Asia Business & Service Innovation 2(1), 111-125. (2024)
https://doi.org//10.23405/ABSI-2024.1.1.06

Information available at the Journal of Asia Business and Service Innovation (http://www.absi.kr/)

Asia Business and Service Innovation

ISSN: 3092-0558 (Print) | eISSN: 3092-1112 (Online)

Exploring Online Customer Reviews of Luxury Hotels in Malaysia

Debasis Paul[1], Nasir Uddin[1] , Hyeon Jeong Choi [2]*

[1] *Graduate Student, Kyungsung University, Busan, Republic of Korea (debasishp@gmail.com, nasir114400@gmail.com)*

[2] *Professor, Department of Tourism, Major in Hotel Management, Baekseok Arts University, South Korea*

* Corresponding author: chj043@bu.ac.kr

Abstract

Online customer reviews can provide service companies with a better insight into customer choices. Understanding online customer feedback can help hotel companies develop an effective marketing strategy and differentiate the levels of service and experience they provide to the customers. This can increase the level of hotel customer satisfaction and positively the impact decision-making process. Therefore the purpose of this study is to analyze the content of online customer reviews to reveal the perception of the hotel customers through text mining and sentiment analysis. The study was based on six top 5-star hotels selected from Google Travel and located in Kuala Lumpur, Malaysia. To achieve the aims of the study more than 15,000 distinct data points of the sample were collected from SCTM 3. Moreover, the R programming language was used for deriving word matrix and word count to analyze the data, while UCINET was applied for semantic network analysis of top frequency words. In addition, CONCOR analysis was performed to divide keywords into four clusters: “Service”, “Scenario”, “Environment”, and “Destination” and visualized by drawing networks and nodes using NetDraw in UCINET. The results of the study provide important empirical evidence with implications for building effective marketing strategies for 5-star hotels in Kuala Lumpur.

Keywords: big data, customer engagement, hotel, text-mining, ranking, Sentimental status.

Received Mar 25, 2024 · Revised May 15, 2024 · Accepted Jun 12, 2024

1. Introduction

The tourism and hospitality industry of Malaysia is the largest and fastest-growing sector which comes up with the country's economic development and employment opportunities. The growth rate of the hospitality market is expected from $3486.77 billion in 2020 to $4132.5 billion in 2021 (HospitalityNet, 2021). Advances in technology are contributing to remarkable improvements and savings in the hospitality market. Malaysia is a glorious jumble of multi-cultural, vibrant cities, and wildlife-rich jungles. Kuala Lumpur is Malaysia's urban hot spots and culturally diverse capital city with gigantic shopping malls, iconic Petronas Tower, the nightlife of golden triangle, and extensively famous for food like 'otakotak' (steamed fish curry), Chinese congee (rice porridge), and grilled satay (Tourism Malaysia, 2021). It is an ideal destination in Southeast Asia for pictures que wildlife, scuba diving, spa and yoga center, national parks, and some rain forests as well. The capital city provides beach resorts, an extensive selection of five-star hotels, and guesthouses (Tourism Malaysia, 2021).

In the era of digitalization, customer reviews are playing one of the most important roles in the success of hotel companies. Through effective use of social media hotels can better engage with their customers, get their customers to stay and spend more, enhance the number of clientele, and positive word-of-mouth (Ban & Kim, 2019). Nowadays, customers can share their opinions and experiences on various social network platforms. These opinions shared online can influence the customer decision process (Bilro, Loureiro, & Guerreiro,2019). It is important to understand customer needs in order to identify key elements for customer satisfaction. Without identifying and properly understanding these key elements, it will be difficult to provide high levels of service and satisfy the customers. Social media platforms can help hotels understand the customers' needs and want from their feedback, upgrade their image, and improve customers satisfaction (Schuckert, Liu, & Law, 2015). Thus, there is increasing attention from researchers as well as industry practitioners regarding online customer reviews (Cantallops & Salvi, 2014).

Text mining and sentimental analysis can be applied to better understand customer needs and desires. Through using text mining, we can gather a large amount of unstructured data which

can be used for deriving meaningful information (Mariani & Borghi, 2020). Previous studies have applied text mining analysis to understand the food hygiene of cruise lines (Shuting, Kang, & Kim, 2018) and the seat comfort experience of airline passengers (Ban & Kim, 2019). Similarly, sentiment analysis can be used to understand the overall emotions, attitudes, and opinions of customer service. It was applied in the study of González-Rodríguez et al., (2016) to explore post-visit and pre-visit tourist destination images. Regrettably, there is still a lack of studies that applied text mining and sentimental analysis to understand the customer experience in the hotel context. This type of study could be of great value to scholars as well as hotel managers. Therefore, the purpose of this study is to analyze the content of online customer reviews of hotels to reveal the perception of the hotel customers through text mining and sentiment analysis. The main contribution of this study is providing better insight into hotel customer experience focusing on previously not explored 5-star hotels in Kuala Lumpur, Malaysia. The results of the study can provide important empirical evidence as well as managerial locations for building efficient marketing strategies necessary for survival and success in today's highly competitive environment. In order to accomplish the purpose of the study.

2. Theoretical Background

2.1. Text Mining

Text mining is the procedure of modifying the shapeless text into a structured layout to recognize significant design and the latest perception. By appealing to the progressive problem-solving activities organizations are capable of research and exploring secret connections surrounding their unstructured data. To explore customers' perceptions from online textual reviews, researchers most often use frequency analysis (Xiang et al, 2015). Nevertheless, the researchers need to study this method of online reviews word by word for better understanding, which may not be favorable in the case of the big data field. For instance, Xu et al., (2017) grouped hotel review text information into groups to find more comprehensive information about the hotel to improve service quality.

2.2. Sentiment Status Analysis

Hundreds of thousands of people rely on user sentiment reviews on the internet. In April 2013, 90 percent of client decisions were based on online reviews (Peng et.al., 2014). The basic purpose of sentiment analysis is to assess the scores of sentiments and analyze there views (Feldman, 2013). Some writers focus on text aggregation to predict sentiment (Zhang et al., 2009), creating tools that automatically analyze posts or reviews (González-Rodríguez et al., 2016), or even mining sentiments, opinions, and emotions (Liu, 2015). Zhao et al., (2016) underlined that each type of object (e.g., hotel) has different types of determinants of customer satisfaction and dissatisfaction, while other authors focus on measuring customers' attitudes towards brands (Ismail et al., 2016). Customer engagement seems to be recommended positively than conflicting. Their opinions and recommendation are always subjective. Some customers might not be satisfied by some features, but they give positive feedback happily. Their reviews most of the time depend on their cognitive thinking that means customers are more concerned about their reviews to others. This shows that online customer reviews, having a social media function, also have significant social commerce value (Lee et al., 2016). The operation methods are different in terms of every individual hotel. Offering quality services like room rates, discount packages, and others, all the things affected customers' perception and satisfaction. This research offers the most valuable effect on the literature of online tourist reviews and its method convenient for marketers.

2.3. Big Data Analysis

In the e-commerce and big data era, customer online reviews of hotels have enormous business value. Big Data refers to data sets that are too massive, complicated, or diversified for typical data processing applications to acquire, manage, and analyze (Shuting, Kang, &Kim, 2018). The explosive rise of social media and user-generated material on the Internet has sparked the creation of "big data analytics" to better understand and solve real-world issues. The technical side of online textual reviews, specifically the linguistic features of

online textual reviews is still generally under explored. This study uses the technical aspects of online textual reviews and customers' activity in the review community top redict total customer happiness using a sample of 15,200 reviews from SCTM3. The data reveal numerous elements of visitor experience with varied weights and more significantly semantic compositions that are innovative and meaningful. The link between customers' happiness and experience appears to be substantial, implying that these two aspects of consumer behavior are inextricably linked.

3. Research Method

3.1. Data Collection

In this study, the qualitative process was applied for the in-depth understanding of customer perception and measuring the quality of service. To achieve the aim of the study customer reviews of the top 5-star hotels in Kuala Lumpur, Malaysia had to be collected. Selection of the most famous hotels in the country was done based on the global travel information websites – Trip advisor and Google Travel. After six of the most famous 5-star hotels were selected from Google Travel, Smart Crawling and Text Mining (SCTM 3) platform created by Wellness and Tourism Big Data Research institute at Kyungsung University were used to collect the raw data on those hotels. Specifically from the SCTM,3 we collected the raw data of the reviewer's ID, review date, rating, review date, and comments for 6 years. Frequency was calculated by using text mining, and keywords were selected from the data sample collected on six hotels. A total of149024 data points were initially collected and 15200 customer reviews remained after unnecessary or not readable data points (like translated by Google) were deleted. The sample of hotels included in the study is presented in Table 1 below.

Table 1. Names of six hotels in Kuala Lumpur

Rank	Hotel names	No. of review
1	Traders Hotel Kuala Lumpur	4991
2	Hilton Kuala Lumpur	2477
3	Le-Meridian Kuala Lumpur	2304

4	Sunway Putra Hotel, Kuala Lumpur	2297
5	Grand Hyatt Kuala Lumpur	1633
6	Impian a KLCC Hotel, Kuala Lumpur	1553

3.2. Data Analysis

R programming was used for word matrix and word count to analyze the data. Thereafter we deleted some redundant data according to the frequency ranking of the arising words by conducting the search for the analysis data. Moreover, understanding the customer satisfaction rating score was used. The overall score was used as a dependable variable which is considered as the main output variable. The semantic network analysis of top frequency words was accomplished by UCINET 6.0 packaged with the visualization tool – NetDraw used to understand the connection structure and characteristics between words. By following this semantic network analysis was performed focusing on degree certainty and eigenvector, which are indicators quantified based on the certainty concept arrangement and measurement method. Finally, CONCOR analysis (Convergence of iterated Correlation) was conducted to obtain the subgroups of these words to understand these interwoven cor-relationships with each other and figure out the facets that customers are interested in. Data analysis process is shown in the Figure 1 below.

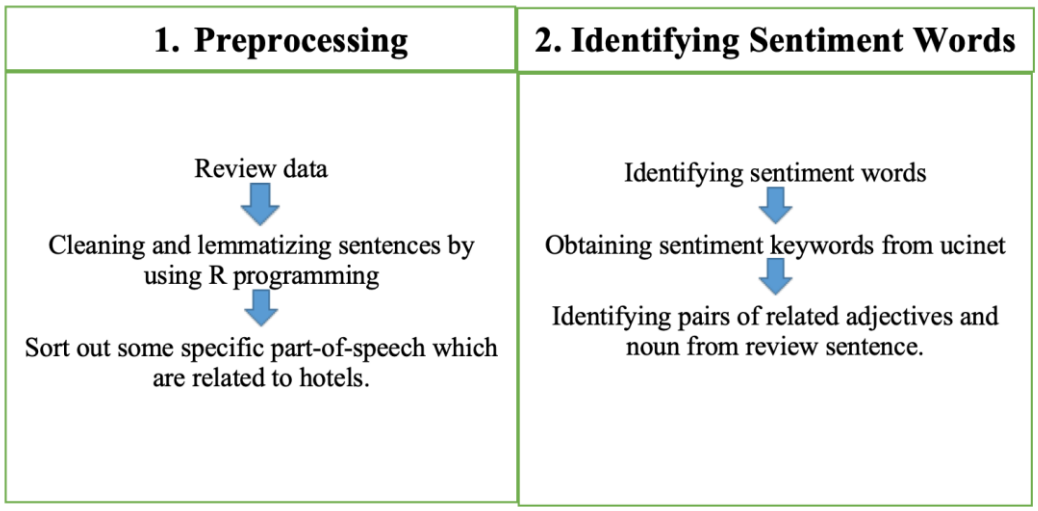

Figure1. Data analysis

4. Results

4.1. Frequency Analysis

As a result of 49024 data points were collected and a total of 50 words for data analysis with high frequency were extracted. It should be noted that someof repeated, low frequency, and unnecessary words were deleted. Table 2 shows most frequent 50 words in the online customer reviews and their frequency percentage.

Table 2. Top 50 frequent words in customer reviews

Most use frequent words in the online customer review							
Rank	Word	Frequency	%	Rank	Word	Frequency	%
1	hotel	5202	10.0	26	like	570	1.1
2	good	3373	6.5	27	trip	565	1.1
3	room	3271	6.3	28	helpful	564	1.1
4	great	2726	5.2	29	comfortable	562	1.1
5	nice	2652	5.1	30	mall	561	1.1
6	staff	2227	4.3	31	star	556	1.1
7	service	2052	3.9	32	sentral	545	1.0
8	stay	1693	3.2	33	towers	536	1.0
9	food	1634	3.1	34	experience	517	1.0
10	location	1590	3.0	35	get	496	1.0
11	place	1512	2.9	36	love	483	0.9
12	view	1427	2.7	37	spacious	480	0.9
13	breakfast	1354	2.6	38	lounge	469	0.9
14	friendly	1133	2.2	39	floor	457	0.9
15	excellent	1015	1.9	40	shopping	440	0.8
16	clean	970	1.9	41	city	438	0.8
17	best	921	1.8	42	stayed	434	0.8
18	pool	879	1.7	43	located	415	0.8
19	one	765	1.5	44	Hilton	411	0.8
20	just	736	1.4	45	station	401	0.8
21	buffet	707	1.4	46	awesome	400	0.8
22	well	698	1.3	47	Kuala	395	0.8
23	time	618	1.2	48	easy	386	0.7
24	amazing	598	1.1	49	convenient	382	0.7
25	really	581	1.1	50	club	376	0.7

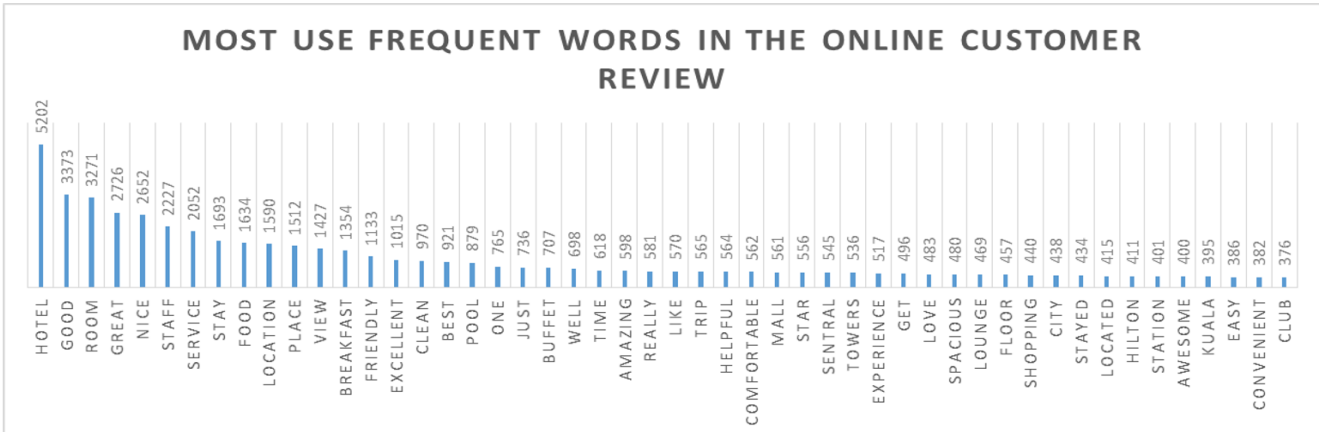

Figure 2. Statistics of top 50 frequent words

4.2. Analysis of Centrality (Freeman's Degree and Eigenvector)

There are two indicators in the centrality analysis used in this study. One is Freeman's degree centrality, which is an index that determines how connected one node is to other nodes in the network (Tao &Kim, 2019). Eigenvector centrality, often known as chain centrality, is the other and It is a valuable index for locating the network's most influential central node (Ban & Kim, 2019; Lee, Lim, & Kim, 2015). The prominence of words is a key indicator of the nodes' link (connection and influence). Table3 summarizes the relationship between top frequency words and their centrality. As it can be seen words such as “Hotel”, “Room”, “Good”, “Staff”, “Great” are characterized by a high frequency and a high centrality value. This indicated that certain words were frequently used by customers in online evaluations and that the link and influence of these words on other nodes was equally substantial. In other words, these words have a lot of weight in the semantic network of high-frequency words. While a term like "Location" has a high frequency (10), Freeman's centrality and Eigenvector centrality are both 33 and33, respectively. It implies that customers frequently used this word to discuss their experiences or opinions about restaurants establishments. The link of these terms to other nodes, on the other hand, is not particularly close. Other words like "Awesome," "City," "Excellent," "Person," and others followed the same trend as "Location." The word "Star," on the other hand, has a frequency rank of 31, with Freeman's degree centrality of 5 and Eigenvector centrality of 9. From the point of view of customers, a word has a significant

relationship and impact on other terms. The centrality rating is obviously higher than the frequency rank. Figure 3 summarizes the keyword visualization of network analysis.

Table 3. Comparison of frequency, the centrality of words extracted from customer review

Word	Frequency		Freeman's degree centrality		Eigenvector centrality	
	Frequency	Rank	Coefficient	Rank	Coefficient	Rank
hotel	5202	1	35.2	10	0.38	10
good	3373	2	23.7	19	0.29	19
room	3271	3	25.7	2	0.29	2
great	2726	4	21.3	20	0.26	20
nice	2652	5	17.8	29	0.23	29
staff	2227	6	22.4	4	0.26	4
service	2052	7	18	7	0.22	7
stay	1693	8	16.5	8	0.19	8
food	1634	9	14.4	12	0.17	12
location	1590	10	16.3	33	0.2	33
place	1512	11	10.5	15	0.13	15
view	1427	12	13.9	38	0.16	38
breakfast	1354	13	16.4	18	0.19	18
friendly	1133	14	12.9	27	0.16	27
excellent	1015	15	9	36	0.11	36
clean	970	16	11	11	0.13	11
best	921	17	7.6	28	0.09	28
pool	879	18	10.6	41	0.12	41
one	765	19	8.4	1	0.1	1
just	736	20	8.4	6	0.1	6
buffet	707	21	6.9	40	0.08	40
well	698	22	8.3	37	0.1	37
time	618	23	7	16	0.08	16
amazing	598	24	5.9	24	0.07	24
really	581	25	6.6	22	0.08	22
like	570	26	6.3	14	0.07	14
trip	565	27	2.1	23	0.02	23
helpful	564	28	7.1	21	0.09	21
comfortable	562	29	6.4	30	0.08	30
mall	561	30	6.3	49	0.07	49

star	556	31	4.5	5	0.06	5
sentral	545	32	5.9	50	0.06	50
towers	536	33	6.3	42	0.07	42
experience	517	34	4.9	26	0.06	26
get	496	35	6.2	13	0.07	13
love	483	36	4.4	39	0.05	39
spacious	480	37	6.4	3	0.07	3
lounge	469	38	5.8	38	0.06	38
floor	457	39	5.6	31	0.06	31
shopping	440	40	5.3	43	0.06	43
city	438	41	5.1	35	0.06	35
stayed	434	42	5.4	44	0.06	44
located	415	43	4.6	46	0.05	46
hilton	411	44	3.7	9	0.04	9
station	401	45	5.1	45	0.05	45
awesome	400	46	3.2	34	0.04	34
kuala	395	47	4	17	0.04	17
easy	386	48	4.4	25	0.05	25
convenient	382	49	4	32	0.05	32
club	376	50	4.1	47	0.04	47

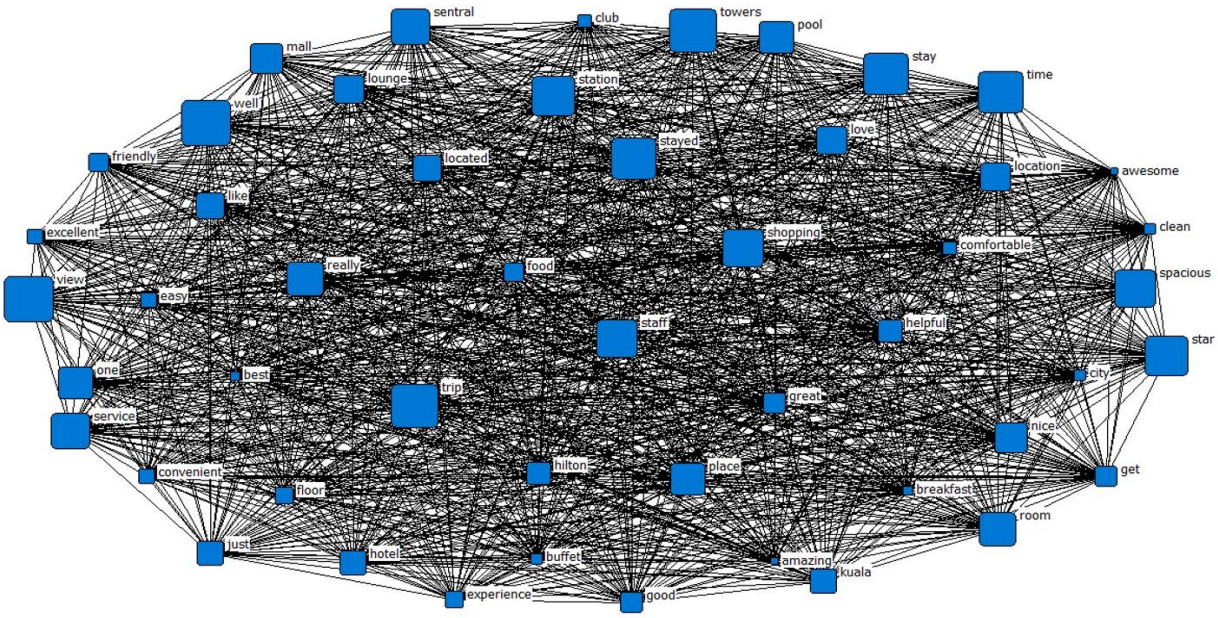

Figure 3. Keyword visualization of network analysis

The result of the CONCOR analysis was illustrated and identified which words belong to each group, the words grouped in each cluster and the noted words were listed in Figures 4 and 5.

Figure 4. Word cloud

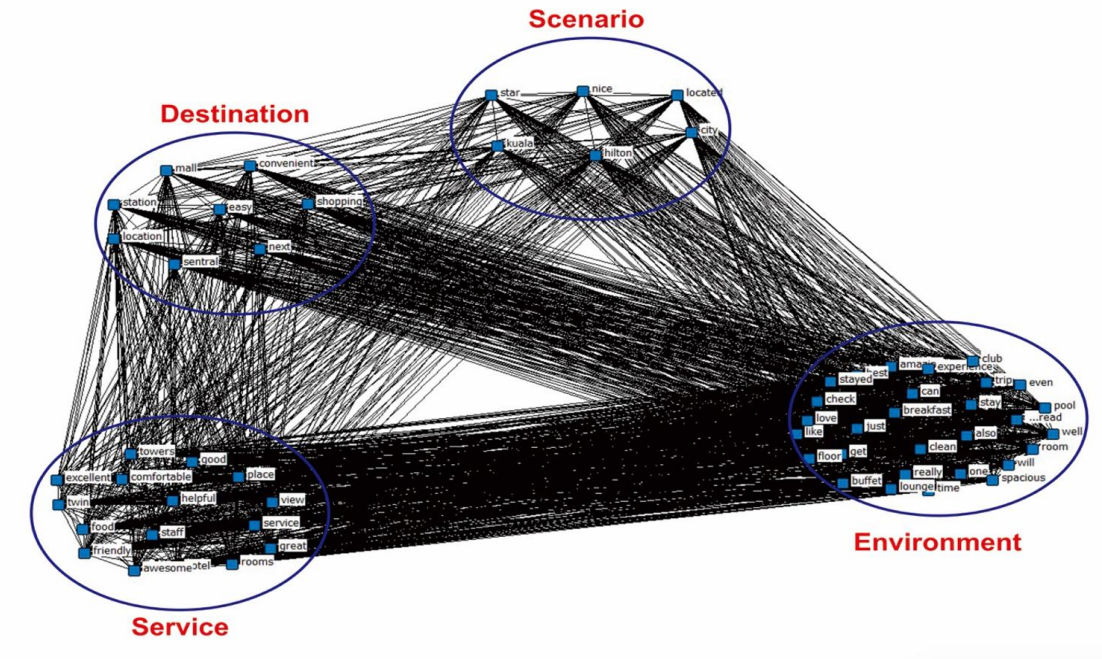

Figure 5. Visualization of convergence of iterated correlation (CONCOR) analysis.

Table 4. Results of CONCOR analysis

	Extracted Words	Significant Words
Service	clean, rooms, hotel, service, breakfast, location, staff, comfortable, food, towers, friendly, good, view, excellent, view, place, helpful, twin, great.	clean, rooms, hotel, service, food, breakfast, friendly, excellent
Scenario	nice, star, City, Kuala, Hilton, located.	Nice, city, kula, Hilton
Environment	Amazing, best, stayed, check, love, clean, love, like, buffet, well, room, floor, check, get, just, breakfast, can, experience, club, trip, pool, also, really, lounge, time, spacious, well, room, read, stay, best, trip.	amazing, best, love, like, spacious,
Destination	mall, location, station, sentral, easy, shopping, convenient.	mall, location, shopping, location, station, sentral

There are four clusters which include: “Service”, “Scenario”, “Environment”, and “Destination”. The first cluster is “Service” with terms such as clean, rooms, hotel, service, breakfast, location, staff, comfortable, food, towers, friendly, good, view, excellent, view, place, helpful, twin, great. Whereas clean, rooms, hotel, service, breakfast, friendly, excellent are the most significant words which indicate the quality of service. The second cluster is a “Scenario” that describes the surrounding view of the place. The third one is an “Environment” with the most important keywords such as amazing, best, clean, love, like, spacious. These words are related to the hospitality of the restaurant. And the final cluster is “Destination” which consists of a mall, location, station, sentral, easy, shopping, convenient. In this group mall, location, shopping, convenient are the highlighted frequent words that mean those places are popular. It is important to note that all words were extracted from CONCOR analysis and are related to customer experience.

5. Discussion and Conclusion

The purpose of this study was to analyze the content of online customer reviews to reveal the perception of the hotel customers through text mining and sentiment analysis. The sample of the study was based on six luxurious 5-star hotels in Kuala Lumpur, Malaysia. To achieve the aims of the study keywords were derived through text mining. Then frequency analysis was performed to identify the most frequent words. Based on this frequency analysis, the top

50frequent words were extracted and the degree and eigenvector centrality analysis was performed to understand the relationship between keywords. At last, the CONCOR analysis was adopted to generate four categories, namely, “Service”, “Scenario”, “Environment”, and “Destination”. Moreover, they were visualized by drawing networks and nodes using NetDraw in UCINET 6.0. Moreover, it was found that hotel customers importantly perceive the quality of service, appearance, communication, staff behavior, and environment. Thus, 5 star-hotel companies in Kuala Lumpur should consider these key factors to generate customer satisfaction and create higher customer ratings. Moreover, these key factors can be used to examine customer satisfaction and even test theoretical models in order to gain a better understanding of the hotel customers. This research is an initial attempt to provide an in-depth understanding of hotel customers’ perceptions in the Malaysian hotel context. Industry practitioners and managers can use the findings of this study in order to build effective marketing strategies and differentiate their hotels from competitors.

This study is not free of limitations. The first limitation is that the study sample was collected using SCTM 3 only from the Google Travel platform. For the collection of the broader and more diverse sample, other platforms such as Trip Advisor, Booking, Yelp.com, or Expedia should be used in future studies. Second, the study was based on a sample of 5-star hotels in Kuala Lumpur, Malaysia. Hence, the results of the studies based on the 5-star hotels in other countries, cities, or categories can significantly differ. Third, the collected text was analyzed based on the frequency of each word. Future studies can emphasize differentiating the role of consumers' expectations and perceptions of hotel services that also depend on the customer’s negative feedback. While most of the studies focus on customers’ positive feedback, negative comments should be also examined together in order to understand better customer opinions in future studies.

References

Ban, H. J., & Kim, H. S.(2019). Semantic network analysis of hotel packages through the big data. Culinary Science & Hospitality Research, 25(2), 110-119. https://doi.org/10.20878/cshr.2019.26.2.014

Bilro, R. G., Loureiro, S. M. C., & Guerreiro, J.(2019). Exploring online customer engagement with hospitality products and its relationship with involvement, emotional states, experience and brand advocacy. Journal of Hospitality Marketing & Management, 28(2), 147-171. https://doi.org/10.1080/19368623.2018.1506375

Cantallops, A. S., & Salvi, F.(2014). New consumer behavior: A review of research on eWOM and hotels. International Journal of Hospitality Management, 36, 41-51. https://doi.org/10.1016/j.ijhm.2013.08.007

Feldman, R.(2013). Techniques and applications for sentiment analysis. Communications of the ACM, 56(4), 82-89. https://doi.org/10.5120/ijca2016912136

González-Rodríguez, M. R.,Martínez-Torres, R., & Toral, S.(2016). Post-visit and pre-visit tourist destination image through eWOM sentiment analysis and perceived helpfulness. International Journal of Contemporary Hospitality Management, 28(11), 2609-2627. https://doi.org/10.1108/ijchm-02-2015-0057

HospitalityNet(2021). Accessed October 19, 2021. https://org/www.hospitality net.org/news/4107041.html

Ismail, M. N. I., Hanafiah, M. H., Aminuddin, N., & Mustafa, N.(2016). Community-based home stay service quality, visitor satisfaction, and behavioral intention. Procedia-Social and Behavioral Sciences, 222, 398-405. https://doi.org/10.1016/j.sbspro.2016.05.192

Lee, C., Lim, S., & Kim, H.(2015). Korean semantic role labeling using structured SVM. Journal of KIISE, 42(2), 220-226. https://doi.org/10.5626/jok.2015.42.2.220

Lee, Y. K., Kim, S. Y., Chung, N., Ahn,K., & Lee, J. W.(2016). When social media met commerce: a model of perceived customer value in group-buying. Journal of Services Marketing, 30(4), 398-410.https://doi.org/10.1108/jsm-04-2014-0129

Liu, B.(2015). Sentiment Analysis: mining sentiments, opinions, and emotions. Cambridge University Press.

Mariani, M., & Borghi, M.(2020). Environmental discourse in hotel online reviews: a big data analysis. Journal of Sustainable Tourism, 29(5), 829-848. https://doi.org/10.1080/09669582.2020.1858303

Peng, L., Cui, G., Zhuang, M., & Li, C.(2014). What do seller manipulations of online product reviews mean to consumers? (HKIBS Working Paper Series 070-1314). Retrieved from Lingnan University website:http://commons.ln.edu.hk/hkibswp/70

Schuckert, M., Liu, X., & Law, R.(2015). A segmentation of online reviews by language groups: How English and non-English speakers rate hotels differently. International Journal of Hospitality Management, 48, 143-149. https://doi.org/10.1016/j.ijhm.2014.12.007

Shuting, T., Kang, B., & Kim, H. S.(2018). Understanding the food hygiene of cruise through the big data analytics using the web crawling and text mining. Culinary science and hospitality research, 24(2), 34-43. https://doi.org/10.20878/cshr.2018.24.2.005

Tao, S., & Kim, H. S.(2019). Cruising in Asia: What can we dig from online cruiser reviews to understand their experience and satisfaction. Asia Pacific Journal of Tourism Research, 24(6), 514-528. https://doi.org/10.1080/10941665.2019.1591473

Tourism Malaysia (2021). Accessed October 30,2021. https://www.tourism.gov.my

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

Xu, X., Wang, X., Li, Y., & Haghighi, M.(2017). Business intelligence in online customer textual reviews: Understanding consumer perceptions and influential factors. International Journal of information management, 37(6), 673-683. https://doi.org/10.1016/j.ijinfomgt.2017.06.004

Zhang, C., Zeng, D., Li, J., Wang, F.Y., & Zuo, W.(2009). Sentiment analysis of Chinese documents: From sentence to document level. Journal of the American Society for Information Science and Technology, 60(12), 2474-2487. https://doi.org/10.1002/asi.21206

Zhao, Y., Xu, Li., & Wang, M.(2016). Predicting overall customer satisfaction: Big data evidence from hotel online textual reviews. International Journal of Hospitality Management, 76(A),111-121. https://doi.org/10.1016/j.ijhm.2018.03.017.