industry, EDT is widely used to understand guest evaluations of service quality, facilities, and overall experiences. Online reviews reflect these evaluations, as guests often comment on aspects such as cleanliness, service, location, value, and comfort. Therefore, EDT provides a useful framework for interpreting customer satisfaction expressed in online hotel reviews.

Customer delight theory suggests that delight goes beyond ordinary satisfaction. Delight occurs when customers experience performance that greatly exceeds expectations and evokes positive emotions such as joy, excitement, or surprise. While satisfaction results from meeting expectations, delight arises when expectations are exceeded in an exceptional way.

In hospitality, delight is often created through personalized service, memorable staff interactions, and outstanding amenities. Online reviews help identify delight because

guests frequently use emotional terms such as “amazing,” “wonderful,” and “exceptional.” Thus, Customer Delight Theory complements EDT by capturing the emotional aspects of hotel customer experiences.

Although previous hospitality studies have extensively applied text mining, sentiment analysis, and online review analytics to examine customer satisfaction and hotel performance, most of this research has focused on developed tourism markets and has primarily investigated general service quality dimensions. Limited attention has been given to how sustainability-related attributes influence customer choice behavior within luxury hotel settings in emerging tourism destinations. Furthermore, existing studies often examine customer satisfaction and sustainability independently rather than exploring how sustainability themes embedded in online reviews shape customers’ hotel selection decisions. This study addresses these gaps by integrating sustainability-oriented review analytics with customer choice behavior analysis in the context of Sri Lankan luxury hotels. By focusing on an emerging tourism market characterized by unique cultural, environmental, and service dynamics, the research also makes both theoretical contributions (testing the integrated framework in an emerging tourism destination) and practical implications (offering data-driven recommendations for hotel managers).

## 3. Methodology

### 3.1. Data Collection

The data for this study was obtained from Google Reviews, a widely used online platform that allows customers to share their hotel experiences through star ratings and written feedback. This platform offers publicly accessible data that combines both qualitative and quantitative information, making it suitable for analyzing customer satisfaction and service quality in the hospitality sector.

Colombo was selected as the study location because it is Sri Lanka's commercial capital and a major hub for business and urban tourism. According to the Sri Lanka Tourism Development Authority (SLTDA), the Western Province, where Colombo is located, accounts for 34.3% of the country's accommodation capacity and contains the highest concentration of upscale business hotels and airport transit accommodations (SLTDA, 2025). Furthermore, Colombo district recorded the largest accommodation inventory in Sri Lanka, with 11,398 rooms (SLTDA, 2025).

The study focused on five-star hotels because they represent the highest category of classified tourist accommodation in Sri Lanka. According to SLTDA statistics, Sri Lanka had 33 classified five-star hotels in 2025 (SLTDA, 2025). An examination of the official accommodation directory identified eleven five-star hotels operating in Colombo at the time of data collection. However, the final sample was restricted to four hotels with the largest and most complete Google Review datasets, while hotels with comparatively limited re-

view volumes were excluded. This approach was adopted because text mining and semantic network analysis require substantial volumes of user-generated textual data. The selected four hotels generated 66,314 customer reviews, providing a sufficiently large dataset for robust text-mining and semantic network analyses of customer satisfaction drivers in Colombo's luxury hotel sector.

The review data was collected using Instant Data Scraper, a Chrome browser extension that facilitates automated extraction of web data without the need for coding. The scraper was applied to each hotel's Google Review page to collect the following information: hotel name, reviewer name or anonymized identifier, review date, star rating, review text, and where available, subcategory ratings (such as Rooms, Service, and Location). Reviews not written in English were

excluded to ensure consistency for text mining and sentiment analysis. The use of a single language reduced potential translation errors and semantic inconsistencies that may arise when analyzing multilingual user-generated content. English was selected because it is widely used by international travelers and represents a substantial proportion of online hotel reviews. Restricting the dataset to English-language reviews therefore enhanced the reliability and comparability of textual analysis across observations. Duplicate entries and incomplete records were removed during the data cleaning process. An example of a review collected for this study is shown in Figure 1, illustrating typical data elements including the reviewer's name, review date, star rating, detailed feedback, and sub-category scores.

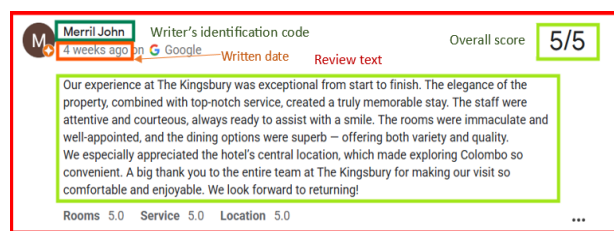


Figure 1. Example of Google review collected for The Kingsbury Colombo.

In total, 4 five-star hotels located in Colombo, Sri Lanka, were selected for this study based on the official listings of the Sri Lanka Tourism Development Authority (SLTDA, 2025). A total of 66,314 Google Reviews were collected from

these hotels, providing a large dataset for the analysis of customer experience and satisfaction. The number of reviews collected for each hotel is shown in Table 1.

Table 1. Selected five-star hotels in Colombo and number of Google Reviews collected

Rank	Hotel Name	Number of google reviews
01	The Kingsbury Colombo	19178
02	Shangri-La, Colombo	16029
03	Taj-Samudra Colombo	11897
04	Cinnamon Grand Colombo	19210

### 3.2. Data Analysis

As for the data analysis of this study, first, the text mining technique was applied to obtain the word frequency from online hotel reviews. As part of the text-mining process, the review texts were preprocessed to improve data quality and analytical reliability. Common stopwords (e.g., articles, prepositions, and pronouns), duplicate terms, and non-informative words unrelated to hotel experiences were removed from the dataset. In addition, semantically equivalent keywords and minor lexical variations were standardized and grouped under a common representative term to improve consistency and reduce redundancy in the dataset. This process included the consolidation of singular and plural forms, spelling variations, and closely related expressions that conveyed the same meaning within the context of hotel customer

experiences. Keywords with limited interpretive value or extremely low occurrence frequencies were excluded from further analysis. To minimize researcher subjectivity, the keyword refinement process followed predefined preprocessing rules that were applied consistently across all review data.

The remaining keywords were then ranked according to their frequency of occurrence across all reviews. To enhance interpretability and reduce noise in the semantic network, only the most frequently occurring keywords were retained for subsequent analysis. This filtering procedure resulted in a final set of 97 keywords, which formed the basis for both the keyword frequency analysis and the semantic network analysis.

Brand-specific hotel names were retained during preprocessing because they formed part of the original customer re-

view content. However, the interpretation of results focused on general customer experience attributes rather than hotel-specific references to minimize potential brand-related bias.

Furthermore, the distribution of hotel experience evaluation based upon overall satisfaction score was performed to generally ascertain the satisfaction level of hotel experience, and this overall score was used as a dependent variable because it was treated as a main output variable. In addition, the word matrix (keywords × keywords) was deduced for fur-

ther data analysis, and this was used to visualize the relationships between keywords in the reviews.

The keyword network visualization was generated using the UCINET 6.0 package with the visualization tool NetDraw. The result of this analysis is presented in Figure 3, which illustrates the relationships among frequently mentioned words in customer reviews and highlights the main keywords associated with the hotel experience.

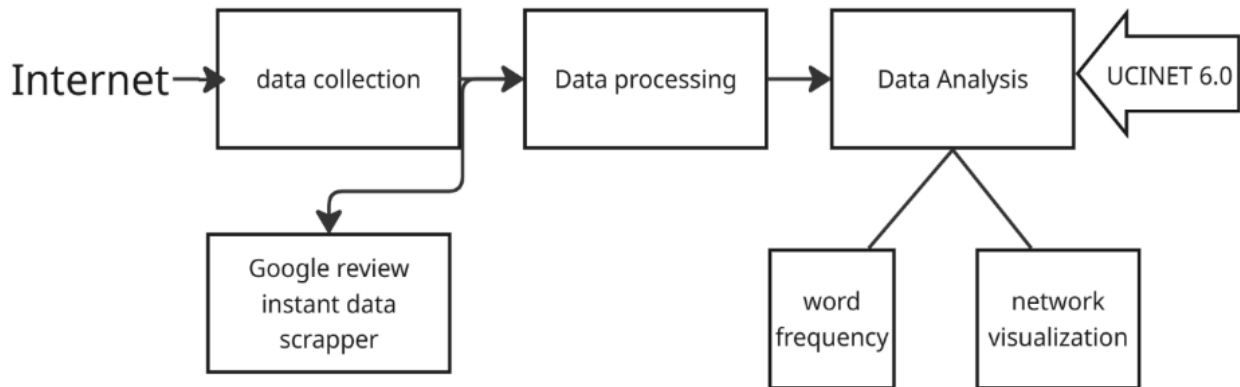


Figure 2. Research process.

The research process shown in figure 2 began with collecting online customer reviews from Google Reviews for four five-star hotels listed by the Sri Lanka Tourism Development Authority. The data were extracted using Instant Data Scraper, followed by filtering and cleaning to remove non-English, duplicate, and incomplete reviews. The cleaned dataset was then processed to extract keywords, and network analysis was conducted using UCINET, with visualization performed through NetDraw to identify key themes in customer experiences.

### 3.3. Theoretical Operationalization

This study applies Expectancy-Disconfirmation Theory (EDT) and Customer Delight Theory (CDT) to online hotel reviews. Instead of using surveys, customer satisfaction is inferred from review content. Keyword frequency analysis identifies key service attributes, such as room quality, cleanliness, location, staff behavior, food quality, and value for money, which reflect whether hotel performance meets or exceeds customer expectations, consistent with EDT.

Co-occurrence network analysis examines relationships between service attributes and emotional expressions. Positive and emotionally charged terms indicating exceptional experiences, excitement, or surprise are interpreted through CDT, as they reflect experiences beyond ordinary satisfaction. Together, these methods capture both the cognitive evaluations emphasized by EDT and the emotional responses highlighted by CDT.

## 4. Results

### 4.1. Frequency Analysis

To find the words most frequently used in customer reviews, Table 2 lists the top 97 frequent words associated with hotel experience in Colombo’s five-star hotels. The top five words were ‘hotel’, ‘room’, ‘service’, ‘good’, and ‘food’. This indicates that guests prioritize overall hotel impression, room quality, service delivery, and dining options in their feedback.

Several words were related to food and beverage experience, such as ‘food’, ‘restaurant’, ‘buffet’, ‘dinner’, ‘breakfast’, ‘delicious’, ‘variety’, ‘dining’, ‘taste’, ‘bar’, ‘lunch’, ‘tea’, and ‘drinks’. These reflect the significant role of culinary offerings in shaping guest satisfaction and the importance of variety and quality in hotel dining experiences. In terms of service, commonly used words included ‘staff’, ‘service’, ‘hospitality’, ‘team’, ‘reception’, ‘manager’, ‘care’, and ‘professional’. These suggest that personalized service and staff behavior remain crucial aspects of the hotel stay. Terms related to the physical environment and amenities were also prevalent, such as ‘room’, ‘pool’, ‘view’, ‘sea’, ‘bathroom’, ‘bed’, ‘floor’, ‘facility’, ‘lobby’, ‘lounge’, ‘club’, ‘swimming’, and ‘rooftop’. These tangible elements significantly contribute to how guests perceive luxury and comfort during their stay. Location-based and brand-specific mentions like ‘Colombo’, ‘city’, ‘area’, ‘beach’, ‘mall’, and ‘Shangrila’ indicate the influence of geographical setting and hotel brand recognition on customer reviews.

Finally, there were many words describing the emotional and experiential aspects of a stay, such as ‘amazing’, ‘excellent’, ‘wonderful’, ‘memorable’, ‘fantastic’, ‘elegant’, ‘pleasant’, ‘lovely’, ‘friendly’, ‘comfortable’, and ‘enjoyed’. These words emphasize how guests emotionally connect with their hotel experience, highlighting satisfaction, joy, and appreciation. The result of visualizing the network that reflects the frequency is Figure 3.

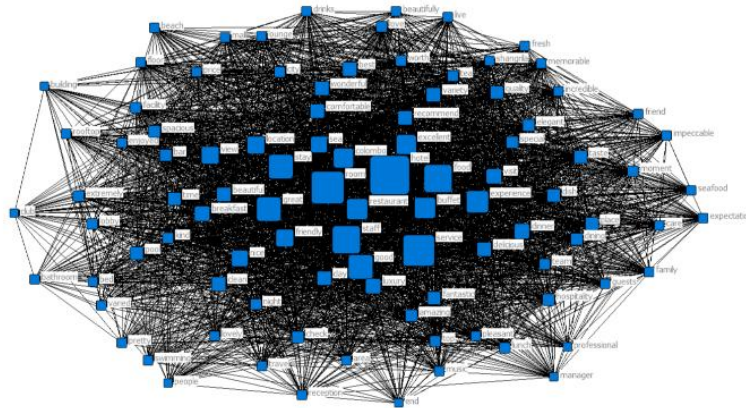
#### 4.2. Semantic Network Analysis

Figure 3 presents the semantic co-occurrence network generated from hotel customer reviews. Node size is propor-

tional to degree centrality, indicating the relative importance of each keyword within the overall semantic structure. The network exhibits a highly dense and centralized configuration, suggesting that customer perceptions are organized around a limited number of strongly interconnected concepts. Several keywords occupy the structural core of the network, including food, good, great, staff, service, restaurant, hotel, and buffet. The prominence of these nodes indicates that dining experiences, service encounters, and overall hospitality quality constitute the dominant themes shaping customer evaluations.

**Table 2.** Top 97 frequent words from the online hotel review

Rank	Word	Freq	Rank	Word	Freq	Rank	Word	Freq
1	hotel	429	34	comfortable	47	67	pretty	19
2	room	279	35	amazing	45	68	worth	19
3	service	262	36	quality	44	69	guests	18
4	good	201	37	dish	43	70	enjoyed	18
5	food	189	38	check	39	71	lovely	18
6	staff	187	39	taste	39	72	fresh	18
7	stay	185	40	bar	38	73	reception	17
8	great	166	41	fantastic	36	74	lounge	17
9	restaurant	139	42	lunch	36	75	rooftop	17
10	buffet	131	43	tea	34	76	care	16
11	experience	131	44	night	33	77	varied	16
12	colombo	119	45	team	32	78	impeccable	15
13	excellent	118	46	dining	29	79	floor	15
14	location	98	47	hospitality	28	80	manager	15
15	nice	95	48	special	27	81	club	15
16	view	90	49	top	27	82	incredible	14
17	friendly	85	50	city	27	83	memorable	14
18	sea	79	51	spacious	27	84	travel	14
19	luxury	70	52	moment	26	85	live	14
20	dinner	68	53	pleasant	23	86	people	13
21	visit	65	54	kind	23	87	beautifully	13
22	beautiful	64	55	lobby	23	88	expectations	13
23	breakfast	61	56	music	22	89	professional	13
24	clean	59	57	price	22	90	bathroom	13
25	wonderful	58	58	shangrila	22	91	friend	13
26	time	55	59	extremely	21	92	seafood	12
27	best	54	60	love	21	93	end	12
28	place	52	61	beach	21	94	drinks	12
29	pool	52	62	family	21	95	swimming	12
30	day	52	63	facility	20	96	mall	12
31	delicious	51	64	bed	20	97	building	12
32	variety	49	65	elegant	19			
33	recommend	48	66	area	19			



**Figure 3.** Keyword network visualization.

The visualization further reveals a core–periphery structure in which highly central concepts are concentrated within the network center, while more specialized attributes appear at the periphery. Keywords such as beach, rooftop, swimming, bathroom, club, and people are positioned toward the outer regions of the network, indicating that these attributes are discussed less frequently in conjunction with the broader experiential themes. This pattern suggests that customers primarily evaluate their hotel experiences through integrated assessments of service quality, staff performance, accommodation experience, and food-related attributes, whereas facility-specific features play a comparatively secondary role.

The variation in node size provides additional evidence of the unequal influence of individual concepts within the network. Larger nodes represent highly connected keywords that frequently co-occur with numerous other concepts, reflecting their central role in customer discourse. In contrast, smaller nodes indicate more specialized or context-dependent topics. Overall, the network structure demonstrates that customer satisfaction is shaped by an interconnected system of experiential factors, with service quality, food experience, accommodation quality, and employee interactions forming the cognitive foundation of customer perceptions.

**Table 3.** Semantic network statistics

Measure	Value
Nodes	97
Links	6,690
Density	3.622
Degree Centralization	0.1447

**Table 4.** Top 15 keywords by degree centrality

Rank	Keyword	Degree	nDegree
1	hotel	1912	0.1732
2	room	1491	0.1351
3	service	1451	0.1314
4	food	1235	0.1119
5	staff	1208	0.1094
6	stay	1037	0.0939
7	good	1007	0.0912
8	great	985	0.0892
9	experience	845	0.0765
10	restaurant	836	0.0757
11	buffet	800	0.0725
12	excellent	756	0.0685
13	colombo	738	0.0668
14	friendly	635	0.0575
15	location	609	0.0552

The semantic network comprised 97 keywords and 6,690 co-occurrence links, indicating a highly interconnected struc-

ture. The density value (3.622) reflects strong associations among review concepts, while the degree centralization

score (0.1447) suggests that customer perceptions are distributed across multiple interconnected themes rather than concentrated around a single concept.

Table 4 presents the top 15 keywords ranked by degree centrality. The findings indicate that *hotel* (Degree = 1912), *room* (1491), and *service* (1451) occupy the most central positions within the semantic network, highlighting their dominant role in customer evaluations. Other highly connected concepts include *food*, *staff*, *stay*, *good*, and *great*, suggesting that accommodation quality, service performance, and dining experiences represent the core dimensions of customer satisfaction. Additionally, keywords such as *restaurant*, *buffet*, *excellent*, *friendly*, and *location* demonstrate substantial connectivity, indicating that both functional service attributes and positive experiential elements contribute significantly to overall customer perceptions. Collectively, these results reveal that customer evaluations of Colombo's five-star hotels

are shaped by a combination of service quality, accommodation experiences, and emotional responses to the hospitality encounter.

Table 5 presents the betweenness centrality results of the semantic network. The findings indicate that *staff* (Betweenness = 36.061, nBetweenness = 0.791) and *room* (36.061, 0.791) function as the most important bridging concepts within the network, followed by *good*, *hotel*, *stay*, and *service*. These keywords occupy strategic intermediary positions that connect multiple thematic areas of customer evaluations. Furthermore, the prominence of *food*, *restaurant*, and *buffet* suggests that dining experiences play a significant role in linking diverse aspects of the hotel experience. Overall, the results demonstrate that accommodation quality, employee interactions, and service delivery serve as the key integrative dimensions shaping customers' overall perceptions of five-star hotels in Colombo.

**Table 5.** Top 15 keywords by betweenness centrality

Rank	Keyword	Betweenness	nBetweenness
1	staff	36.061	0.7908
2	room	36.061	0.7908
3	good	34.228	0.7506
4	hotel	34.218	0.7504
5	stay	33.354	0.7315
6	service	33.249	0.7291
7	food	31.037	0.6806
8	buffet	30.448	0.6677
9	restaurant	30.528	0.6695
10	colombo	28.757	0.6306
11	excellent	27.913	0.6121
12	experience	28.255	0.6196
13	visit	27.782	0.6092
14	beautiful	25.821	0.5662
15	great	25.563	0.5610

## 5. Discussion

The findings reveal that customer satisfaction in Colombo's five-star hotels is primarily shaped by service quality, accommodation quality, and food-related experiences. Across the frequency, centrality, and network analyses, keywords such as *hotel*, *room*, *service*, *food*, and *staff* consistently occupied the most prominent positions. This indicates that customers evaluate luxury hotel experiences through a combination of core service attributes and experiential elements rather than isolated service encounters.

The results are consistent with prior hospitality research that identifies service quality, employee behavior, and accommodation quality as key determinants of customer satisfaction (Torres, 2014; Wirtz & Jerger, 2016). In particular, the high centrality of *staff* and *service* reinforces the argument that interpersonal interactions remain fundamental to guests' evaluations of hotel experiences. Similarly, the prominence of *food*, *restaurant*, and *buffet* supports previous studies highlighting the importance of dining experiences in luxury hospitality settings (Suwanamas et al., 2015). These findings suggest that customer satisfaction is influenced by both

tangible service outcomes and the quality of customer–employee interactions.

The network analysis provides insights beyond those obtained through frequency analysis alone. While frequency analysis identifies the most frequently mentioned attributes, the semantic network reveals the structural relationships among customer experience dimensions. The high degree and betweenness centrality values of *staff*, *room*, *service*, and *food* indicate that these concepts function not only as important evaluation criteria but also as key connecting mechanisms linking multiple aspects of the hotel experience. This finding demonstrates that customer satisfaction emerges from an interconnected system of service encounters, accommodation experiences, and dining activities rather than from individual service attributes considered in isolation.

From a theoretical perspective, the findings provide support for the integrated expectancy-disconfirmation and customer delight framework adopted in this study. Expectancy-disconfirmation theory suggests that customers evaluate service experiences by comparing their expectations with actual performance outcomes. The prominence of service-related, accommodation-related, and food-related keywords indi-

cates that these attributes were frequently associated with customers' evaluations of their hotel experiences. Functional attributes such as service, staff, room, and food appear to represent dimensions through which guests assess whether their expectations have been met or exceeded. At the same time, emotionally oriented keywords such as "excellent," "great," "friendly," "wonderful," and "experience" reflect positive affective responses that align with the concept of customer delight. These findings are consistent with previous hospitality studies that emphasize the importance of both functional service performance and emotional engagement in customer evaluations. The coexistence of functional and emotional concepts within the semantic network suggests that customer satisfaction discussions in online reviews encompass both cognitive assessments and affective responses. Therefore, the findings extend previous research by demonstrating how these dimensions coexist within customer-generated online review content in the context of luxury hotels in Colombo.

The findings are also broadly consistent with previous studies that identified service quality, employee behavior, accommodation quality, and dining experiences as prominent themes in hotel customer evaluations. While earlier research has primarily relied on survey-based approaches to examine customer satisfaction, the present study demonstrates that similar themes emerge from large-scale online review data. This consistency strengthens confidence in the relevance of these attributes across different methodological approaches. At the same time, the study contributes to the literature by applying semantic network analysis to uncover the structural relationships among customer experience dimensions. Unlike traditional approaches that focus primarily on individual determinants of satisfaction, the network perspective highlights how service, accommodation, food, and emotional experiences are interconnected within customer discussions. These findings provide additional insight into the multidimensional nature of hotel customer satisfaction and illustrate the value of combining keyword frequency analysis with semantic network analysis to better understand customer perceptions. However, the results should be interpreted as associative patterns derived from textual data rather than evidence of causal relationships between specific hotel attributes and customer satisfaction.

## 6. Conclusion

This study examined online Google reviews of four five-star hotels in Colombo, Sri Lanka, to identify the main attributes shaping customer experience and satisfaction. The findings indicate that guest evaluations were most frequently associated with three interconnected areas: service quality, accommodation quality, and food and beverage experience.

The frequency analysis revealed that words such as "hotel," "room," "service," "food," and "staff" dominated the reviews, suggesting that these aspects were prominent in customers' descriptions and evaluations of luxury hotel experi-

ences. In addition, positive Evaluative terms such as "excellent," "wonderful," and "amazing" suggest that emotional response is an important part of the overall hotel experience.

The keyword network analysis reinforced these patterns by showing strong relationships among service-related, room-related, and food-related terms, suggesting that customer satisfaction discussions in five-star hotel reviews are associated with multiple interconnected attributes rather than a single factor. This means that hotel managers should pay equal attention to frontline staff performance, room comfort and cleanliness, and the quality and variety of food and beverage offerings.

From a practical perspective, the study highlights the value of online review data as a rich source of real customer feedback. Hotels can use this kind of analysis to identify strengths and weaknesses more accurately, support service improvements, and strengthen their competitiveness in a growing tourism market such as Colombo.

Overall, the study demonstrates that big data analysis of online reviews is a useful approach for understanding hotel customer satisfaction. It also contributes to hospitality research by showing how customer-generated content can reveal key experience attributes that are frequently associated with customer satisfaction in luxury hotels.

## 7. Implications

### 7.1. Theoretical Implications

This study contributes to the hospitality and customer satisfaction literature in several ways. First, it extends existing research by integrating expectancy-disconfirmation theory and customer delight theory within a big data analytics context. The findings demonstrate that customer satisfaction in luxury hotels is shaped not only by the fulfillment of core service expectations, reflected in highly central keywords such as service, staff, room, and food, but also by positive emotional responses captured through keywords such as excellent, great, friendly, and experience. This suggests that both cognitive evaluations and affective reactions jointly influence customer satisfaction.

Second, the study demonstrates the value of semantic network analysis as a complementary approach to traditional hospitality research methods. While previous studies have primarily relied on survey-based techniques, the present study reveals the structural relationships among customer experience attributes by identifying influential and bridging concepts through centrality analysis. Finally, the study provides empirical evidence from an emerging tourism destination, thereby extending the geographical scope of customer satisfaction research within the luxury hotel sector.

### 7.2. Practical Implications

The findings provide several actionable implications for hotel managers. The prominence of service, staff, and room suggests that investments in employee training, service re-

sponsiveness, and room maintenance should be prioritized, as these attributes occupy central positions in customers' evaluations. Managers should establish continuous service quality monitoring systems and regularly assess guest feedback to identify operational weaknesses.

The high centrality of food, restaurant, and buffet indicates that dining experiences are a critical component of overall hotel satisfaction. Therefore, hotel operators should focus on menu variety, food quality, and restaurant service standards to enhance guest experiences. Furthermore, the presence of emotionally oriented keywords such as excellent, great, and friendly suggests that hotels should move beyond basic service delivery and create memorable experiences through personalized interactions and customer engagement initiatives. Finally, the study demonstrates that online review analytics can serve as a strategic decision-support tool, enabling managers to monitor customer perceptions in real time and make evidence-based service improvements.

## 8. Limitation and Future Research

A further limitation of this study relates to the exclusion of non-English reviews. Although this decision improved consistency in text processing and semantic interpretation, it may have reduced the representativeness of the dataset. Sri Lanka attracts tourists from diverse linguistic and cultural backgrounds, and guests who write reviews in languages other than English may emphasize different aspects of hotel experiences and satisfaction. Consequently, the findings may more strongly reflect the perceptions of English-speaking travelers than those of the broader international tourist population. Future research should incorporate multilingual review datasets and advanced natural language processing techniques to capture a wider range of customer perspectives and improve the generalizability of the findings.

Future researchers should incorporate multilingual text mining and advanced sentiment analysis to capture nuanced positive and negative experiences across diverse guest demographics. Comparative studies with hotels in other Asian cities could test the universality of identified satisfaction clusters. Longitudinal analysis tracking review trends over time would reveal how service improvements impact satisfaction evolution. Integrating machine learning models, such as topic modeling, could uncover latent themes beyond word frequency.

**Funding:** This research received no external funding.

**Institutional Review Board Statement:** This study followed ethical standards for human research. Participants were informed about the study purpose, voluntary participation, confidentiality, and their right to withdraw at any time, and all data were used solely for academic purposes.

**Data Availability Statement:** The data supporting the findings of this study are available from the corresponding author upon reasonable request.

**Conflicts of Interest:** The author declares no conflict of interest.

## References

- Abrate, G., Quinton, S., & Pera, R. (2021). The relationship between price paid and hotel review ratings: Expectancy-disconfirmation or placebo effect? *Tourism Management*, 85, 104314. <https://doi.org/10.1016/j.tourman.2021.104314>
- Ali, F., Hussain, K., & Ragavan, N. A. (2014). Memorable customer experience: Examining the effects on customers' behavioral intentions. *Journal of Quality Assurance in Hospitality & Tourism*, 15(3), 293–313. <https://doi.org/10.1080/19368623.2014.945223>
- Berman, B. (2005). How to delight your customers. *California Management Review*, 48(1), 129–151. <https://doi.org/10.2307/41166331>
- Bronner, F., & de Hoog, R. (2011). Vacationers and eWOM: Who posts, and why, where, and what? *Tourism Management*, 32(1), 15–26. <https://doi.org/10.1016/j.tourman.2009.10.016>
- Buhalis, D., & Law, R. (2008). Progress in information technology and tourism management: 20 years on and 10 years after the Internet—The state of eTourism research. *Tourism Management*, 29(4), 609–623. <https://doi.org/10.1016/j.tourman.2008.01.005>
- Carraher-Wolverton, C., & Hirschheim, R. (2023). Utilizing expectation disconfirmation theory to develop a higher-order model of outsourcing success factors. *Journal of Systems and Information Technology*, 25(1), 1–29. <https://doi.org/10.1108/JSIT-07-2021-0143>
- Cheng, B. L., Gan, C. C., Imrie, B. C., & Mansori, S. (2019). Service recovery, customer satisfaction and customer loyalty: Evidence from Malaysia's hotel industry. *International Journal of Quality and Service Sciences*, 11(2), 187–203. <https://doi.org/10.1108/IJQSS-09-2017-0081>
- Ding, K., Bao, Y., Li, L., Zhang, R. R., & Chen, Y. (2025). From crisis to change: Exploring the lasting influence of COVID-19 on Airbnb users through structural topic modeling. *Humanities and Social Sciences Communications*, 12(1), 1–15. <https://doi.org/10.1038/s41599-025-05153-8>
- Drieger, P. (2013). Semantic network analysis as a method for visual text analytics. *Procedia - Social and Behavioral Sciences*, 79, 4–17. <https://doi.org/10.1016/j.sbspro.2013.05.053>
- Fortanier, F., & Van Wijk, J. (2010). Sustainable tourism industry development in sub-Saharan Africa: Consequences of foreign hotels for local employment. *International Business Review*, 19(2), 191–205. <https://doi.org/10.1016/j.ibusrev.2009.11.007>
- Glaveli, N., Manolitzas, P., Palamas, S., Grigoroudis, E., & Zopounidis, C. (2023). Developing effective strategic decision-making in the areas of hotel quality management and customer satisfaction from online ratings. *Current Issues in Tourism*, 26(6), 1003–1021. <https://doi.org/10.1080/13683500.2022.2048805>
- Kwon, H. J., Ban, H. J., Jun, J. K., & Kim, H. S. (2021). Topic modeling and sentiment analysis of online review for airlines. *Information*, 12(2), 78. <https://doi.org/10.3390/info12020078>
- Ladhari, R. (2009). Service quality, emotional satisfaction, and behavioural intentions: A study in the hotel industry. *Managing Service Quality: An International Journal*, 19(3), 308–331. <https://doi.org/10.1108/09604520910955320>

- Lu, W., Huang, S., Yang, J., Bu, Y., Cheng, Q., & Huang, Y. (2021). Detecting research topic trends by author-defined keyword frequency. *Information Processing & Management*, 58(4), 102594. <https://doi.org/10.1016/j.ipm.2021.102594>
- Marcos, A. M. B. F., & Coelho, A. F. M. (2022). Service quality, customer satisfaction and customer value: Holistic determinants of loyalty and word-of-mouth in services. *The TQM Journal*, 34(5), 957–978. <https://doi.org/10.1108/TQM-10-2020-0236>
- Nisar, T. M., Prabhakar, G., Ilavarasan, P. V., & Baabdullah, A. M. (2020). Up the ante: Electronic word of mouth and its effects on firm reputation and performance. *Journal of Retailing and Consumer Services*, 53, 101726. <https://doi.org/10.1016/j.jretconser.2018.12.010>
- Ordenes, F. V., Theodoulidis, B., Burton, J., Gruber, T., & Zaki, M. (2014). Analyzing customer experience feedback using text mining: A linguistics-based approach. *Journal of Service Research*, 17(3), 278–295. <https://doi.org/10.1177/1094670514524625>
- Osman, Z., & Sentosa, I. (2013). Mediating effect of customer satisfaction on service quality and customer loyalty relationship in Malaysian rural tourism. *International journal of economics business and management studies*, 2(1), 25–37.
- Schiebler, T., Lee, N., & Brodbeck, F. C. (2025). Expectancy-disconfirmation and consumer satisfaction: A meta-analysis. *Journal of the Academy of Marketing Science*, 54, 91–112. <https://doi.org/10.1007/s11747-024-01078-x>
- Sparks, B. A., & Browning, V. (2011). The impact of online reviews on hotel booking intentions and perception of trust. *Tourism Management*, 32(6), 1310–1323. <https://doi.org/10.1016/j.tourman.2010.12.011>
- Sri Lanka Tourism Development Authority. (2025). *Official hotel classification list*. <https://www.sltda.gov.lk>
- Sri Lanka Tourism Development Authority. (2025). *Accommodation directory*. [https://www.srilanka.travel/hotels\\_and\\_other\\_accomodation](https://www.srilanka.travel/hotels_and_other_accomodation)
- Sri Lanka Tourism Development Authority. (2025). *Year in review 2025*. Sri Lanka Tourism Development Authority.
- Sukhu, A., Choi, H., Bujisic, M., & Bilgihan, A. (2019). Satisfaction and positive emotions: A comparison of the influence of hotel guests' beliefs and attitudes on their satisfaction and emotions. *International Journal of Hospitality Management*, 77, 51–63. <https://doi.org/10.1016/j.ijhm.2018.06.013>
- Suwanamas, C., Trimetsontorn, J., & Fongsuwan, W. (2015). The effect of perceived service quality on customer emotions, customer satisfaction and customer loyalty: A PLS-SEM approach towards a luxury hotel in Bangkok, Thailand. *Journal for Global Business Advancement*, 8(4), 374–398. <https://doi.org/10.1504/JGBA.2015.074018>
- Tolkach, D., Pratt, S., Masiero, L., Jørgensen, M. T., Zoltan, J., Schuckert, M., & Chon, K. (2024). *Self-interest, ethical and environmental considerations of air traveler weight policies*. *International Journal of Tourism Research*, 26, e2691. <https://doi.org/10.1002/jtr.2691>
- Torres, E. N. (2014). Deconstructing service quality and customer satisfaction: Challenges and directions for future research. *Journal of Hospitality Marketing & Management*, 23(6), 652–677. <https://doi.org/10.1080/19368623.2014.846839>
- Torres, E. N., Fu, X., & Lehto, X. (2014). Examining key drivers of customer delight in a hotel experience: A cross-cultural perspective. *International Journal of Hospitality Management*, 36, 255–262. <https://doi.org/10.1016/j.ijhm.2013.10.007>
- Wirtz, J., & Jerger, C. (2016). Managing service employees: literature review, expert opinions, and research directions. *The Service Industries Journal*, 36(15–16), 757–788. <https://doi.org/10.1080/02642069.2016.1278432>
- World Tourism Organization. (2023). *International tourism highlights*. UNWTO.
- Xiang, Z., Du, Q., Ma, Y., & Fan, W. (2017). A comparative analysis of major online review platforms: Implications for social media analytics in hospitality and tourism. *Tourism Management*, 58, 51–65. <https://doi.org/10.1016/j.tourman.2016.10.001>
- Zeng, B., & Gerritsen, R. (2014). What do we know about social media in tourism? A review. *Tourism Management Perspectives*, 10, 27–36. <https://doi.org/10.1016/j.tmp.2014.01.0011>

Received 2026.04.07 - Revised 2026.06.08 - Accepted 2026.06.19

© 2026 The Author(s). Published by Academic Society of Food and Service Management. This is an open access article distributed under the terms of the Creative Commons Attribution License (CC BY 4.0, <https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.