

# Exploring Customer Feedback on Their Hotel Experiences in Vietnam

Ha Thi Thu Nguyen, Department of Greenwich, FPT University, Vietnam

 <https://orcid.org/0000-0003-4414-6944>

Thao Phan Huong, Management Accounting Department, Thuongmai University, Vietnam\*

Anh Le Thi Tram, Accounting-Auditing Department, Thuongmai University, Vietnam

Thao Viet Tran, Department of Research Administration, Thuongmai University, Vietnam

## ABSTRACT

Over the past twenty years, customer experience has attracted the attention of researchers and business executives. For the hotel service industry, understanding customer experience becomes a necessity, as it is one of the top goals for hotel survival. With the development of e-commerce and the globalized hotel industry, customers are easily shared on online booking sites, making hotel managers work harder to design a good customer experience management plan. This chapter proposes an approach to analyzing data from customers' online reviews with their experiences to understand their emotional and psychological states after using Vietnamese hotel services. The Python language is used for statistical analysis of these data, and the Vader library measures customers' positive and negative views after the hotel service experience. The results show that most customers are satisfied with Vietnamese hotel services, and less than 10% are dissatisfied with aspects such as staff, price, check-in, and location.

## KEYWORDS

Customer Experience, Data Analysis, Hotel Service, Online Review, Python, Tripadvisor

## 1. INTRODUCTION

The COVID-19 pandemic dealt a severe blow to Vietnam's tourism industry in 2020, causing losses of approximately 23 billion USD and an 80% drop in international visitors compared to the previous year. As the country moves towards a post-pandemic green economy, it is imperative that businesses in the hotel and tourism service sectors not only ramp up their digital transformation efforts, but also take the lead in enhancing customer experience (CX) through in-experience design. Customer experience is a complex and multi-dimensional concept, with definitions ranging from psychological to marketing and economic perspectives (Akhtar et al., 2017; Rahimian et al., 2021). Effective customer experience management is crucial for businesses, as it involves understanding customer emotions and designing strategies to improve the quality of their experience (Luturlean & Anggadwita, 2016). By utilizing customer feedback, businesses can identify areas for improvement and enhance

DOI: 10.4018/IJEEL.330023

\*Corresponding Author

This article published as an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0/>) which permits unrestricted use, distribution, and production in any medium, provided the author of the original work and original publication source are properly credited.

product quality or personal customer care. In the current literature, researchers have confirmed that the difference in prices, products, and services is no longer a sustainable competitive strategy of businesses, but “customer experience” is a new strategy for making a “difference” (Akhtar, Zubair, Kumar, & Ahmad, 2017; Mohamed, 2021). Measuring customer experience has become a critical component of a business’s digital transformation, as outlined in the six main pillars of the Ministry of Information and Communications in Vietnam. Companies must prioritize customer experience management to stay competitive (Luturlean & Anggadwita, 2016; Mohamed, 2021; Paulose & Shakeel, 2022; Rahimian et al., 2021; Sampetua Hariandja & Vincent, 2022).

The hotel industry is flourishing, with growth and development observed (Blomberg-Nygaard, & Anderson, 2016). To cater to the diverse needs of consumers, hotels are introducing a wide range of services and business models (Luturlean & Anggadwita, 2016; Rahimian et al., 2021). In today’s hospitality industry, customer experience is a critical factor influencing customer loyalty, reputation, and revenue (Paulose & Shakeel, 2022; Rahimian et al., 2021; Camilleri & Filieri, 2023). As a result, scholars and hotel managers have increasingly focused on researching customer experience in the hotel industry. Customer experience research in the hotel industry evaluates the impact of customer experience on emotions, psychology, perspectives, and customer relationship building, just like in other industries (Sampetua Hariandja & Vincent, 2022).

Online booking sites have revolutionized the tourism and hospitality industries, with data analytics from customer-generated reviews playing a crucial role (Chalupa & Petricek, 2022; Chen et al., 2019). Recent research has shown that big data analysis of these reviews offers valuable insights that enable businesses to adapt to environmental changes and develop effective long-term strategies. By utilizing data analytics, companies can extract meaningful value from vast data, empowering them to make informed, data-driven decisions (Chen et al., 2019; Chen et al., 2020; Godnov & Redek, 2018). However, analyzing customer reviews in text form can be challenging, requiring the identification of keywords and patterns that reflect customer experience (Li et al., 2019). To this end, conducting text analysis of hotel experiences provides valuable insights into the genuine emotions experienced by customers. By leveraging this information, businesses can improve their offerings and increase customer satisfaction (Zhao et al., 2019).

Customer-generated data can prove pivotal to business operation (Chen et al., 2020). However, difficulties still exist. First, data collection can only use manual methods when the number of online reviews is manageable. Secondly, it is challenging to discover many online reviews by manual methods (Barnes et al., 2020; Liu et al., 2017). Many studies have mentioned collecting and mining this data to analyze the emotions of customers experiencing customer satisfaction or dissatisfaction expressed in a positive or negative emotion (Barnes et al., 2020; Roy, 2023; Song et al., 2022; Wu et al., 2023). Some techniques are related to deep learning, natural language processing, or statistical machine learning to mine customer opinions (Akhtar et al., 2017; Adjei Peter & Decui, 2022; Wu et al., 2022; Leal et al., 2019; Alrawadieh & Law, 2019). Although these techniques are modern, the implementation process is complicated, and the techniques and learning models are challenging to understand and apply in practice. On the other hand, those methods mostly use processed data, so they become even more remote upon actual application. Currently, hotel managers want to find suitable and straightforward approaches to solve part of the decision-making problem in their management. Moreover, not all hotels have the conditions to have IT support, so finding a simple data collection and analysis solution becomes the most useful for the job. When hotel managers understand the customer’s experience, changing the management model or improving hotel service quality is more attainable.

To address these limitations, we recognize the importance of establishing a comprehensive data collection process that utilizes simple and suitable analysis tools. In this study, we first show a simple process to analyze online reviews, including two phases: collection data and analysis data. Next, the study uses WebHarvy to gather online reviews from TripAdvisor for hotels in six major cities in Vietnam, namely Hanoi, Ho Chi Minh City, Quy Nhon, Nha Trang, Hue, and Da Nang, resulting in a total of 20,550 reviews. To measure customer satisfaction for Vietnamese hotels, the authors use

the Python language's Vader library for data analysis, enabling us to analyze this data to measure customer satisfaction, and then identify the hotel services that customers most highly regard. The study also aims to answer two questions: Q1. "Is it possible to understand the customer experience of Vietnamese hotels through online reviews by an easier process?" and Q2: "How many customers are satisfied with the Vietnamese hotel service?"

The remaining part of the paper is organized as follows: Section 2 thoroughly examines customer experience with hotel services through related research, while Section 3 provides an explicit breakdown of the research methodology. Section 4 presents the measured results and delivers the concluding remarks in Section 5.

## **2. LITERATURE REVIEW**

### **2.1 Hospitality Industry Data and Analysis Techniques**

Online booking systems provide the travel industry with valuable big data. Hotel managers offer millions of records of descriptive data, while customer reviews act as a social network and offer additional unstructured data (D'Acunto et al., 2023; Dong et al., 2020; Gómez-Suárez & Veloso, 2022). Analyzing this data benefits hotel managers and customers, allowing for a deeper understanding of the industry (Bian et al., 2022; Mohamed, 2021; Xu et al., 2022).

The utilization of big data has become a cost-effective means of collecting information from computer databases for analysis and extraction of valuable insights (Padma & Ahn, 2020). This method is faster and less challenging compared to manual surveys. As businesses aim to understand customer psychology, objectively, they are turning to extensive data collection, instead of limiting questions on a ballot. However, mining such vast amounts of data is no longer manual, and conventional methods must be revised. For instance, collecting and analyzing millions of online hotel reviews using traditional methods is unfeasible. Therefore, data analysis is shifting towards utilizing data mining models, computer statistics, and artificial intelligence to enhance accuracy and efficiency.

To fully comprehend customer needs and enhance service quality, hotels utilize several data analysis techniques in data mining. These include emotion analysis, which allows for a deeper understanding of customers' psychological states and opinions; aspect extraction, which focuses on determining the most important and preferred aspects of hotel services; and complaint detection, which helps identify and address customer grievances (Chen et al., 2019).

Analyzing customer sentiment using NLP is being researched by companies and organizations to optimize personalized customer care. Using sentiment analysis to identify the feelings behind the words will provide a lot of knowledge about customer behavior and their choices to make better management decisions (Alaei et al., 2019)

Sentiment and opinion mining is a new approach to understanding the opinions and emotions of customers, as most transactions today are done online. This is the first step for researching customer emotions and measuring customer satisfaction with the business's brand, products, and services. This allows the exploration of research perspectives on opinions, feelings, subjective conceptions, evaluations, attitudes, appraisals, emotions, etc., expressed in the text. These are expressed through comments, blogs, discussions, news, feedback, etc., or other materials (Alaei et al., 2019; Dong et al., 2020; Gómez-Suárez & Veloso, 2022). Opinion mining is based on computational languages, information retrieval, text data mining, natural language processing, machine learning, statistics, and predictive analytics.

### **2.2 Analyzing Data From Online Reviews**

Research on data mining generated by customers on the Internet is one of the research directions on hotel experience in recent years (Mohamed, 2021). Yabing Zhao researched customer-generated online big data (Zhao, Xu, & Wang, 2019). When looking for hotel information, online reviews are

an excellent source of detail. They cover a range of features such as cleanliness, location, rooms, services, sleep quality, and facilities. By examining these reviews, hotel managers can see how satisfied customers are with the service they receive (Wu et al., 2023; Gómez-Suárez & Veloso, 2022). Customer satisfaction is measured by comparing perceived service quality with expected service quality. Researchers have found that analyzing hotel attributes extracted from online reviews is an effective way to assess service quality. Researchers can gain valuable insights into customer satisfaction by studying individual hotel attributes as independent variables and overall satisfaction as the dependent variable (Camilleri & Filieri, 2023).

In studies in the field of customer experience measurement, the first step is to use an automated tool to collect online reviews on booking sites (Song et al., 2022; Wu et al., 2022; Wu et al., 2023; Dong et al., 2014; Hu et al., 2019). Currently, quite a few online booking platforms allow customers to share data. But in recent studies, the data is often collected from Ctrip and TripAdvisor. Ctrip (China's one-stop travel shop) is a leading provider of accommodation booking, transportation ticketing, package tours, corporate travel management, and other travel-related services in China. It has grown rapidly since its founding in 1999 and has become China's largest travel agency (Xu et al., 2022; Wu et al., 2023; Bian et al., 2022). Meanwhile, TripAdvisor is an online travel information and booking site that features reviews, photos, and informative forums about various hotels and resorts around the globe. Users rate the places they have stayed for other users (Hu et al., 2019; Godnov & Redek, 2018; Alrawadieh & Law, 2019). These data platforms provide many customer reviews about hotels they have experienced. Data are collected using various automated tools and formats; common collection tools in studies include Octopus (Wu et al., 2023), Houyi (Song et al., 2022), and other methods. Collected data ranges from several thousand to hundreds of thousands of reviews, or even millions, from hotels in one or more regions (Xu et al., 2022). This is a rather large amount of data, difficult to use manually to mine.

After collecting large amounts of data, researchers must use methods to mine the collected data in different directions. The main method used is text mining or natural language processing (Roy, 2023; Dong et al., 2014; Zhao et al., 2019). Techniques that refer to machine learning, statistical, CNN, Word2vec, or LDA networks to design models for processing this big data include separating words, identifying topics, and grouping topics (Xu et al., 2022). Lexical recognition is a task that extracts frequently used words representing important aspects of the customer experience. These words are categorized into service quality topics that include hotel-related attributes such as "restaurant," "breakfast," "food," "pool," "reservation," "buffet," "lounge," "bar," "check-in," "floor," "view," "cleanliness," "bed," "bathroom," or "staff." In addition, studies have mentioned combining different types of words to explore customers' meanings and emotional states in online reviews (Bian et al., 2022; Zhao et al., 2019; Dong et al., 2014). The tools used to assist in the research are quite diverse, including the programming languages R and Python (Godnov & Redek, 2018; Zhao et al., 2019; Roy, 2023; Li et al., 2019) or tools such as RapidMiner, Kansei Engineering, AFINN Lexicon, or ROST CM6.0 (Wu et al., 2023; Godnov & Redek, 2018; Dong et al., 2014).

The discovery of lexical combinations gives quite vast results, with many different feelings and perspectives of customers. Satisfaction is considered a central task when measuring customer experience for hotel services (Song et al., 2022; Roy, 2023; Godnov & Redek, 2018). Some studies explore customer complaints about hotel services in online reviews (Alrawadieh & Law, 2019). They consider customer satisfaction as a function that depends on many independent factors such as price, location, staff, and room (Roy, 2023; Song et al., 2022; Xu et al., 2022; Wu et al., 2022), from which factors such as service, room, cleanliness, location, and value significantly impact customer satisfaction (Song et al., 2022). While other studies have focused on vocabulary combining rules, (Dong et al., 2014; Li et al., 2019; Zhao et al., 2019) the discovery of key findings from lexical combinations gives more surprising results because it helps to explore many aspects that are more than just measuring customer satisfaction. These studies can discover negative topics based on the statistics of these topics' appearances in negative reviews compared to positive reviews or explore a list

of negative words in customer complaints in separate topics like: facilities (noisy, loud, air, heat, unit, sound, hear), service (smoke, Priceline, double, nonsmoking, twin), location (distance, central, within, block, subway) (Hu et al., 2019). In addition, the satisfaction measurement based on the regression model by combining the independent variables only measures the overall satisfaction level by each aspect, not the satisfaction level. Studies after the 2020s have yielded quite surprising results as the number of people who are satisfied with the service is increasing, and they become more receptive to the services provided, with less complaining about the service. It can also be said that during the pandemic, hotels have also focused on improving service quality, and safety and epidemic prevention measures have also become more intentional (Song et al., 2022; Xu et al., 2022; Bian et al., 2022).

The studies all show that it is possible to use modern tools to measure customer satisfaction by analyzing online reviews. They used many approaches, from statistical tools to modern machine learning, to measure customer experiences (Xu et al., 2022; Song et al., 2022; Bian et al., 2022; Li et al., 2019; Alrawadieh & Law, 2019; Godnov & Redek, 2018). Discovering opinions in customer reviews will help to improve service quality (Godnov & Redek, 2018; Dong et al., 2014). Therefore, the analysis of online reviews is becoming increasingly important in the context of the globalization of the hotel industry. It will make a significant contribution to improving service quality, as well as tourists' choice of destinations. But in these methods, we can see the complexity of the models. It becomes difficult for hotel managers, who often don't have the skills to use high-technology tools. Additionally, there are many hotels managed by individuals who may not have any knowledge about CNN, machine learning or LDA, etc. For Vietnamese hotel managers, a simplified algorithm or process would be more effective to solve their problems when the online review data becomes vast.

In this paper, we try to find a simpler model to process textual data. First, WebHarvy is used to crawl data. WebHarvy is a straightforward and easy-to-use user interface. WebHarvy supports searching all categories and subcategories on a web page, submitting multiple keywords for search forms, and more. As a general web scraping tool, WebHarvy can be used to collect data from any web page, including data from websites in various categories such as E-Commerce (Amazon, eBay, Google Shopping), real estate, and travel (TripAdvisor, Booking.com, Hotels.com), etc. Second, using Python is very convenient when it is open, and many online platforms can interpret and run, such as Kaggle. It is free of cost, and many open-source codes support users. Their libraries can help users extract words, measure overall satisfaction, and identify nouns, adjectives, or adverse text to understand customers' sentiments in detail.

The case studies all refer to analytics on hotel customer experience data from customers on online travel or booking sites. Leveraging the availability of data to uncover customer sentiments, emotions, and perspectives can lead to new business opportunities, understanding customer experiences that help businesses improve service quality and have a better customer care plan, and attracting more tourists to the hotel.

### 3. METHODOLOGY

#### 3.1 Concepts and Procedures for Evaluating Customer Experience

**Definition 1:** Set of customer experience

A Set of customer experiences is a set of guests who have used hotel services in Vietnam and are reviewed on the TripAdvisor system. Each customer has a review  $r_i$ . So, it is expressed by:

$$\mathcal{G} = \{g_1, g_2, \dots, g_m\} \quad (1)$$

**Definition 2: Satisfaction**

Satisfaction of guest  $g_j$  is who wrote then review, and it is measured by the value of sentiment of  $r_i$  with a combination function  $\lambda$ :

$$Sas(g_j) = \lambda \cdot (neg(r_i) + pos(r_i) + neu(r_i)) \tag{2}$$

**Definition 3: Aspect**

Aspects are the set of attributes and services of the hotel provided to guests during their stay at the hotel, such as rooms, staff, and hotel location. Aspects are nouns:

$$A = \{a_1, a_2, \dots, a_m\} \tag{3}$$

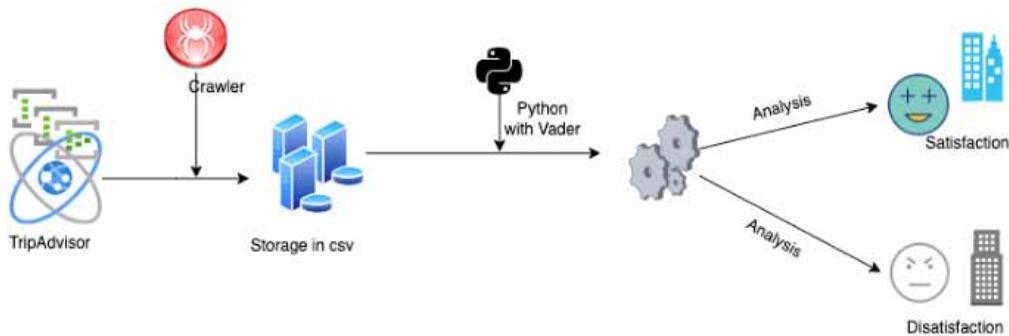
with  $a_k$  is the aspect of the hotel.

Measuring customer experience is measuring the emotions of customers experiencing Vietnamese hotel services through their shares on the TripAdvisor page. Customers' emotions are expressed in satisfaction. Customer satisfaction can be measured in two ways through data analysis: overall satisfaction with all hotel services and satisfaction with each specific service (such as room, staff, meal, and location).

**3.2 Analyzing Customer Experience Sentiment Using the VADER Library**

VADER (Valence Aware Dictionary and Sentiment Reasoner) is a rule-based and vocabulary-based sentiment analysis tool suitable to the sentiment analyzed on social media. It uses a list of lexical features (e.g., words) labeled as positive or negative according to their semantic orientation to calculate text sentiment. Vader returns the probability of a given input sentence being positive, negative, or neutral. Vader is optimized for social media data and can yield good results with data from Twitter, Facebook, and more. As the above results show, the polarity of words and their probabilities are pos, neg neu, and compound words. The process of measuring customer satisfaction in this study is proposed in Figure 1.

Figure 1. Measuring customer satisfaction process



There are five steps in our research process, from data collection to outcome measurement:

- Step 1 - Data collection:** Customer review data about Vietnam hotels is collected by an automatic collection tool from the TripAdvisor site and stored in a .csv file.
- Step 2 - Data storage and processing:** Data is processed by removing non-conforming patterns.
- Step 3 - Use Python language to analyze statistics:** Measure the overall hotel satisfaction from the customer's opinion.
- Step 4 - Using Python with a natural language processing library:** To identify hotel aspects that customers mention and analyze the satisfaction and dissatisfaction from these reviews.
- Step 5 - Measurement and display of results:** Algorithmic models and tools are used in this step to run the quantified data, thereby displaying the results of the method.

### 3.3 Assessing the Satisfaction Level Regarding the Services Provided by the Hotel

The satisfaction rate for all hotel services in Vietnam is determined by calculating the percentage of customers who are completely satisfied. This calculation is performed using the following formula:

$$Sas_{overall} = \frac{\text{number of positive reviews}}{\text{total of reviews}} \times 100\% \quad (4)$$

These are the steps we use to measure:

- Step 1:** Use the Vader library to measure customer reviews by calculating positive, negative, and neutral scores. Then, aggregate these values by a unique score. In which overall customer satisfaction is measured according to formula (2).
- Step 2:** We select positive reviews and use formula (4) to measure customer satisfaction with Vietnamese hotel services.

### 3.4 Assessing the Level of Contentment Regarding Various Aspects of a Hotel

Reviews from customers on online booking sites are presented in the form of unstructured data. This data type can be read and understood but needs to be aggregated to analyze and extract information. These reviews often express customers' opinions on issues related to the service they have experienced at the hotel. Usually, hotel managers want to understand customer psychology through these reviews. They have to read them manually. When the amount of data is too large, it is not easy to read, so a tool is needed to mine the information and analyze it automatically to summarize the results for the manager to know if the customer is satisfied with the service offer or not. Managers can also know which attributes customers are satisfied or unsatisfied with, or what aspect of the hotel makes them happy. This task is also known as aspect-based extraction in customer sentiment analysis.

The customer's opinion is expressed as satisfaction or dissatisfaction in a certain aspect. It has positive or negative connotations. To express positivity or negativity (also called polarity), usually adjectives and accompanying adverbs are analyzed in the sentence. For example, "great" and "good" are two adjectives, and "not" is an adverb. Therefore, extracting attributes and aspects of customer review data about hotel service quality involves finding keywords that represent hotel attributes for measuring quality. Hotel services are mentioned in customer reviews. These attributes and aspects are nouns; if many customers mention them, their frequency will be high.

Call  $R$  is the set of customer reviews  $R = \{r_1, r_2, \dots, r_n\}$ , then the attribute set of the hotel is the set of keywords  $W = \{w_1, w_2, \dots, w_n\}$  that is extracted from  $R$  with  $w_i$  is the noun and ensure that  $f(w_i) > k$ , in which  $f(w_i)$  is the frequency of keywords and  $k$  is a bias of specific value, and it is

used to select an important keyword. This hotel's attributes are combined with several other parameters to become a set of factors affecting customer satisfaction. (Thu, H.N. T. et al., 2020).

The measuring satisfaction of customers for each hotel aspect can be divided into two sub-tasks:

1. Extract the aspects of Vietnamese hotel services that customers need:

This task will select a list of aspects and services of Vietnamese hotels. It is the noun that customers mention in total reviews. It is a set of A:

$$Ext(A) = \{\langle a_1, f_1 \rangle, \langle a_2, f_2 \rangle, \dots, \langle a_n, f_n \rangle\} \quad (5)$$

In which,  $a_i$  are the aspects (services) of Vietnamese hotels and  $f_j$  are the number of occurrences of the mentioned aspect in all the reviews.

2. Percentage of satisfied customers by service aspect.

Measure the rate of customer satisfaction according to the service aspects extracted in the previous step according to the following formula:

$$Sas(a_i) = \frac{\text{number of positive reviews}(a_i)}{\text{total of reviews}(a_i)} \times 100\% \quad (6)$$

where:

- *number of positive reviews*( $a_i$ ): Number of positive reviews with aspect  $a_i$ .
- *total of reviews*( $a_i$ ): Total reviews about aspect  $a_i$ .

This study uses data analysis techniques with Python programming language. The study collected and analyzed data from the online booking page with textual data from customer reviews. From this data source, we build an algorithm model for mining to convert into a quantifiable digital form and measure the results. Extract the service aspects customers were satisfied with according to the steps: (1). Filter positive reviews; (2). Identify and Extract the nouns in the text; (3). Find the nouns that appear with frequencies  $>$  threshold  $\delta$ ; (4). Filter and display the nouns in order of occurrence frequency from high to low, evaluate the aspects that are most satisfied by customers.

## 4. DATA ANALYSIS RESULTS

### 4.1 Data Collection

The study collected online customer reviews from TripAdvisor about 12 hotels in Vietnam with 3-5 star ratings. TripAdvisor is the world's largest travel community, with over 200 million monthly visits and 100 million market members in around 80 countries worldwide. Travel companies and hotel managers also visit the website to read customer reviews. 3-4 star hotels were selected from the Vietnam Hotel Association website (<http://vietnamhotel.org.vn/Default.aspx>). The website listed all hotels with their star rates.

In 2017, the hotel industry in Vietnam experienced a growth rate that was 6th in the world. Data in the time range from 2017 to 2022 on TripAdvisor is vast. This study used Web Harvy to collect and store hotel review data from 2017 to 2022, including reviewer name, review date, content, and title.

Next, we preprocess the data with Python, resulting in 20,551 reviews, 2,268,646 total words, and a vocabulary size of 32,687. Figure 2 depicts the processed data.

The length of customer reviews from 3-5-star hotels in Vietnam, totaling 20,551 reviews, varies greatly, from 10,335 words to just one word.

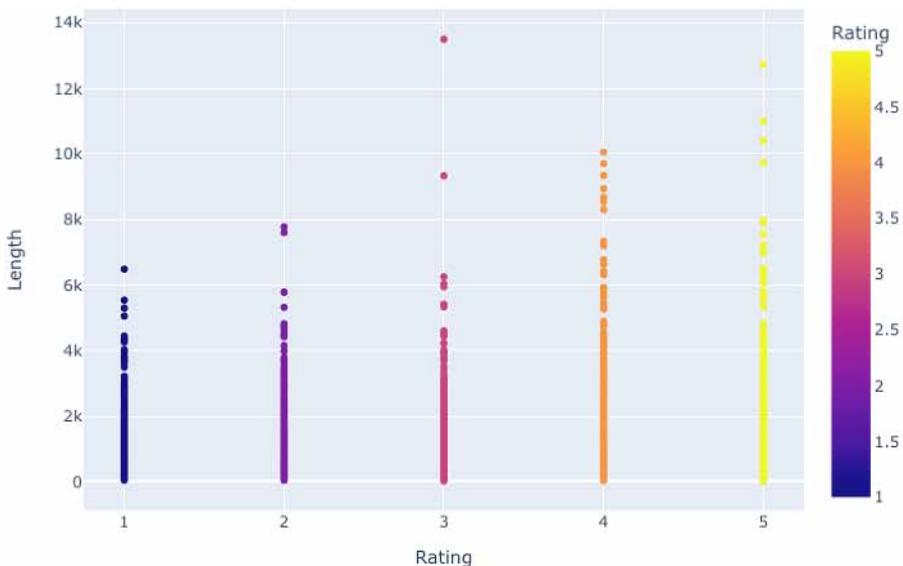
The study found that customers satisfied with hotel services tend to write longer sentences, indicating that they carefully considered the hotel's positive aspects. The figure below illustrates the relationship between customers who rated the hotel quality highly and review length. Similarly, Figure 3 shows that the number of positive review sentences is longer than the number of negative sentences. Negative customer reviews are written short and rate the service quality with a lower number of stars.

Figure 2. Description of data that were collected from the TripAdvisor site

```
                                text
0      My husband and I traveled here for business an...
1      We booked the full board package with 3 meals ...
2      Nice location and rooms are spacious with sea ...
3      Yes I would recommend Avani Quy Nhon, check by...
4      My family and I had stayed in Avani Quy Nhon i...
...
...
20546  We stayed for four nights recently and simply ...
20547  We stayed in the Crowne Plaza after a disaster...
20548  Expecting the worst after all the bad reviews ...
20549  When our tour guide said of our hotel - 'it's ...
20550  The hotel is ahead of its time but I liked the...

[20551 rows x 4 columns]
```

Figure 3. The relationship between review length and customer satisfaction



## 4.2 Assessing the Overall Satisfaction of Customers Staying at a Vietnamese Hotel

The scores for positive, negative, neutral, and compound sentiments of the text have been analyzed using the Vader library and are presented in Figure 4 with great detail. In which, negative scores are represented by Neg, neutral scores by Neu, and positive scores by Pos. The Compound score determines overall sentiment.

The authors' classification of online reviews includes three distinct types: negative for expressing dissatisfaction, positive for indicating satisfaction, and neutral for neither. Negative reviews serve as an indicator of unsatisfactory service, while positive reviews signify satisfactory service. Meanwhile, neutral reviews do not have explicit opinions. There are 19,504 reviews expressing satisfaction, 975 reviews are dissatisfied, and 71 neutral reviews. We calculate the rate according to the formula to measure the percentage of satisfied customers given in formula number (4) in Section 3.3. Satisfaction and dissatisfaction rates of customers after measuring are equivalent to: satisfied: 94.9%, dissatisfied: 4.75% and 0.35% are not clearly expressed attitudes. The measurement results are shown in Figure 5.

## 4.3 Measuring Customer Satisfaction Based on Service Aspects

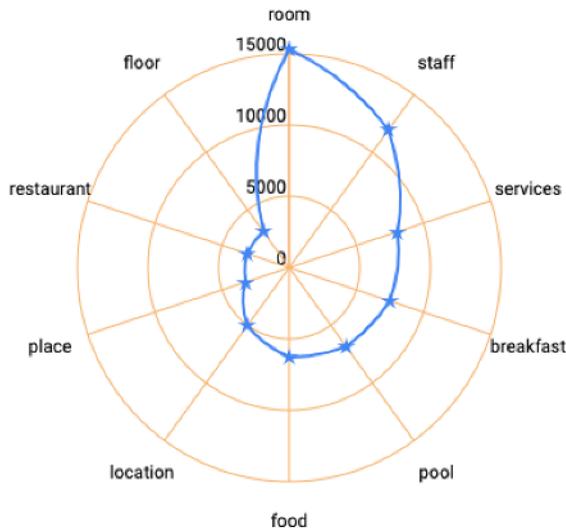
The hotel aspects that satisfy customers will be found in the positive reviews. As an overview is depicted in Figure 6, words frequently in positive reviews include hotel service aspects such as room, staff, food, service, place, etc. It is easy to see that customers are most satisfied in terms of room, staff, and buffet when they stay in 3-5-star hotels (Figure 6).

Figure 4. Scores of reviews

| text           | description_lengths | scores                                                          |
|----------------|---------------------|-----------------------------------------------------------------|
| My husband     | 172                 | {'neg': 0.033, 'neu': 0.77, 'pos': 0.197, 'compound': 0.9885}   |
| We booked      | 115                 | {'neg': 0.017, 'neu': 0.716, 'pos': 0.268, 'compound': 0.9895}  |
| Nice location  | 60                  | {'neg': 0.0, 'neu': 0.888, 'pos': 0.112, 'compound': 0.6055}    |
| Yes I would r  | 72                  | {'neg': 0.0, 'neu': 0.71, 'pos': 0.29, 'compound': 0.9769}      |
| My family      | 203                 | {'neg': 0.09, 'neu': 0.835, 'pos': 0.075, 'compound': -0.3153}  |
| We spent 8 c   | 70                  | {'neg': 0.0, 'neu': 0.881, 'pos': 0.119, 'compound': 0.8435}    |
| I got good     | 44                  | {'neg': 0.035, 'neu': 0.815, 'pos': 0.15, 'compound': 0.674}    |
| This is by     | 167                 | {'neg': 0.021, 'neu': 0.822, 'pos': 0.157, 'compound': 0.979}   |
| Our family st  | 138                 | {'neg': 0.117, 'neu': 0.803, 'pos': 0.081, 'compound': -0.6662} |
| A beautiful    | 108                 | {'neg': 0.0, 'neu': 0.779, 'pos': 0.221, 'compound': 0.9763}    |
| The hotel is i | 41                  | {'neg': 0.076, 'neu': 0.597, 'pos': 0.327, 'compound': 0.9304}  |
| The hotel is i | 41                  | {'neg': 0.076, 'neu': 0.597, 'pos': 0.327, 'compound': 0.9304}  |
| Unfortunatel   | 40                  | {'neg': 0.117, 'neu': 0.742, 'pos': 0.142, 'compound': 0.2926}  |
| This private r | 193                 | {'neg': 0.0, 'neu': 0.731, 'pos': 0.269, 'compound': 0.9967}    |
| This hidden    | 147                 | {'neg': 0.0, 'neu': 0.788, 'pos': 0.212, 'compound': 0.9889}    |
| What a         | 37                  | {'neg': 0.0, 'neu': 0.496, 'pos': 0.504, 'compound': 0.9808}    |
| Beautiful roc  | 125                 | {'neg': 0.0, 'neu': 0.718, 'pos': 0.282, 'compound': 0.9918}    |
| Slow           | 133                 | {'neg': 0.077, 'neu': 0.822, 'pos': 0.1, 'compound': 0.6697}    |
| I stayed with  | 71                  | {'neg': 0.0, 'neu': 0.784, 'pos': 0.216, 'compound': 0.9612}    |
| Celebrated     | 129                 | {'neg': 0.014, 'neu': 0.672, 'pos': 0.314, 'compound': 0.9942}  |
| Great service  | 38                  | {'neg': 0.061, 'neu': 0.543, 'pos': 0.396, 'compound': 0.9455}  |
| Have a confi   | 37                  | {'neg': 0.0, 'neu': 0.78, 'pos': 0.22, 'compound': 0.8622}      |
| Convenient     | 35                  | {'neg': 0.0, 'neu': 0.613, 'pos': 0.387, 'compound': 0.9538}    |
| I visited      | 110                 | {'neg': 0.025, 'neu': 0.76, 'pos': 0.215, 'compound': 0.968}    |
| I stayed for   | 114                 | {'neg': 0.026, 'neu': 0.759, 'pos': 0.215, 'compound': 0.9769}  |
| We had a       | 130                 | {'neg': 0.028, 'neu': 0.643, 'pos': 0.329, 'compound': 0.9943}  |
| I had a wond   | 75                  | {'neg': 0.041, 'neu': 0.736, 'pos': 0.223, 'compound': 0.9305}  |
| Excellent exc  | 49                  | {'neg': 0.0, 'neu': 0.564, 'pos': 0.436, 'compound': 0.9863}    |



Figure 7. Extracting the highest aspects



While Figure 7 shows only customer interest in a certain aspect but does not measure satisfaction, Figure 8 shows customer satisfaction with individual aspects of hotel service. The study used formula (6) in section 3.4 to measure satisfaction with each service aspect.

Figure 9 shows polarity words, also known as sentiment words, that indicate customer satisfaction in reviews. These words include positive adjectives like good, great, nice, excellent, and friendly, which are used frequently.

Next, the study can find the pairing of sentiment words and aspects. This pairing represents how often they occur together. Figure 10 below shows an example of how often the aspects “room,” “staff,” and “food” appear when paired with sentiment words. The size between the cells indicates more or less chance of co-occurrence.

Figure 8. The list of aspects with the ratio of customer satisfaction

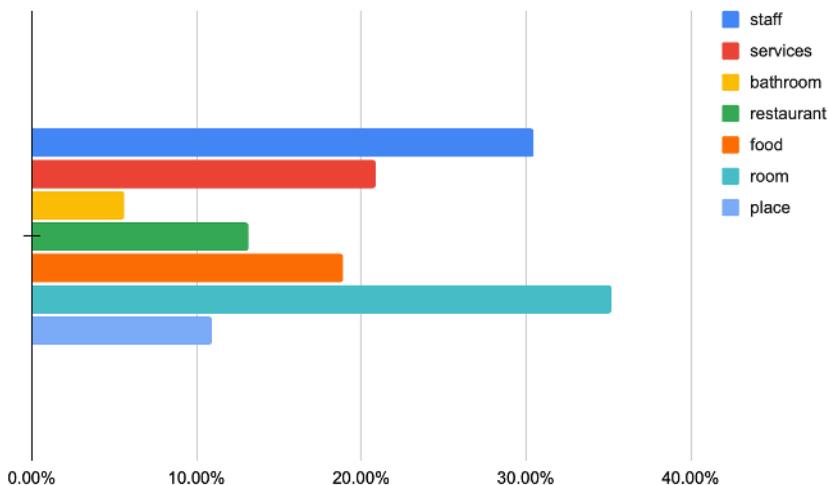


Figure 9. List of polarity words in positive reviews

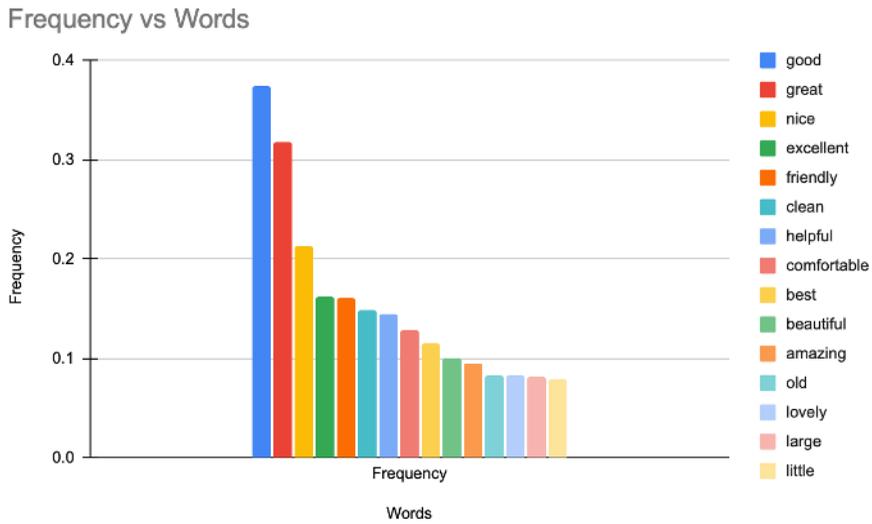


Figure 10. The combination of sentiment words and aspect



## 5. DISCUSSION AND CONCLUSION

### 5.1 Discussion

This study does not approach complex machine learning models, linear regression functions, or influencing factors like previous studies (Wu et al., 2022; Wu et al., 2023; Xu et al., 2022). The study only proposes two simple application tools, namely WebHarvy with convenient API and open source Python, but also solves three important problems in analyzing a large number of online reviews, including: (1). Automated data collection with WebHarvy tools; (2). Measure overall customer satisfaction; (3). Measure customer satisfaction in each aspect of the hotel. The main results of the study include:

1. The customer satisfaction rate with the service quality of Vietnamese hotels reaches more than 94% overall.
2. Although the overall level of satisfaction is very high, when analyzing the aspect detail level, the highest aspect “room” only achieved more than 35% satisfaction. This shows that there are negative reviews mixed in with positive sentences.
3. Combining sentiment words and aspects gives a complete picture of customer satisfaction. Because sentiment words will have many different levels of emotional expression, they will measure the scale of customer satisfaction states (For example, on a satisfaction scale from 1-5).

## **5.2 Limitations and Future Work**

The customer-centric approach is a modern business strategy that increases customer satisfaction and loyalty, leading to higher revenue and profits. It involves offering products and services based on customers’ needs and desires, considering their opinions and feelings to improve their shopping experience. This creates a positive company image, attracts customers to return, and builds trust and loyalty. One of the most significant disruptions in the hospitality industry has been the sharing economy, where private individuals coordinate transactions through online platforms. This trend has evolved into highly profitable business models in recent years.

Vietnam is a developing country in Southeast Asia and has become the region’s second most popular tourist destination. To boost the hospitality industry further, hotel managers should establish a more substantial online presence to improve global governance. In today’s world of interconnectedness, social listening across various interactive channels is crucial for understanding customer opinions. This market research model helps hotels evaluate customer satisfaction, assess their psychology and emotional views, and manage media crises affecting their reputation.

This study has proposed a simple but effective solution supporting 3-5 star Vietnamese hotels more efficiently instead of complex algorithms and tools. After COVID-19, Vietnam’s tourism is on track to accelerate again, and opportunities to develop the hotel industry in Vietnam have gradually improved. Using social network mining solutions to listen to customer feedback is necessary to understand customers and provide services that are more tailored to their needs. However, for the study, it was impossible to show a clear separation between pre-COVID-19 and post-COVID19 data to explore customer expectations for hotel services in more detail. In addition, it is necessary to analyze at a deeper level to measure unsatisfied customer service issues to have a plan to improve service quality.

In the future, there will be many open directions when applying data analysis techniques to measure and exploit customer views at a deeper level. Future work includes measuring customer complaints about the quality of hotel services in Vietnam and evaluating hotel quality based on analysis by combinations of words.

## **ACKNOWLEDGMENT**

We would like to express our gratitude to Thuongmai University for their support in conducting this research.

## **FUNDING AGENCY**

This research is funded by Vietnam Thuongmai University.

## **COMPETING INTERESTS**

All authors of this article declare there are no financial or non-financial competing interests.

## REFERENCES

- Adjei Peter, D., & Decui, L. (2022). Modeling customer satisfaction through online reviews: A FlowSort group decision model under probabilistic linguistic settings. *Expert Systems with Applications*, 195(116649), 116649. doi:10.1016/j.eswa.2022.116649
- Akhtar, N., Zubair, N., Kumar, A., & Ahmad, T. (2017). Aspect based sentiment oriented summarization of hotel reviews. *Procedia Computer Science*, 115, 563–571. doi:10.1016/j.procs.2017.09.115
- Alaei, A. R., Becken, S., & Stantic, B. (2019). Sentiment analysis in tourism: Capitalizing on big data. *Journal of Travel Research*, 58(2), 175–191. doi:10.1177/0047287517747753
- Alrawadieh, Z., & Law, R. (2019). Determinants of hotel guests' satisfaction from the perspective of online hotel reviewers. *International Journal of Culture, Tourism and Hospitality Research*, 13(1), 84–97. doi:10.1108/IJCTHR-08-2018-0104
- Barnes, S. J., Mattsson, J., Sørensen, F., & Jensen, J. F. (2020). Measuring employee-tourist encounter experience value: A big data analytics approach. *Expert Systems with Applications*, 154(113450), 113450. doi:10.1016/j.eswa.2020.113450
- Bian, Y., Ye, R., Zhang, J., & Yan, X. (2022). Customer preference identification from hotel online reviews: A neural network based fine-grained sentiment analysis. *Computers & Industrial Engineering*, 172(108648), 108648. doi:10.1016/j.cie.2022.108648
- Blomberg-Nygaard, A., & Anderson, C. K. United Nations World Tourism Organization (UNWTO). (2016). United nations world tourism organization study on online guest reviews and hotel classification systems: An integrated approach. *Service Science*, 8(2), 139–151. doi:10.1287/serv.2016.0139
- Camilleri, M. A., & Filieri, R. (2023). Customer satisfaction and loyalty with online consumer reviews: Factors affecting revisit intentions. *International Journal of Hospitality Management*, 114(103575), 103575. doi:10.1016/j.ijhm.2023.103575
- Chalupa, S., & Petricek, M. (2022). Understanding customer's online booking intentions using hotel big data analysis. *Journal of Vacation Marketing*, 135676672211221. doi:10.1177/13567667221122107
- Chen, M.-C., Hsiao, Y.-H., Chang, K.-C., & Lin, M.-K. (2019). Applying big data analytics to support Kansei engineering for hotel service development. *Data Technologies and Applications*, 53(1), 33–57. doi:10.1108/DTA-05-2018-0048
- D'Acunto, D., Filieri, R., & Amato, S. (2023). Hotels' environmentally-framed eWOM. The moderating role of environmental culture. *Tourism Management*, 98(104776), 104776. doi:10.1016/j.tourman.2023.104776
- Dong, J., Li, H., & Zhang, X. (2014). Classification of customer satisfaction attributes: An application of online hotel review analysis. In *IFIP Advances in Information and Communication Technology* (pp. 238–250). Springer Berlin Heidelberg.
- Breda, Z., Costa, R., Dinis, G., & Martins, A. A. (2020). ewow of guests regarding their hotel experience: Sentiment analysis of TripAdvisor reviews. In *Handbook of research on social media applications for the tourism and hospitality sector* (pp. 295-308). IGI Global.
- Godnov, U., & Redek, T. (2018). Good food, clean rooms and friendly staff: Implications of user-generated content for Slovenian skiing, sea and spa hotels' management. *Management*, 23(1), 29–57. doi:10.30924/mjcmi/2018.23.1.29
- Gómez-Suárez, M., & Veloso, M. (2022). Designing Facebook publications focused on hotel customer experience: How to improve brand attitude and booking intention. In *Brand, Label, and Product Intelligence* (pp. 247–258). Springer International Publishing.
- Leal, F., Malheiro, B., & Burguillos, J. C. (2019). Analysis and prediction of hotel ratings from crowdsourced data. *Wiley Interdisciplinary Reviews. Data Mining and Knowledge Discovery*, 9(2), e1296. doi:10.1002/widm.1296
- Li, H., Liu, Y., Tan, C. W., & Hu, F. (2020). Comprehending customer satisfaction with hotels: Data analysis of consumer-generated reviews. *International Journal of Contemporary Hospitality Management*, 32(5), 1713–1735. doi:10.1108/IJCHM-06-2019-0581

Li, Q., Li, S., Zhang, S., Hu, J., & Hu, J. (2019). A review of text corpus-based tourism big data mining. *Applied Sciences (Basel, Switzerland)*, 9(16), 3300. doi:10.3390/app9163300

Liu, Y., Teichert, T., Rossi, M., Li, H., & Hu, F. (2017). Big data for big insights: Investigating language-specific drivers of hotel satisfaction with 412,784 user-generated reviews. *Tourism Management*, 59, 554–563. doi:10.1016/j.tourman.2016.08.012

Luturlean, B. S., & Anggadwita, G. (2016). A framework for conceptualizing customer experiences management in the hotel industry. *Proceedings of the 3rd International Seminar and Conference on Learning Organization (Isclo-15)*. IEEE. doi:10.2991/iscl-15.2016.25

Mohamed, E. S. A. (2021). The impact of customer experience and relationship quality on corporate reputation in the hotel sector. *International Journal of Customer Relationship Marketing and Management*, 12(2), 53–79. doi:10.4018/IJCRMM.2021040104

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

Paulose, D., & Shakeel, A. (2022). Perceived experience, perceived value and customer satisfaction as antecedents to loyalty among hotel guests. *Journal of Quality Assurance in Hospitality & Tourism*, 23(2), 447–481. doi:10.1080/1528008X.2021.1884930

Rahimian, S., ShamiZanjani, M., Manian, A., & Esfidani, M. R. (2021). A framework of customer experience management for hotel industry. *International Journal of Contemporary Hospitality Management*, 33(5), 1413–1436. doi:10.1108/IJCHM-06-2020-0522

Roy, G. (2023). Travelers' online review on hotel performance – Analyzing facts with the Theory of Lodging and sentiment analysis. *International Journal of Hospitality Management*, 111(103459), 103459. doi:10.1016/j.ijhm.2023.103459

Sampetua Hariandja, E., & Vincent, F. (2022). Linking customer experience, satisfaction, and loyalty to brand power and performance in international hotels. *Innovative Marketing*, 18(3), 59–71. doi:10.21511/im.18(3).2022.06

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

Wu, J., Liu, C., Wu, Y., Cao, M., & Liu, Y. (2022). A novel hotel selection decision support model based on the online reviews from opinion leaders by best worst method. *International Journal of Computational Intelligence Systems*, 15(1), 19. doi:10.1007/s44196-022-00073-w

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

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

Zhao, Y., Xu, X., & Wang, M. (2019). Predicting overall customer satisfaction: Big data evidence from hotel online textual reviews. *International Journal of Hospitality Management*, 76, 111–121. doi:10.1016/j.ijhm.2018.03.017

*Ha Nguyen Thi Thu is a lecturer researcher from the Department of Greenwich Vietnam at FPT University. She received a Master's in IT at Guilin University of Electronic Technology in 2006 and now is a Doctor. Her major is in E-commerce and Data analysis. She has more than 20 international papers in this field. She is interested in AI, big data, and how to apply it to Business management.*

*Thao Phan Huong is a lecturer, researcher from Management Accounting Department at Thuongmai University. She received a Master's in Accounting at Thuongmai University in 2011 is a Ph. Doctor now. Her major is in Management Accounting, Sustainable Development. She has more than 18 international papers in this field. She is interested in Accounting, Strategic Management Accounting, and how to apply them to Business management.*

*Anh Le Thi Tram is a lecturer and researcher from Accounting - Auditing Department at Thuongmai University. She received a Master's in Accounting at Thuongmai University in 2010 and now is a Doctor. Her major is in Accounting and Auditing, Information Disclosure. She has more than 12 international papers in this field.*

*Thao Tran Viet is a lecturer, researcher from Department of Research Administration at Thuongmai University. His major is Economy Management, Business Administration. He received a Master of Business Administration degree at Hanoi National University in 2009 and is a Doctor now. He has more than 25 international papers in this field. He is interested in Business Administration, Economy Management, and how to apply them to business management and sustainable development.*