Tourism Recreation Research

ISSN: (Print) (Online) Journal homepage: https://www.tandfonline.com/loi/rtrr20

Customer expectations in the hotel industry during the COVID-19 pandemic: a global perspective using sentiment analysis

Mihir P. Mehta, Gopal Kumar & M. Ramkumar

https://doi.org/10.1080/02508281.2021.1894692

Customer expectations in the hotel industry during the COVID-19 pandemic: a global perspective using sentiment analysis

Mihir P. Mehta, Gopal Kumar and M. Ramkumar

Operations & Quantitative Methods, Indian Institute of Management Raipur, Raipur, India

ABSTRACT

Hotel industry is the one which has confronted the unprecedented effect of the coronavirus disease 2019 (COVID-19) pandemic to significant social and economic risks. The COVID-19 pandemic has challenged the tourism across the globe and impacted hospitality in hotel industry severely. This study aims to assess customer satisfaction by carrying sentiment analysis and topic modelling over customer reviews on the hospitality provided by hotels in different continents during January to September 2020, i.e. the COVID-19 pandemic. We formulate an improved new scale of metrics to categorize customer satisfaction assessed by sentiment analysis in an elaborate way. Topic modelling was deployed to understand various topics most often discussed by customers. We find that North America and Europe could perform up to customer expectation. In Asia, Sri Lanka did well, Indonesia could maintain its customer satisfaction, while India consistently improved the satisfaction level. We identified 12 most discussed topics, and main reasons of dissatisfaction appear in staff, service, room, cleanliness, slow booking, and pandemic response by hotel. Findings of this study will help senior managers of hotels of developed as well as developing countries in providing new and effective services that can satisfy customers and restore their confidence.

ARTICLE HISTORY
Received 6 November 2020
Accepted 13 February 2021

KEYWORDS
Hospitality; hotel; satisfaction; sentiment analysis; topic modelling; COVID-19

Introduction

Probably, coronavirus disease 2019 (COVID-19) brought the most devastating effect of all the effects in recent human memory. As per the latest record, the COVID-19 pandemic directly affected the health of more than 45 million people and caused more than one million death worldwide. More of its effects are now apparent and reflecting in interrupted cash flow resulting from business interruption and closure followed by lockdown order by governments (Filimonau et al., 2020; Hall et al., 2020). As domestic and international travel was halted (Das & Tiwari, 2020; Yeh, 2020) and trusting service providers for health and hygiene was risky, tourism industry was hardest hit (Polyzos et al., 2020) and, hence, demand for hotel and hospitality service decreased significantly. Decreased demand was fuelled by different and uncompromising service expectations by customers. Social distancing norm imposition forced the hospitality sector to operate at significantly lower capacity in terms of rooms and staffs. Sudden and devastating impacts of the COVID-19 pandemic raised a serious question that if the hospitality sector, especially hotel industry, is prepared to meet customer expectation under uncertainties which may disrupt human life and businesses on the scale COVID-19 is causing.

Tourism industry accounts for 10% of the global GDP (Faus, 2020). Uncertainty in the form of COVID-19 event disrupted tourism and, therefore, the business of hotel industry (Filimonau et al., 2020; Hao et al., 2020). Tourism industry could realize a shock of 60–80% decline in international tourism economy (OECD, 2020), 50 million jobs in travel and tourism industry are at risk, and Asia being the most affected continent as per the World Travel and Tourism Council (Faus, 2020). The probability of similar disaster (such as earthquake, Tsunami, terrorist attack, and extreme climate) in near future that can disrupt human life and business is increasing. McKinsey Global Institute reminds us that the COVID-19 pandemic is not an outlier that disrupted business and its supply chain, but it is only latest in a series of disruptions (Lund et al., 2020). Noticeably, now companies should expect disruptions in their supply chain lasting a month or longer may occur every 3.7 years (Lund et al., 2020). Uncertain events may unfold so rapidly that managers of hotel can't wait for how these events and developments shape

CONTACT Gopal Kumar gopalkumar@iimraipur.ac.in Operations & Quantitative Methods, Indian Institute of Management Raipur, Atal Nagar, P. O. – Kurru (Abhanpur), Raipur 493 661, India

out (Olsen, 1996). So, managers must become boundary spanners who will be efficient in scanning business environment to assess the pattern of change in expectation and adjusting internal operations (Olsen, 1996). Post-event in general and post-lockdown in particular, the primary concern must be to retain the customer base by satisfying their new expectations, new lifestyle, displaying a sense of sensitivity, and restoring their confidence. Therefore, it is very important for hotel industry to have strategies that can effectively deal with sudden disruptions and managers should be capable of sensing the trend of customer expectations and their new lifestyle (Das & Tiwari, 2020; Yeh, 2020), such that they can provide them with satisfying experience during and after the uncertain event.

Drawing on the above premises, this study is formulated to study the impact of COVID-19 on hotel industry in handling customer satisfaction and expectation. Especially, we analysed customer reviews during January–September 2020 if their service experience is satisfactory during pre-lockdown, lockdown, and post-lockdown. As social media and public forums are becoming popular for research in tourism and hospitality (Rasul & Hoque, 2020), we collected customer reviews from TripAdvisor.com in the form of big data and employed sentiment analysis and topic modelling. The literature (Alamoodi et al., 2020; Borg & Boldt, 2020) recommends sentiment analysis to improve hotel performance and to handle uncertainties. Using polarity score, we developed a new scale of metrics for unsupervised sentiment analysis. Customer reviews of hotels from North America, Europe, Africa, and Asia are analysed and compared. As previous studies gave little attention to hotel industry or tourism in developing countries, we preferred to go deeper in analysing customer reviews from India, Indonesia, Malaysia, and Sri Lanka.

This study, therefore, contributes to practical implications for the hospitality sector by recommending the usage of customer review big data and gain on the major trend (changing expectations and demands of the customers) in different continents with different cultures. This study motivates the hotel industry of developed and developing continents to get prepared to handle uncertainties and shows a way to embed an improved resilience capability. Furthermore, we provide managers and decision makers with information about the performance of domestic and international hospitality business competitors. Customers of hotels can also get benefitted from this study, as it provides with information that would help them make better decisions during the efficient selection of hotels.

The structure of the remainder of the paper is as follows. The next section briefly discusses literature and highlights literature gaps, followed by description of the theory of boundary spanning which motivates us to conduct this study, methodology, results, discussion, and the conclusions and limitations.

Literature review

Customer satisfaction along with dissatisfaction has been an important topic in the hospitality sector. In this sector, service quality or satisfaction is critical (Rauch et al., 2015), and it is the gap between perceived and expected service quality (Padma & Ahn, 2020). Customers gain good experience (Godovykh & Tasci, 2020), and they are satisfied when their perceived service quality exceeds the expected service quality. Traditionally, service quality was assessed by different attributes, such as location, room service, cleanliness, comfort, attitude of staffs, booking process, and complaint handling. Through conjoint analysis, Rhee and Yang (2015) found that the importance of hotel attributes varies based on the classification of hotels. Based on luxury hotel experience in Turkey, Cetin and Walls (2016) found that design, amenities, staff professionalism, attentiveness, and attitude towards guest are pivotal attributes influencing customer experience. Customers from different cultures have different services and amenity expectations (Torres et al., 2014), and therefore their overall positive and negative sentiment will vary with the same type and level of services.

In addition to traditional service expectation, COVID-19-like pandemic gives rise to new expectation and service delivery, such as social distancing, sanitization, usage of mask (Das & Tiwari, 2020) and providing it to guests, managing and scheduling common area differently, etc. During COVID-19, tourism and travel intention are affected by the perceived severity of COVID-19, desire, and willingness to adapt personal non-pharmaceutical intervention (Das & Tiwari, 2020). Customer experience is positively correlated with the overall revenue and competitiveness of a hotel. In COVID-19-like situations, it is necessary to maximize customers' footfall by serving customers in ways that can boost their confidence (Hao et al., 2020). A pleasant, healthy, and mentally secure experience can boost customers' confidence, and consequently it can enhance positive sentiments.

Table 1 summarizes relevant literature. Researchers suggested different methods to respond to disasters or crisis. The previous literature used communication, dissemination of information (Jia et al., 2012; Yeh, 2020), collaboration (Nguyen et al., 2017), and influence of belief and psychological factors (Wang & Wu, 2018) to suggest approaches to handle disasters, especially in the hospitality sector. Using neural network, Polyzos

Table 1. Literature summary of methods used.

Studies	Main theme	Methods	Geography	Sector	Major findings
Alamoodi et al. (2020)	Sentiment analysis in COVID-19-like time	Literature review	Not specific		Sentiment analysis can help reduce response time and risks.
Borg and Boldt (2020)	Predicting customer response sentiment	VADER sentiment and support vector machine models, analysis of emails	Swedish	Telecom	Considering even unread email, action plan can be prepared to serve customers better.
Sparks et al. (2016)	Impact of hotel response on customer inference of trust and concern	Empirical, confirmatory factor analysis	Australia	Hotel	Hotel should respond online to negative customer reviews. Agency that responds, voice used, and its timing are impactful.
Hao et al. (2020)	Developing framework to manage and understand disaster management strategies.	Conceptual	China	Hotel	The framework suggests anti-pandemic phases, principles, and strategies. Hotels need to restore costumer's confidence by providing disaster specific services.
Filimonau et al. (2020)	Influence of organizational resilience, response to COVID-19, and perceived job security on organizational commitment	Quantitative, structural equation modeling with partial least squares (PLS) method	Spain	Hotel	Response to COVID-19 is beneficial for organizational resilience, improved job security, and organizational commitment of senior managers.
Padma and Ahn (2020)	Guest satisfaction and dissatisfaction	Empirical, big data analysis in the form of online review	Malaysia	Hotel	Expectation increases with increase in room price. Expectation of guests of luxury hotels is different from those of budget hotels. Guests of luxury hotels are more interested in delight attributes while those of budget hotels are interested in basic attributes.
Tsai et al. (2020)	Disaster prevention management	Interviews, Delphi method	Taiwan	Hotel	Key dimensions to manage disaster are disaster prevention knowledge, attitude, skills, and services. Under hotel disaster prevention, training to hotel employees will impact COVID-19 scenario. management positively.
Melián-Alzola et al. (2020)	Assessing organizational resilience under uncertainties	Quantitative (questionnaire), PLS	Spain	Hotel	Strategy and change management (such as continuous improvement, innovation, etc.) are important for hotel resilience.
Torres et al. (2014)	Assessing drivers of customers delight from cross cultural perspective	Qualitative (interview) and quantitative	United States, South America, Northern Europe, Canada	Hotel	While there are some common expectations, guests from different cultures are delighted by different services and amenities.

et al. (2020) predict that recovery time from arrival to pre-crisis levels can take 6–12 months. The recent literature (Alamoodi et al., 2020; Borg & Boldt, 2020) emphasized that sentiment analysis is useful in the situation of pandemic-like COVID-19 and epidemic, and it can be used to prepare actions for customers. The sentiment analysis on social media should be used to understand customers more deeply which can help different businesses, governments, and non-government organizations to handle COVID-19-like tough times (Alamoodi et al., 2020). Responding to online negative reviews is required, and in this process, voice used to respond and timing of response are impactful to hotel's image and customer's satisfaction (Sparks et al., 2016). Considering the COVID-19 situation, Hao et al. (2020) proposed a framework for anti-pandemic journey. In this direction, hotel industry requires to devise strategies for leadership and communication, human resource, corporate social responsibility (CSR), finance, and standard operating procedure to handle COVID-19-like disaster (Hao et al., 2020). In uncertainties, researchers (Filimonau et al., 2020; Melián-Alzola et al., 2020) attempted to improve hotel resilience. Hotel resilience is how rapidly the hotel can return to its initial state after getting affected by adverse situations. In COVID-19-like situation, the commitment of senior hotel managers to the organization may become low. However, CSR and effective response to COVID-19 can improve manager's commitment and hotel resilience (Filimonau et al., 2020). COVID-19 imposes an adverse situation for hotels, and the hotels require better resilience (Melián-Alzola et al., 2020). Strategy and change dimensions exert a significant effect on hotel resilience (Melián-Alzola et al., 2020).

Researchers should examine if different types of scenarios and service failure require different responses by hotel (Sparks et al., 2016). More research is warranted to understand online customer reviews and sentiments on hospitality in order to respond to customers' demand effectively (Sparks et al., 2016). Previous studies have attempted to improve the perception of the hospitality sector in disaster and crisis, but mostly these disasters imposed local impact (Filimonau et al., 2020). In the light of global disaster and business disruption, such as COVID-19, improving the perception of hotels and customer sentiment requires a customized response and strategies to retain and restore customer's confidence. Previous studies mostly focused on the (immediate) recovery plan, but the unexpected emergence of COVID-19 that devastated global business (Filimonau et al., 2020) and McKinsey Global Institute estimating events that may occur more frequently and lasting for an extended period of time (Lund et al., 2020) pressurizes hotel industry to plan for the long-term disaster preparedness. Research on the link between service quality and customer satisfaction in luxury hotels in developing countries is limited (Padma & Ahn, 2020). When tourism in many developing countries, such as India, Indonesia, Malaysia, and Sri Lanka, is catching up, studies on sentiment analysis and strategy formulation, especially pertaining to pandemic situation, are limited. Studies employing data mining, machine learning, content analysis, and ethnography on big data collected from social media and popular online forums are further limited.

Theory of boundary spanning

It is becoming increasingly important for manager's intellectual capacity to identify what it takes to succeed in the increasingly complex and volatile world of hospitality (Olsen, 1996). The most challenging part is to understand and handle the speed of change. Managers must become boundary spanners who will be capable of scanning business or external environment to identify the pattern of changes that is developing and creating threats and opportunities for their business (Olsen, 1996). The theory of boundary spanning describes the roles that managers can play to tightly link their internal operations to external changing business environment by scanning business environment (Cheng et al., 2020; Tushman & Scanlan, 1981) in order to incorporate critical trends into daily operating decisions. Boundary spanners are crucial who are responsible to transform external meaningful information across organizational boundary (Cheng et al., 2020). Today, when online communities are becoming more important and vocal, boundary spanners' roles in hotel industry in participating in online discussion, extracting online public information, and transforming them into daily operational decision become critical to satisfy changing needs of customers. The theory of boundary spanning is used in tourism and hospitality literature (Cheng et al., 2020; Cooper, 2015), but to the best of our knowledge, this theory in handling uncertainties and preparing future courses by utilizing social media and public forum is rarely used.

Methodology

This section elaborates on the methodology of the text analysis. Given the exponential usage of TripAdvisor by national and international travellers, researchers (Padma & Ahn, 2020; Sparks et al., 2016) use and recommend TripAdvisor as an important source of customers' sentiments and reviews. Hence, we extracted customer reviews from TripAdvisor using a web portal, and these reviews were used for analysis using topic modelling and sentiment analysis. Both methods require review data to be preprocessed first. Steps required for both methods used in this work have been presented in Figure 1.

Data collection and preprocessing

To analyse a wide range of customer demands and expectations, we have scraped travellers' reviews from TripAdvisor. The dataset includes 7568 reviews along with their headline, review date, and stay date. The data came from different continents, namely North America, Europe, Africa, Australia, and Asia that include popular travel countries: USA, UK, India, Indonesia, Malaysia, Singapore, and Sri Lanka. The review dates considered from 1 January 2020 to 31 September 2020 which includes the pre-lockdown, lockdown, and unlock period. To offer an unbiased analysis, we selected ten four- and five-star-rated hotels and resorts from each continent. We separately extracted reviews from popular tourism destination countries around India, namely

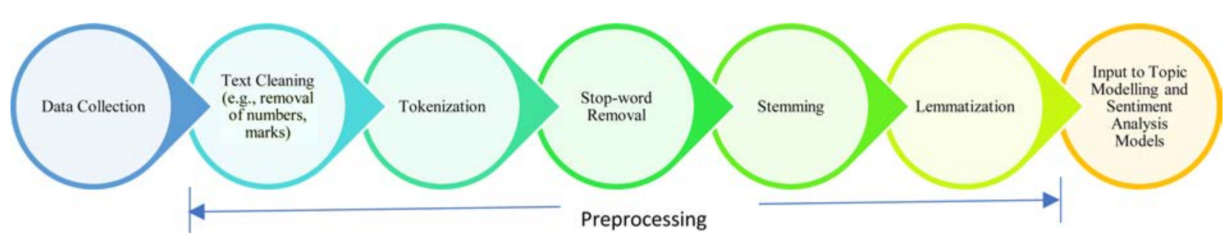

Figure 1. Text preprocessing workflow.

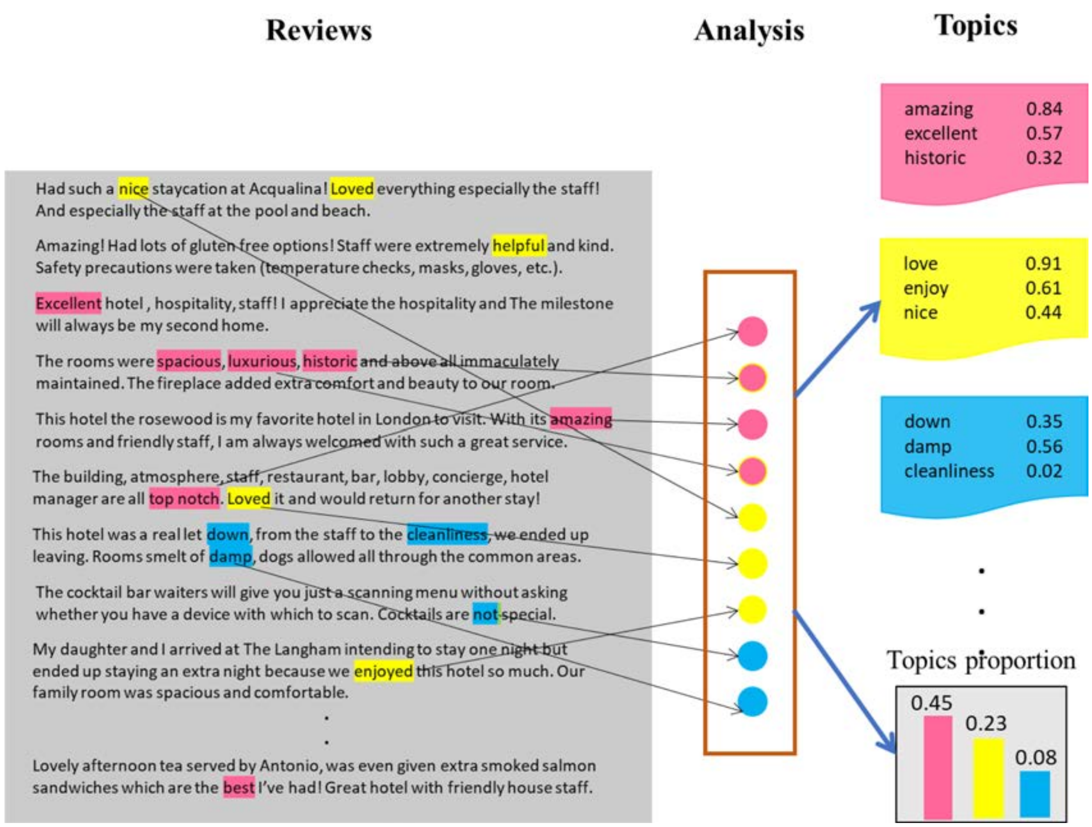

Figure 2. Topic modelling using Latent Dirichlet Allocation (LDA).

Indonesia, Malaysia, Singapore, and Sri Lanka to understand local competition and to contrast their hospitality performance. We conducted continent-, country-, and hotel-based analysis, covering 63 hotels and resorts in total.

To reduce noise in textual data, raw text needs to be preprocessed before feeding it to topic modelling or sentiment analysis algorithm. As shown in Figure 1, the preprocessing includes five steps (Bastani et al., 2019): (1) cleaning (converting into lowercase, stripping html

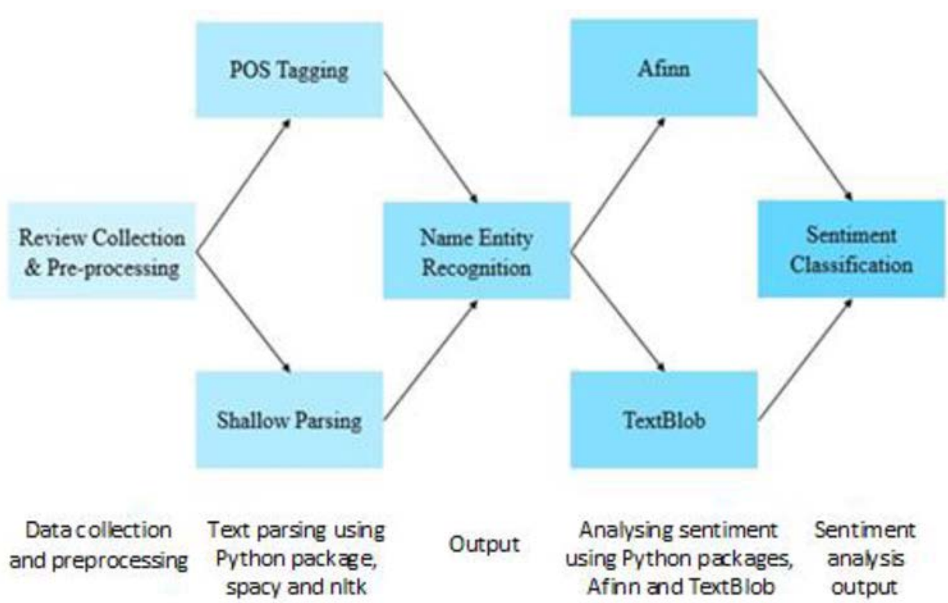

Figure 3. Unsupervised sentiment analysis steps.

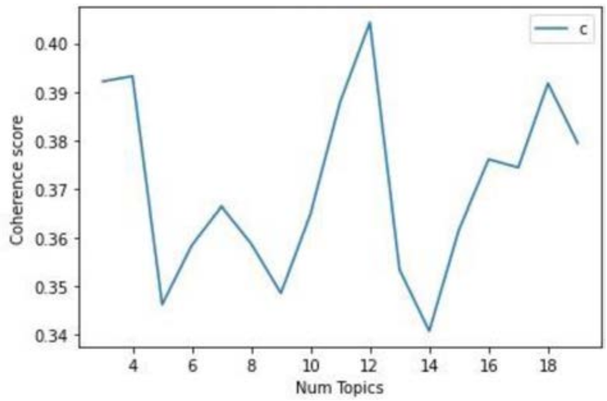

Figure 4. Number of topics versus coherence score (C) plot.

tags, expanding contracted words, removing special characters (e.g. !%$#&*?./.) and digits, and removing newlines), (2) tokenization, (3) stop-words removal, (4) stemming, and (5) lemmatization. Tokenization in Step 2 involves tokenizing contents into terms. To keep the focus on important words, stop-words, such as 'the', 'is', 'a', 'an', and 'be', were removed. Stemming and lemmatization reduce word variation by converting inflected words to their common base/form.

Topic modelling

Topic modelling is a machine learning technique that determines topics from a set of documents to represent them as a cluster of words that cover most of the document and represent the narrative of documents. Each individual group of words represents a topic. We used the Latent Dirichlet Allocation (LDA) technique (Bastani et al., 2019; Blei et al., 2003) that classifies input text into a meaningful number of topics to get insights about the words frequently used by people all over the world while writing reviews. LDA is a generative probabilistic modelling approach that highlights latent semantic structures in a collection of text documents. LDA utilizes Bayesian learning to extract unobservable constructs by estimating their posterior distribution from the joint distribution (Bastani et al., 2019). It works as a two-level allocation. First, it allocates suitable topics to the document, and secondly, it allocates probability distribution over words to each topic.

Figures 2 and 3 illustrate exploratory outputs of LDA. Specifically, LDA gives two outputs, namely topics and their importance or weights in each document. Figure 2 shows that LDA analysis is producing three topics with the largest weights. Each topic is represented by a combination of the top three keywords with the highest probability of occurrence that distinguish various topics from each other. To have a more precise understanding, these topics need to be labelled.

Following the previous literature (Hu et al., 2019; Kuhn, 2018), a coherence score was assessed to find out the number of meaningful topics which is based on the semantic coherence and interpretability (or human judgement) of topics (Schmiedel et al., 2019). The coherence score can be anything between 0 and 1. The number of topics can be any integers which are more than one. A higher coherence score implies that most frequently words in a topic may often appear together. For visualizing the reviews according to dominant topics, we used t-distributed Stochastic Neighbor Embedding (t-SNE) for dimensionality reduction.

Sentiment analysis

Unsupervised sentiment analysis is a machine learning technique utilizing natural language processing to classify the sentiments or opinions of people through their reviews, tweets, comments, and the likes by analysing sequences of words and word nearest neighbours to classify text into positive, negative, or neutral sentiment. Sentiment analysis has multiple applications in business and management to exploit users' reviews and opinions and get insights about their requirements and feedback. Sentiments of people have also been observed to alter stock prices of brands, thereby making it a useful apparatus of study to predict major and minor business fluctuations occurring as a result of social media sentiments.

There are two main methods for sentiment analysis: supervised machine learning-based and sentiment lexicon-based. Supervised machine learning-based methods use text features and sentiment classes. Here, large training dataset and features drive the performance, and bag-of-words is the most common feature type (Liu et al., 2019). We extract the frequency of occurrence of words using bag-of-words approach (Belford & Greene, 2020) by creating a dictionary to calculate how many times a word appears in a document. Bag-of-words disregards the order of words in which they appear and focuses on what words appear in a document. A sentiment lexicon is a collection of words or phrases commonly used to express feeling. While classifying a sentence, the simplest way is to check each word against the sentiment lexicon and count the number of positive and negative words as a sentiment score. We adapted a sentiment lexicon-based or unsupervised approach in which the text is unclassified, unlikely from supervised sentiment analysis where text requires to be classified or labeled for training. The unsupervised approach doesn't require training on specific data, rather provides sentiment scores directly on the basis of length, sequences, and frequency of words. For more discussion on sentiment analysis, interested readers can refer to the literature (Alamoodi et al., 2020; Ma et al., 2018).

Having the data preprocessed, we have used two text parsing approaches, namely POS Tagging and Shallow Parsing (also known as chunking), as shown in Figure 4. POS tagging tags each word to a corresponding part of speech depending on the context of the words used in the text. We performed shallow parsing to label words with their part of speech and present them as a group of grammatically linked words or phrases.

New scale of metrics

Tools to generate a sentiment score usually do not provide flexibility to analyse the sentiments in depth. It only provides an aggregate index based on the numerical output of the overall mix of positive and negative words used in the review. One could not predict the amount of negativity in the low positive scored reviews and the positivity in low negative scored reviews. It is of due importance to analyse if the experience was pleasurable or not, which is not derived from the polarized output from sentiment analyser used in this study. Therefore, we started analysing the collected reviews manually. The fact that the reviews were already segregated with respect to their polarities made the task simpler. Table 2 represents the examples of such ambiguously scored reviews for which the new scale of metrics developed.

Therefore, upon careful analysis of all the reviews, we came up with a scale along with a score range that affixes them into five satisfaction categories, namely Extremely Dissatisfied (ED), Dissatisfied (D), Neutral (N), Satisfied (S), and Extremely Satisfied (ES), as shown in Table 3. The categorization with different score ranges clearly shows why mere polarity would be an inefficient measure to analyse the reviews accurately. Negative reviews of 214 are now converted into a true picture of 1368 below neutral (D and ED) reviews. Even 12 neutrally polarized reviews now expand into 1771 neutrally categorized reviews and thereby capturing the true essence of the sentiments of customer.

Table 2. Examples of ambiguously scored reviews.

Type of review	Review heading	Review content[a]
Negatively scored with positive feedback (−1 score)	Birthday treat	This was our second visit for a meal in the restaurant, the first being a little spoiled by an overly talkative waiter. This time, no such problems and staff attentive, and food and wine was very enjoyable. Small delay to be seated from allotted booking, but a drink from the bar filled the void. The location of the dining room makes the most of its harbourside location, and heating lamps are provided to enjoy tables al fresco, but choose your night as it cools pretty quickly.
Positively scored with negative feedback (5 score)	A little disappointing	This is our fourth visit to this property as it is one of our favorites. I am not sure if it's the time of year but we decided to try February to see the whales which we haven't done before. I can't help but feel like the resort is understaffed, particularly by the pool. One of the things that separated the four seasons from other resorts, particularly the other 5-star resorts in their category that we tend to stay at, is they just have a different level of service. It has been two days now and I'm having to go find my own water, ask where the food we ordered for our kids is (after 50 min, it arrived), look around when I want to order a cocktail or beer, and often have to stand up and flag someone down. It's annoying and makes me just want to get up and go to the bar myself to get a beer. That's not the kind of service I expect (or pay for) when choosing to stay here.

aTypo's or any other grammatical or punctuation error has not been filtered from reviews. It has been presented as it was provided by customers.

Results

The number of reviews that surfaced on TripAdvisor each month for Africa, Australia, Europe, North America, and Asia is counted and presented in Figures 5 and 6. The increase or decrease in the number of reviews indirectly indicates an increase or decrease in footfall of the hotel, respectively. Data preprocessing and LDA-based topic modelling were performed in Python. The NTLK Python library (Bird et al., 2009; NLTK, n.d.) was used to preprocessed the review

Table 3. Categories in new scale of metrics.

Category	Score range	Number of reviews
Extremely dissatisfied	Less than −10 ($x < -10$)	46
Dissatisfied	From −10 to less than 10 ($-10 \leq x < 10$)	1322
Neutral	From 10 to less than 15 ($10 \leq x < 15$)	1771
Satisfied	From 15 to less than 50 ($15 \leq x < 50$)	4614
Extremely satisfied	50 and above (10 ($x \geq 50$)	184

Table 4. Family of words.

ID	Topic keywords	Topic weight	Label
0	Say, call, meet, point, tell, waiter, bad, receive, reservation, review	0.2413	Communication
1	Pleasant, problem, entrance, park, situation, access, wave, pandemic, average, strong	0.2032	Pandemic response
2	Stay, hotel, staff, good, food, service, great, make, thank, excellent	0.6983	Stay experience
3	Villa, holiday, play, spread, flower, photo, picture, driver, privacy, calm	0.0193	Recreation
4	Raffle, colonial, modern, style, course, building, traditional, Singapore sling, handle	0.1256	Aesthetic experience
5	Awesome, booking, fast, country, decorate, crowd, rush, employee, box, imagine	0.2205	Booking experience
6	Drink, use, price, include, water, private, arrival, pay, cold, free, hot, wine	0.3244	Beverage on arrival
7	Polite, ambience, vacation, appear, excursion, massive, baby, line, fill, ice	0.2354	Aura
8	Delicious, activity, attend, dine, club, possible, sweet, legoland, son, ambience	0.1065	Event experience
9	Room, restaurant, pool, hotel, breakfast, view, night, clean, well, bar	0.4371	Hotel infrastructure
10	Day, check, get, take, go, tea, even, time, book, guest	0.4821	Routine
11	Grand, host, Indian, group, watch, nature, environment, shout, control, swim	0.0579	Sightseeing

report, and its outcome was imported to the Gensim Python library (https://radimrehurek.com/gensim/) (Rehurek & Sojka, 2010) for LDA analysis.

The spaCy and NTLK Python libraries were used for the parsing. To detect polarity and subjectivity, we used Python packages, Afinn and TextBlob, as sentiment analyser tools (Al-Natour & Turetken, 2020; Hasan et al., 2018; Loria et al., n.d.). The output thus obtained is in the form of polarity scores between specified ranges, as the algorithm compares them and returns values with a scale of positive, negative, and neutral polarity index.

Topic modelling results

Figure 3 shows the variation in a coherence score as the number of topics increases. By looking at a higher coherence score along with interpretability of topics, 12 topics were kept. We also identified dominant topics and their keywords among the 12 topics and calculated topic percentage contribution for each. To make these 12 topics clearer and more understandable, topics were labelled. Two experts independently looked at keywords of topics and labelled all the topics. They, then, discussed with each other, and finally topics were labelled with a consensus. These topics are communication, entrance experience during pandemic, stay experience, recreation, aesthetic experience, booking experience, beverage on arrival, aura, event experience, hotel infrastructure, routine, and sightseeing. As described in Topic modelling, 12 topics with their weight and corresponding most important ten words are presented in Table 4. Topic weight signifies the cumulative occurrence of these words in the reviews. The results are explained in Discussion section.

Sentiment analysis results

Based on the new metrics scale shown in Table 3, we calculated the average polarity score of reviews that fall under each of those five categories of new scale of metrics. We calculated the polarity score on monthly basis for each of the five continents and for

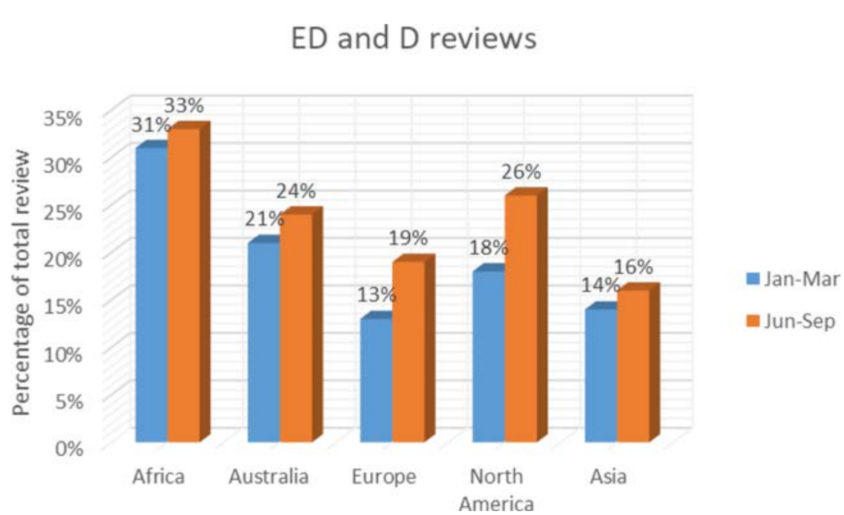

Figure 5. Percentage of ED (Extremely Dissatisfied) and D (Dissatisfied) reviews with respect to total reviews in that period.

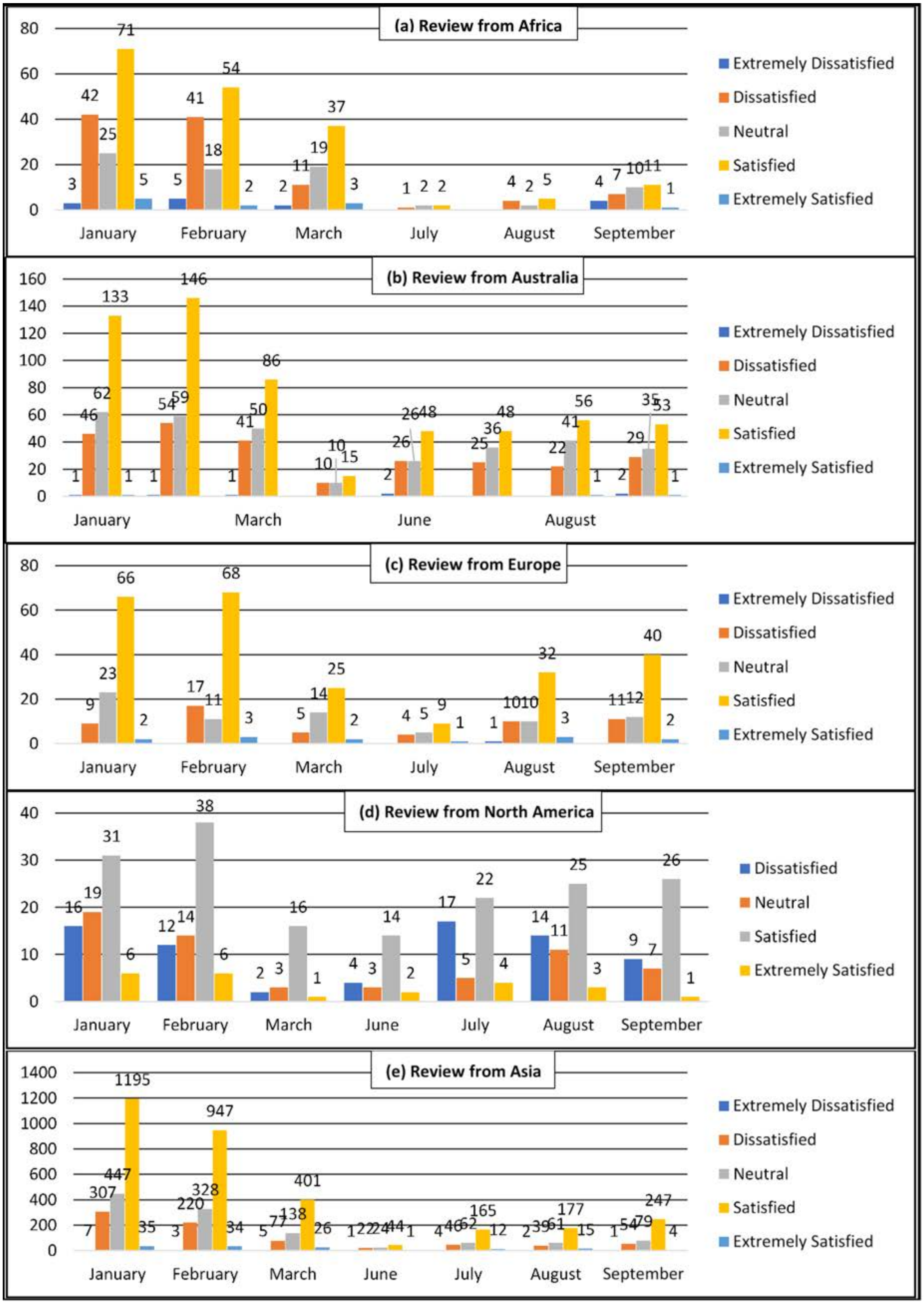

Figure 6. Number of reviews from Africa, Australia, Europe, North America, and Asia.

Figure 7. Category-wise average polarity score of Africa, Australia, Europe, North America, Asia (India, Indonesia, Malaysia, Singapore, and Sri Lanka).

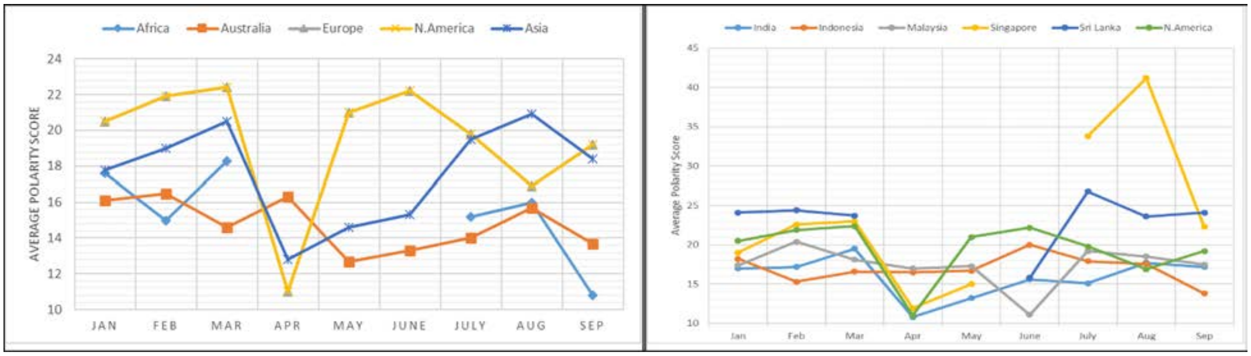

Figure 8. Average polarity score of continents and Asian countries.

Table 5. List of drivers and frequency of their occurrence.

Topic label (No.)	Word	% of ED and D Reviews (1315) containing the word	% of all (7568) Reviews containing the word
Stay experience (2)	Staff	43.3	60.9
	Service	41.3	49.3
	Experience	19.2	22.2
Routine (10)	Time	26.9	26.9
	Book	11.7	8.3
Hotel infrastructure (9)	Clean	13	18.4
	Room	49.8	51.3
	Restaurant	13.8	18.6
Event experience (8)	Delicious	1.3	7.5
Beverage on arrival (6)	Water	7.4	6.2
Booking experience (5)	Fast	15.6	19.8

Note: ED = Extremely Dissatisfied; D = Dissatisfied.

the five countries in the Asian continent. As shown in Figure 7, we plotted the average scores of all those distinct satisfaction categories over months for Africa, Australia, Europe, North America, Asia, India, Indonesia, Malaysia, Singapore, and Sri Lanka to observe the trend in the variation of the polarity score of the hotel. To get better relative insights, Figure 8 presents relative variations in the average polarity score for continents and Asian countries. These figures are explained in Discussion section.

Discussion

Drawing on the theory of boundary spanning, the objective of this study is to understand the voice of the customers of hotel industry and to capture their prevailing sentiment experienced during the COVID-19 pandemic, such that hotel industry can satisfy customers in a difficult time. This study further aims to provide a numerical indication of the levels of quality service, availability of premium facilities, and aesthetic environment facilitated to the customer upon the uplift of the lockdown barrier by conducting the sentiment analysis of the reviews on TripAdvisor. Table A1 shows the average polarity scores in all the five categories of satisfaction level of all the 63 hotels included in the study. Average polarity scores of different continents and different Asian countries are shown in Table A2 and Table A3, respectively.

Polarity score analysis

As shown in Figure 6, of the 387 reviews obtained from the African continent, 120 of them contributed to D and ED categories. A whopping 196 of them were below satisfactory levels, i.e. N, D, and ED. Almost 33% of the reviews constituted to D and ED categories, and almost 66% belonged to the combination of N, D, and ED categories, while the proportion of D and ED category reviews increased from 30.7% in the January–March period to 32.6% in the July–September period.

During the period January–February, the trend of scores of ES, N, and ED was on increase (refer Figure 7), with a significant rise, despite a decline in footfall (refer Figure 6). From January to September, the continent observed a decline in the average polarity score. Despite some fluctuations, overall a decline is observed for all categories of satisfaction level as shown in Figure 7. From February to September, the score of both ES and ED deteriorated drastically. It exhibits the degradation of services and customer experience at top-notch hotels in Africa. This means that hotels were not well equipped and prepared to address a marginally smaller crowd than the regular one with better service. The servicing became poorer with every passing month post the month of July as the guests started arriving. The lowest polarity score of −41 in the month of September was observed for the African continent, stating the diminishing levels of services.

A similar trend of decay of scores is observed in Australia as shown in Figure 7. The drop in average polarity scores of September from January score for all categories of satisfaction (Except N and ES categories) shows that Australia, despite being a popular tourist spot of the world, could not keep up to the expectations of customers, even at its top tier hotels. It also conveys a lack of preparedness and resilience in hotel industry. Despite a nearly 50% drop in footfall (refer Figure 6) in Australia, hotels with relatively more resources failed to provide expected service to the smaller population and improve the customer experience and overall rating of the hotels. As shown in Table 5, the proportion of D and ED category reviews increased from 21.1% in the January–March period to 23.5% in the June–September period. Approximately 25% of customer ratings belonged to D and ED categories and over 52% belonged to D, ED, and N categories. Furthermore, average sentiment scores of Australia in the months of July and August were 14.0 and 15.7, respectively, which are lower than the average sentiment score of Africa, 15.2 and 16.0, in the same period. Also, among continents and countries, the overall average polarity scores are lowest for Australia (except in April). One can although attribute this statistic post the May month to a quick surge and suppression in the second wave of the COVID-19 pandemic (BudgetDirect, 2020), but the scores have been the same during the months of January, February, and March when the country was affected by bushfires. It seems that the top management of hotels in the continent is quite slow in learning and transforming themselves as per changing expectations of customers during COVID-19 or other situations. The present suppression of the pandemic now paves path to managers to recognize the falling satisfactory experience levels and formulate strategies that would help the hotel with better hospitality to secure positive sentiments in the minds and heart of customers. It becomes more important as domestic population contributes 74% to its tourism and tourism GDP has grown at a faster rate than the national economy and thereby contributing 3.1% to it (WHO, 2020).

For Europe, we observe that there has been a rapid increase in people visiting the hotels (perceived based on increasing reviews) from July through September. This can be attributed to gradually rising domestic and outbound tourism in Europe. This rise can again be attributed to a temporary suppression of COVID-19 cases along with an increase in the recovery rate in countries like Germany and Italy, during the months of June and July, which facilitated the movement of people in the continent in subsequent months. This rise can also be due to better-perceived service at hotels.

The miniscule dip in the satisfaction levels can be observed via a drop in average polarity scores in each category in the months of July and August (except D) and an increase in the proportion of ED and D category reviews from 12.6% in the January–March period to 18.6% in the July–September period. However, September saw an increase in satisfaction levels restoring the experience back to normal or at least the way it had been in January. Approximately 15% or less reviews form a part of ED and D categories in September and nearly 34% contributes to ED, D, and N categories together. The majority of travellers (i.e. 66%) in Europe obtained satisfactory (S) to extreme satisfactory (ES) experience over months. This can be a learning model for other countries on how Europe has not only been the hub for tourism but also been a provider of good hospitality to its visitors and travellers irrespective of diversity in the population of visitors coming from different cultures, demography, and in large volume even when the circumstances are unfavourable and unforeseen. It conveys that hotels in Europe provided better and close to the expected service demanded by guests from different cultures. Our findings complement previous studies (Filimonau et al., 2020; Melián-Alzola et al., 2020) that a resilient organization can better cope with COVID-19-like uncertainties. However, the continent, especially the UK, is now witnessing another serious new coronavirus strain. From our findings, it can be said that Europe may be able to maintain customer satisfaction and their positive sentiment by providing expected and desired service to customers. At the same time, researchers across the world would have their eyes and pens active to observe how the continent rises from new coronavirus strain and the hotels keep up with their track of providing hospitality to their customers.

North America recorded the highest polarity scores of positive sentiments. The sentiment is dominated by S and ES categories. The polarity score of ES is highest among continents and countries. Notably, there is no review that falls in the bracket of ED category. During the main period of the COVID-19 pandemic, the percentage share of D category reviews has increased from nearly 20% in the initial two months to nearly 30% in July. The score of D improved from June to July and again from August to September, and it seems gradually reaching the initial score of January. The satisfaction level of all categories improved from August to September. Among continents, as shown in Figure 8, overall North America maintained its domination in satisfaction, except a significant drop in satisfaction in April, perhaps during the peak of the pandemic fear. In August, a further drop in satisfaction was observed and, in the same month, Asia surpassed North America in satisfaction. However, again in September North America's domination was restored. The analysis of hospitality levels of North America is hindered by merely 72 reviews available (in January) which is less than any of the continents and countries in our analysis.

Finally, the continent, Asia has the highest volume of data with five countries considered for analysis. We observe that the proportion of reviews in ED and D categories is 14.48% in the first three months and 15.9% in June–September, with nearly 60% of its customer population having S and ES experience. ES, S, and average polarity scores deteriorated from March to April, and then they improved through July. From August to September, scores of ES, S, N, and average polarity again deteriorated. The ES score of Asia is the second-best figure observed across the globe. As shown in Figure 8, the satisfaction level dropped in April, but it consistently improved through August when it even surpassed the satisfaction level of North America. To have a deeper understanding, we analysed five popular Asian tourist countries separately.

The largest contributor to the data on the Asian continent is India, one of the ancient nations, fastest-growing developing nations in the world, with tourism industry being one of the important sources of its foreign exchange in economic aspects. Its government has taken several measures, framed various policies, and launched several plans and schemes to attract a large amount of inbound and outbound tourist base, laying a platform to exploit its contributing ability in economic growth and at the same time compete with Europe and the USA in the field of tourism. The hospitality sector too has improved significantly to support the government initiatives.

Looking at the polarity values that provide the measure of the hospitality facilitated at present, especially during the COVID-19 pandemic, we observe that there has been a tremendous decline in the number of reviews, indicating a decline in footfall of customers at hotels. This can be attributed to the irregularity in the spread of pandemic across the country, varying both in time and rate of spread across various regions that have led to a decline in not only the international tourist visit but also restricted domestic travel within the country. Indian railway that transported 23 million passengers in 2019 is still not fully operational. A densely large population led to an increase in the lockdown period up to six months, thereby affecting the economy-boosting sector of its GDP severely. Yet, the standards of hospitality served have remained nearly steady. Nearly 17% of the consumers provided review in ED and D categories cumulatively in the months of July, August, and September, and close to 60% of the reviews belong to S and ES categories which are highest among Asian countries considered in this study. A significant drop in the satisfaction level is observed (see Figure 8) in April when the first lockdown was implemented. After April, impressively, the satisfaction level consistently kept improving through August. Overall, the average sentiment score (nearly 17%) for the country has been the second lowest among all the countries and continents of the study, only ahead of the African continent. This insight obtained from sentiment analysis portrays that the country lacks in terms of providing premium level infrastructure and hospitality experience, which is expected from top-rated luxury hotels, also evident from the reviews. This provides the answer as to why despite having a higher proportion of ES and S category reviews, the overall sentiment score is not commendable.

Furthermore, the extremely low number of reviews in Indonesia and Singapore from July to September, when other countries saw the surge in tourist movement, can be attributed to the increasing spread of pandemic throughout that period in both the countries. The pandemic remained more dormant in the other two countries, Malaysia and Sri Lanka until the end of September with the exception of some peaks in their pandemic count chart which led to a significant increase in domestic tourism in these countries. While in both Malaysia and Sri Lanka the average polarity scores remained almost the same when compared to the initial month of the year 2020, Sri Lanka secured the highest average score in the world. It also accounted for 80% of the reviews in ES and S categories, the highest proportion across the world. As shown in Figure 8, the average satisfaction level in Sri Lanka was the highest during January–March, and they again dominated in later part of the year, July–September. Review was not available in April and May, but it observed low satisfaction in June. In terms of the satisfaction level, Singapore did exceptionally well in July and August, and it fell behind Sri Lanka in September. On the other hand, small sample from Indonesia and Singapore restricts us to conclude anything convincingly about their trend, despite these countries being popular tourist destinations.

Voice of customer

Inspired by the theory of boundary spanning, top managers of hotels can sync their operations and services with changing requirements and expectations of customers. In this direction, to help the hospitality sector in recognizing the changing requirements and expectations with crisp and adequacy, we exploit the usage of topic modelling and sentiment analysis that bring insights from the reviews in the form of keywords (or topics) and satisfaction level and thereby facilitating quick analysis of the customer expectations and requirements from the top tier hotels.

We call all words which have direct implications on defining the hospitality experience of the customer as 'drivers' of hospitality. Table 5 highlights the list of all such drivers along with their probability of occurrence or, in other words, their frequency, in all the category of reviews combined and in reviews belonging to ED and D categories taken together. Looking at Topic 2, i.e. stay experience, the one with the highest weightage and therefore made up of words with most frequent usage, we observe that stay, hotel, staff, and good are top words to describe the hospitality experience at the hotel. As shown in Table 5, a whopping proportion of over 43.3% and 41.3% reviewers complained about unsupportive staff and poor service, respectively.

Other important areas of dissatisfaction include time and book of the topic 10, routine; clean, room, and restaurant of topic 9, hotel infrastructure; fast of topic 5, booking experience; water of the topic 6, beverage on arrival; and delicious of topic 8, event experience. Time signifies time taken to deliver a service, hence, measuring the responsiveness of the staffs, and book implies the availability and booking of specific types of rooms and resources available at hotels. Noticeable dissatisfaction related to hotel infrastructure was due to issues with room (49.8%), cleanliness (13%), and restaurant (13.8%). Significant dissatisfaction was also due to booking experience was not fast (15.6%), only water or even water was not offered on arrival, and food was not delicious during event (1.3%). Customers are observed expressing their experience about the pandemic response (i.e. topic 1) by hotel. They seem discussing about accessibility and arrangements made at the entrance and parking area. These drivers highlight the importance of touchpoints that impact the experience of the customer and display the criteria that the decision makers could use during their planning stages. These drivers can be used to improve the hospitality of customers.

Managerial implications

Observing strengths and shortcomings in hotel operations through customers' sentiments, is important for hotel managers to become boundary spanners by scanning and noticing the speed of changes in external environment (Cheng et al., 2020). As changes in external environment can bring opportunities as well threat, managers must transform their existing hotel operations by considering the trend of customer sentiments (Olsen, 1996), especially in situations such as COVID-19. We highlight some key takeaways tailor-made for these managers.

Before the COVID-19 situation eases, hotel industry would get time to prepare themselves to address the new expectations of customers alongside improving the existing services. They will also get opportunity to put the new plans and services into operations on a small number of inflow of guests after the COVID-19 situation eases. In this fashion, the hotel industry can portray their efficiency and dynamic capability by exhibiting the quality of hospitality that is expected by customers, mandated by governments and social responsibility. This would be useful as the pandemic has been rising in different regions at different rates at different points of time. The cyclic nature of rise and fall of infection, termed as 'waves', would allow such preparedness action plans to test hotels' resilience and service effectiveness until the final wave of dissemination recedes.

Managers can try to understand protocols and special arrangements made by European and North American hotels which are observed doing well, and hence it resulted in achieving zero reviews in ED categories (see Figure 7). Sri Lanka as a country did well in terms of satisfaction level as shown in Figure 8. As discussed above, the relatively small size of Sri Lanka's tourist base lays down a good picture for managers to observe ways of providing high levels of hospitality in hotels of Sri Lanka. As the uplift of lockdown will initially bring customers in smaller number, managers can easily experiment by deploying practices followed in Sri Lanka's hotels. If these practices bring more customer satisfaction, they can be deployed later over a larger population when the volume of footfall increases with time, especially in countries like Australia, Indonesia, Singapore, and India.

Although there are abundant tourist spots in India, the country faces a potential threat of losing revenue from its domestic and international tourists to the neighbour country, Sri Lanka, as they seem to provide much better hospitality at hotels (see Figure 8). Managers in India must focus aggressively on improving hospitality, especially safety concern during the COVID-19 pandemic, to dissipate competition from their neighbour country and attract its large domestic population that flies outside for vacation and recreational purpose as the majority of them flies to countries within its continent like Indonesia, Singapore, and Malaysia. As shown in Figure 8, Indonesia could maintain the satisfaction level without any significant drop in any month of the year of the pandemic. However, Malaysia with a significant drop in the satisfaction level in June could also reasonably maintain customer satisfaction.

Initial drop in polarity scores, in general indicates the lack of preparedness upon the arrival of customers in India's hotels after the uplift of lockdown, can be turned around and managers must take necessary steps by understanding customer requirements and expectations and ensure that all the facilities are sufficient. This will help the hotels worldwide to provide satisfactory experience after the second wave of coronavirus settles down. This will also help Asian countries after the suppression of the first wave.

As shown in Tables 4 and 5, staff, service, overall experience, booking, cleanliness, room, restaurant, slow booking, beverage on arrival, food at events, and hotel response to pandemic are major reasons of dissatisfaction. Managers of hotels should improve these aspects of hotel operations which will increase the satisfaction level. Staffs must be trained in prior; protocols of operation must be neatly laid down to ensure that the entire workforce is organized under the vision of improving hospitality. Clean atmosphere, premium aesthetics, and service with minimal contact and quick servicing must be ensured, as these were key drivers of sentiments from July to September.

This study helps hotels in considering the potential impact of negative word of mouth and negative e-word of mouth during COVID-19 recovery periods that could cause them losing customers. There is a growing concern about safety and facilities at public spaces after pandemic (Al-Natour & Turetken, 2020).

Conclusions

We analysed the hotel reviews on TripAdvisor using the combination of topic modelling and sentiment analysis and exploited the intricacies of the reviews by developing a new scale of metrics, the first of its kind applied in this field of research. The scale better categorizes the review and feedback of the customers rather than rating them simply on a polarity scale, thereby covering the gap in analysing the hospitality experience from customer reviews more effectively. We attempted to jot down more accurate inferences by using the combination of sentiment analysis and keywords extracted from topic modelling, which give direction to managers and decision makers to boost their preparedness in dealing with COVID-19 pandemic-like uncertainties.

Overall, we observed a significant drop in the satisfaction level during April, May, or June when perhaps fear due to the COVID-19 pandemic was at peak. Among continents, North America and Europe performed exceptionally well in terms of customer satisfaction level. Almost always North America maintained its domination in customer satisfaction, and no customer found extremely dissatisfied. Among Asian countries, customer satisfaction in Sri Lanka was the highest. Topic modelling revealed 12 topics including hotel response to pandemic as most often discussed by customers. The analysis also revealed that staff, overall service, cleanliness, room, booking experience, and time are main sources of dissatisfaction. Performing the study to analyse the conditions in this sector during the pandemic encourages us to imbibe better resilience in handling emergency like the COVID-19 pandemic that prevailed in the top hotels across the world and helped us extract possible solutions that hotels must adapt and ensure increase in customer satisfaction levels. From the perspective of the theory of boundary spanning, we recommend managers and decision makers to keep scanning the external changes by deploying topic modelling and sentiment analysis and satisfy customers by improving hotel operations and services.

The study can be extended to other sectors of tourism industry which require breaking constricted mindsets of the management which claims to be superior and does not take the customer-centric approach into account. It can also be extended to analyse any form of review to test the new scale of metrics and determine its efficiency and reliability. The study was performed over a limited period, and the inferences were drawn considering the circumstances of pandemic that defined the nature of reviews. The analysis also excluded the period of March to June and in some cases April to June due to the unavailability and low amount of data during the lockdown period. The better analysis of trend can be deduced by taking data of continuous time periods to analyse the reviews in different business fields.

Acknowledgements

We thank Ms. Chahat Raj of Delhi Technological University who supported in refining the programme codes of this study.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Notes on contributors

Mihir P Mehta is pursuing MBA at IIM Raipur after the completion of his Bachelor of Technology. He is interested in computer application in business and management.

Gopal Kumar is an Assistant Professor of Operations Management at IIM Raipur. He received his PhD from IIT Kharagpur and postdoctoral fellowship at the Dublin City University, Ireland. His research articles have been published in *International Journal Production Economics*, *Journal of Business and Industrial Marketing*, *Journal of Cleaner Production*, *Benchmarking*: an international journal, *IIMB Management Review*, *Measuring Business Excellence*, *International Journal of*

Productivity and Performance Management, and *International Journal of Services and Operations Management*. His primary research interests include supply chain collaboration, sustainable and green supply chain, and operations management.

M. Ramkumar is currently working as an Assistant Professor in the Department of Operations Management at the Indian Institute of Management Raipur. He was a Postdoctoral Researcher at the Chair of Logistics Management at the Swiss Federal Institute of Technology Zurich, Switzerland. His research is interdisciplinary and lies on the interface between operations management and information systems, and encompasses supply chain technologies, supply chain sustainability, and humanitarian operations. His research has been published in international journals such as *Service Science (INFORMS)*, *International Journal of Production Economics*, *International Journal of Production Research*, *Production Planning & Control*, *Journal of Cleaner Production, Annals of Operations Research*, *Computers & Industrial Engineering*, *IEEE Systems Journal*, and other peer-reviewed journals.

References

Al-Natour, S., & Turetken, O. (2020). A comparative assessment of sentiment analysis and star ratings for consumer reviews. *International Journal of Information Management*, *54*, 102132. https://doi.org/10.1016/j.ijinfomgt.2020.102132

Alamoodi, A., Zaidan, B., Zaidan, A., Albahri, O., Mohammed, K., Malik, R., Almahdi, E., Chyad, M., Tareq, Z., Albahri, A., Hameed, H., & Alaa, M. (2020). Sentiment analysis and its applications in fighting COVID-19 and infectious diseases: A systematic review. *Expert Systems with Applications*, *167*, 114155. https://doi.org/10.1016/j.eswa.2020.114155

Bastani, K., Namavari, H., & Shaffer, J. (2019). Latent Dirichlet allocation (LDA) for topic modeling of the CFPB consumer complaints. *Expert Systems with Applications*, *127*, 256–271. https://doi.org/10.1016/j.eswa.2019.03.001

Belford, M., & Greene, D. (2020). Ensemble topic modeling using weighted term co-associations. *Expert Systems with Applications*, *161*, 113709. https://doi.org/10.1016/j.eswa.2020.113709

Bird, S., Loper, E., & Klein, E. (2009). *Natural language processing with Python*. O'Reilly Media Inc.

Blei, D. M., Ng, A. Y., & Jordan, M. I. (2003). *Latent Dirichlet Allocation*, *3*, 30.

Borg, A., & Boldt, M. (2020). Using VADER sentiment and SVM for predicting customer response sentiment. *Expert Systems with Applications*, *162*, 113746. https://doi.org/10.1016/j.eswa.2020.113746

BudgetDirect. (2020). *Australian tourism industry 2020: An in-depth analysis of the latest tourism statistics for Australia*. Budget Direct. https://www.budgetdirect.com.au/travel-insurance/research/tourism-statistics.html.

Cetin, G., & Walls, A. (2016). Understanding the customer experiences from the perspective of guests and hotel managers: Empirical findings from luxury hotels in Istanbul, Turkey. *Journal of Hospitality Marketing & Management*, *25* (4), 395–424. https://doi.org/10.1080/19368623.2015.1034395

Cheng, M., Zhang, G., & Wong, I. A. (2020). Spanning across the boundary of airbnb host community: A network perspective. *International Journal of Hospitality Management*, *89*, 102541. https://doi.org/10.1016/j.ijhm.2020.102541

Cooper, C. (2015). Managing tourism knowledge. *Tourism Recreation Research*, *40*(1), 107–19. https://doi.org/10.1080/02508281.2015.1006418

Das, S. S., & Tiwari, A. K. (2020). Understanding international and domestic travel intention of Indian travellers during COVID-19 using a Bayesian approach. *Tourism Recreation Research*, 1–17. https://doi.org/10.1080/02508281.2020.1830341

Faus, J. (2020). *This is how coronavirus could affect the travel and tourism industry*. World Economic Forum. https://www.weforum.org/agenda/2020/03/world-travel-coronavirus-covid19-jobs-pandemic-tourism-aviation/.

Filimonau, V., Derqui, B., & Matute, J. (2020). The COVID-19 pandemic and organisational commitment of senior hotel managers. *International Journal of Hospitality Management*, *91*, 102659. https://doi.org/10.1016/j.ijhm.2020.102659

Godovykh, M., & Tasci, A. D. A. (2020). Customer experience in tourism: A review of definitions, components, and measurements. *Tourism Management Perspectives*, *35*, 100694. https://doi.org/10.1016/j.tmp.2020.100694

Hall, C. M., Scott, D., & Gössling, S. (2020). Pandemics, transformations and tourism: Be careful what you wish for. *Tourism Geographies*, *22*(3), 577–98. https://doi.org/10.1080/14616688.2020.1759131

Hao, F., Xiao, Q., & Chon, K. (2020). COVID-19 and China's hotel industry: Impacts, a disaster management framework, and post-pandemic agenda. *International Journal of Hospitality Management*, *90*, 102636. https://doi.org/10.1016/j.ijhm.2020.102636

Hasan, A., Moin, S., Karim, A., & Shamshirband, S. (2018). Machine learning-based sentiment analysis for twitter accounts. *Mathematical and Computational Applications*, *23* (1), 11. https://doi.org/10.3390/mca23010011

Hu, N., Zhang, T., Gao, B., & Bose, I. (2019). What do hotel customers complain about? Text analysis using structural topic model. *Tourism Management*, *72*, 417–426. https://doi.org/10.1016/j.tourman.2019.01.002

Jia, Z., Shi, Y., Jia, Y., & Li, D. (2012). A framework of knowledge management systems for tourism crisis management. *Procedia Engineering*, *29*, 138–143. https://doi.org/10.1016/j.proeng.2011.12.683

Kuhn, K. D. (2018). Using structural topic modeling to identify latent topics and trends in aviation incident reports. *Transportation Research Part C: Emerging Technologies*, *87*, 105–122. https://doi.org/10.1016/j.trc.2017.12.018

Liu, Y., Jiang, C., & Zhao, H. (2019). Assessing product competitive advantages from the perspective of customers by mining user-generated content on social media. *Decision Support Systems*, *123*, 113079. https://doi.org/10.1016/j.dss.2019.113079

Loria, S., Keen, P., Honnibal, M., & Yankovsky, R. (n.d.). *TextBlob: Simplified text processing – TextBlob 0.16.0 documentation*. Retrieved January 5, 2021, from https://textblob.readthedocs.io/en/dev/.

Lund, S., Manyika, J., Woetzel, J., Barriball, E., Krishnan, M., Alicke, K., Birshan, M., George, K., Smit, S., Swan, D., & Hutzler, K. (2020). *Risk, resilience, and rebalancing in global value chains | McKinsey*. McKinsey Global Institute. https://www.mckinsey.com/business-functions/operations/our-insights/risk-resilience-and-rebalancing-in-global-value-chains#.

Ma, E., Cheng, M., & Hsiao, A. (2018). Sentiment analysis – A review and agenda for future research in hospitality contexts. *International Journal of Contemporary Hospitality Management*, *30*(11), 3287–308. https://doi.org/10.1108/IJCHM-10-2017-0704

Melián-Alzola, L., Fernández-Monroy, M., & Hidalgo-Peñate, M. (2020). Hotels in contexts of uncertainty: Measuring organisational resilience. *Tourism Management Perspectives*, *36*, 100747. https://doi.org/10.1016/j.tmp.2020.100747

Nguyen, D. N., Imamura, F., & Iuchi, K. (2017). Public-private collaboration for disaster risk management: A case study of hotels in Matsushima, Japan. *Tourism Management*, *61*, 129–140. https://doi.org/10.1016/j.tourman.2017.02.003

NLTK. (n.d.). *Natural language toolkit—NLTK 3.5 documentation*. Retrieved January 7, 2021, from https://www.nltk.org/.

OECD. (2020). *Tourism policy responses to coronavirus (Covid-19) – OECD*. https://read.oecd-ilibrary.org/view/?ref=124_124984-7uf8nm95se&title=Covid-19_Tourism_Policy_Responses.

Olsen, M. D. (1996). Events shaping the future and their impact on the multinational hotel industry. *Tourism Recreation Research*, *21*(2), 7–14. https://doi.org/10.1080/02508281.1996.11014770

Padma, P., & Ahn, J. (2020). Guest satisfaction & dissatisfaction in luxury hotels: An application of big data. *International Journal of Hospitality Management*, *84*, 102318. https://doi.org/10.1016/j.ijhm.2019.102318

Polyzos, S., Samitas, A., & Spyridou, A. E. (2020). Tourism demand and the COVID-19 pandemic: An LSTM approach. *Tourism Recreation Research*, 1–13. https://doi.org/10.1080/02508281.2020.1777053

Rasul, T., & Hoque, M. R. (2020). A review of social media research in the tourism and hospitality Industry through the lens of social structure. *Tourism Recreation Research*. https://www.tandfonline.com/doi/pdf/10.1080/02508281.2020.1721140?redirect=1.

Rauch, D. A., Collins, M. D., Nale, R. D., & Barr, P. B. (2015). Measuring service quality in mid-scale hotels. *International Journal of Contemporary Hospitality Management*, *27*(1), 87–106. https://doi.org/10.1108/IJCHM-06-2013-0254

Rehurek, R., & Sojka, P. (2010). *Software framework for topic modelling with large corpora*. ELRA. http://radimrehurek.com/gensim.

Rhee, H. T., & Yang, S.-B. (2015). Does hotel attribute importance differ by hotel? Focusing on hotel star-classifications and customers' overall ratings. *Computers in Human Behavior*, *50*(C), 576–587. https://doi.org/10.1016/j.chb.2015.02.069

Schmiedel, T., Müller, O., & vom Brocke, J. (2019). Topic modeling as a strategy of inquiry in organizational research: A tutorial with an application example on organizational culture. *Organizational Research Methods*, *22*(4), 941–968. https://doi.org/10.1177/1094428118773858

Sparks, B. A., So, K. K. F., & Bradley, G. L. (2016). Responding to negative online reviews: The effects of hotel responses on customer inferences of trust and concern. *Tourism Management*, *53*, 74–85. https://doi.org/10.1016/j.tourman.2015.09.011

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.2013.09.007

Tsai, C. H., Linliu, S. C., Chang, R. C., & Mak, A. H. (2020). Disaster prevention management in the hotel industry: Hotel disaster prevention literacy. *Journal of Hospitality and Tourism Management*, *45*, 444–455.

Tushman, M. L., & Scanlan, T. J. (1981). Boundary spanning individuals: Their role in information transfer and their antecedents. *The Academy of Management Journal*, *24*(2), 289–305. https://doi.org/10.2307/255842

Wang, J., & Wu, X. (2018). Top-down or outside-in? Culturally diverse approaches to hotel crisis planning. *Journal of Hospitality and Tourism Management*, *36*, 76–84. https://doi.org/10.1016/j.jhtm.2018.08.002

WHO. (2020). *WHO coronavirus disease (COVID-19) dashboard*. https://covid19.who.int.

Yeh, S.-S. (2020). Tourism recovery strategy against COVID-19 pandemic. *Tourism Recreation Research*, 1–7. https://doi.org/10.1080/02508281.2020.1805933

Appendix A

Table A1. Category-wise average polarity score of continents.

Polarity score category	Continents	Jan	Feb	Mar	Apr	May	June	July	Aug	Sep
Average polarity score	Africa	17.6	15.0	18.3	NA[a]	NA	NA	15.2	16.0	10.8
	Australia	16.1	16.5	14.6	16.3	12.7	13.3	14.0	15.7	13.7
	Europe	20.5	21.9	22.4	11.0	21.0	22.2	19.8	16.9	19.2
	N. America	20.5	21.9	22.4	11.0	21.0	22.2	19.8	16.9	19.2
	Asia	17.8	19.0	20.5	12.8	14.6	15.3	19.5	20.9	18.4
Average ED polarity score	Africa	−16.3	−15.4	−18.5	NA	NA	NA	NA	NA	−22.8
	Australia	−12.0	−16.0	−12.0	NA	NA	−13.5	NA	NA	−15.5
	Europe	NA	NA	NA	NA	NA	NA	NA	NA	NA
	N. America	NA	NA	NA	NA	NA	NA	NA	NA	NA
	Asia	−14.3	−13.0	−14.4	NA	NA	−11.0	−21.0	−13.0	−11.0
Average D polarity score	Africa	3.9	2.6	4.4	NA	NA	NA	4.0	2.3	−0.6
	Australia	4.6	5.4	5.3	6.5	4.0	4.2	4.2	5.2	2.6
	Europe	3.6	4.5	5.6	NA	NA	NA	3.5	3.9	3.1
	N. America	5.0	4.8	2.0	NA	NA	−1.0	5.2	3.7	4.1
	Asia	5.4	4.9	4.0	−0.5	5.8	3.3	6.0	4.3	4.8
Average N polarity score	Africa	12.2	12.3	11.9	NA	NA	NA	12.0	11.0	11.9
	Australia	12.2	12.4	12.0	12.3	12.3	12.2	11.9	12.4	12.3
	Europe	12.1	12.6	12.1	NA	NA	NA	11.8	11.6	12.0
	N. America	13.0	12.1	12.0	11.0	NA	12.3	11.0	11.0	11.7
	Asia	12.2	12.1	12.0	12.0	11.2	12.6	12.2	12.3	12.1
Average S polarity score	Africa	25.6	25.4	23.7	NA	NA	NA	24.0	29.0	25.5
	Australia	21.9	22.5	20.8	22.0	18.8	19.9	20.6	21.6	21.1
	Europe	24.5	23.7	26.8	NA	NA	NA	25.4	23.7	26.4
	N. America	23.1	24.0	25.2	NA	21.0	24.1	24.6	21.0	24.5
	Asia	21.9	22.8	24.1	17.7	22.0	22.4	23.8	23.2	22.8
Average ES polarity score	Africa	65.4	90.0	66.7	NA	NA	NA	NA	NA	50.0
	Australia	50.0	NA	NA	NA	NA	NA	NA	52.0	54.0
	Europe	59.0	68.3	58.0	NA	NA	NA	58.0	63.7	63.0
	N. America	72.8	65.3	50.0	NA	NA	69.5	66.8	66.3	69.0
	Asia	64.4	73.3	65.5	NA	NA	54.0	63.4	77.4	62.0

[a]NA indicates no reviews available in that category.

Table A2. Category-wise average polarity score of Asian countries.

Polarity score category	Countries	Jan	Feb	Mar	Apr	May	June	July	Aug	Sep
Avg. P. score	India	17	17.2	19.5	10.8	13.2	15.6	15.1	17.7	17.2
	Indonesia	18.2	15.3	16.6	16.5	16.7	20	17.9	17.6	13.8
	Malaysia	17.4	20.4	18.1	17	17.3	11.1	19.2	18.5	17.5
	Singapore	19	22.6	23	11.9	15	NA[a]	33.8	41.2	22.3
	Sri Lanka	24.1	24.4	23.7	NA	NA	15.8	26.8	23.6	24.1
ED	India	−14.5	−14	−14.7	NA	NA	−11	−25.5	−13	NA
	Indonesia	−13	NA	NA	NA	NA	NA	NA	NA	NA
	Malaysia	−15.5	NA	−12	NA	NA	NA	NA	NA	−11
	Singapore	−11	NA	−16	NA	NA	NA	NA	NA	NA
	Sri Lanka	−16	−11	NA	NA	NA	NA	−16.5	NA	NA
D	India	6.1	5.4	4.3	−8	7	3	6.4	4.3	5.7
	Indonesia	5.1	4.6	6.8	NA	8	NA	7	1.5	1
	Malaysia	4	3.9	3.1	NA	−6	1.3	4.7	5.1	3.3
	Singapore	4.1	4.9	3.3	7	NA	NA	7	2.6	2.3
	Sri Lanka	4	2.7	4.3	NA	NA	7.3	5	5.3	7
N	India	12.2	12	11.9	13	11.5	12.6	12.4	12.5	12.2
	Indonesia	12.5	12	11.9	10.5	NA	11	14	12	13
	Malaysia	12.3	12	12.2	NA	NA	13.3	11.9	12.5	12.4
	Singapore	12.1	12.1	12.2	12.2	10	NA	13	11.8	11.4
	Sri Lanka	12	13.1	11.8	NA	NA	12.6	11.6	12.1	11.8
S	India	20.8	21.5	23.3	15.7	22.7	22.6	21.7	21.1	21.4
	Indonesia	23.1	21.1	20	22.5	21	29	18.4	23	19
	Malaysia	24.4	25.3	24.3	17	22	22	23.8	22.3	24.4
	Singapore	24.2	25.3	25.6	15	20	NA	33.7	38.5	32
	Sri Lanka	26.9	25.2	26.2	NA	NA	21.1	26.1	25.6	25.9
ES	India	58.9	64.4	63.4	NA	NA	54	NA	79.7	52
	Indonesia	69.5	NA	NA	NA	NA	NA	51	NA	NA
	Malaysia	67.3	65.8	64	NA	NA	NA	53	57.7	NA
	Singapore	57.6	74.4	71.2	NA	NA	NA	58	86.3	71
	Sri Lanka	77.6	78.4	60.3	NA	NA	NA	67.6	72.5	54

[a]NA indicates no reviews available in that category.