Determination of Customer Satisfaction by Text Mining: Case of Cappadocia Hotels

Ozan ÇATIR [iD] [a] **Kenan SAKALLI** [iD] [b] **Hasan ARMUTLU** [iD] [c] **Şükran ARMUTLU** [iD] [d]

[a] Usak University, Ulubey Vocational School, Usak, Turkey. ozan.catir@usak.edu.tr
[b] Usak University, Ulubey Vocational School, Usak, Turkey. kenan.sakalli@usak.edu.tr
[c] Usak University, Karahanlı Vocational School, Usak, Turkey. hasan.armutlu@usak.edu.tr
[d] Usak University, Usak, Turkey. sukranarmutlu@gmail.com

ARTICLE INFO	ABSTRACT
Keywords: Customer satisfaction Text mining Cappadocia Hotels Turkey Big Data Received 2 October 2019 Revised 29 February 2020 Accepted 12 March 2020 **Article Classification:** Research Article	**Purpose** – The aim of this study is to analyze the comments of travel web sites on 28 hotels that have 4 or 5 stars, operating in Cappadocia region by text mining and to determine the satisfaction level of customers from hotels. For this purpose, text mining methods were used and 10 263 comments were downloaded and analyzed. **Design/methodology/approach** – The research focuses on 5 and 4-star hotel reviews of the Cappadocia region in the English language version. For this study, 10263 comments were obtained from 28 different hotels. The excel file created in the context of data acquisition was added to the RapidMiner data mining program. Then, a model was created for the individual separation of the words mentioned in the headlines and comments. The aim is to make the word frequencies by determining the words in the comments as attributes, and also to make them suitable for the association analysis. **Findings** – It was determined that the tourists made comments on the room experience, staff and hotel dimensions. They have stated that the rooms were great, clean, gorgeous, amazing, unique and spacious. The hotel staff was found to be helpful and friendly. It is concluded that hotels are generally clean, beautiful and unique. It can be said that there is a high level of satisfaction from the hotels. **Discussion** – The fact that the data obtained by the method of text mining has facilitated generalizations. This method can be defined as the most common factors that tourists perceive when expressing their opinions and sharing their experiences online. In this respect, they have directly received the service quality, physical condition, etc., they expressed their comments about many factors as easily as they felt. It is thought that providing online environments with such facilities is important in obtaining reliable and valid information.

INTRODUCTION

With the rapid development of the Internet, an interactive interaction between businesses and customers has increased. Customers perform the services they have received online on their websites, and hotel managers are trying to manage this situation by reading these comments and identifying the satisfaction and dissatisfaction of the customers. Especially in the accommodation and travel industry, online reviews have had significant implications for tourists' purchasing decisions (Gao, Hu, and Bose, 2017; Pantano, Priporas, and Stylos, 2017).

The data that emerged as a result of sharing the customer experience that emerged with the use of the service provided on the online platforms have become important for the development of these services (Ghose and Ipeirotis, 2011; Chau and Xu, 2012). The comments shared by thousands of people online have become a great source of data especially for the tourism and travel sector with social awareness (Surowiecki, 2005). Online data has turned into a dimension that everyone can easily access it from anywhere and facilitates and influences their purchasing decision as customers express their criticism when they receive service (Ho-Dac, Carson, and Moore, 2013; Clemons, Gao and Hitt, 2006). Online data facilitates the planning of tourism business, marketing strategies, in short, getting the business success. In terms of marketing, it has a great effect on consumers' perceptions and behaviors and positive or negative effects of their choices. The analysis and interpretation of the data generated by the online platform contribute to the expected level of experience of the customer with the development of future strategic plans of the tourism enterprises, the solution of their

problems and the development of products and services. Although the study is not comprehensive, it is a representative study that can produce effective results in the region.

Manual text analysis has been performed in many studies for online reviews. This type of researches usually require intensive work and expertise of the researchers and also the number of samples is limited too. At the same time, it is difficult to prevent subjective prejudices (Mankad, Han, Goh and Gavirneni, 2016; Levy et al., 2013; Zhou, Ye, Pearce, and Wu, 2014).

In line with the online reviews, the studies carried out in the tourism sector so far have been mostly carried out on hotel sales (Xie et al. 2011, Ye et al. 2009, Ye Law, Li and Li., 2011) through traditional econometric and statistical analyses. In this research, unlike the previous studies, to reveal the dimensions of customer satisfaction, natural language processing, and advanced software and mathematical model of text mining were used. An important contribution of this study is to analyze the big data online by using text mining methods to reveal the dimensions of customer satisfaction. The results of this analysis are the customer's real perception of the dimensions of customer satisfaction. Revealing the customers' perception by their satisfaction dimensions, the real perception of customers can be practically used by hotel owners and managers.

The aim of this study is to analyze the comments of travel web sites on 28 hotels that have 4 or 5 stars, operating in Cappadocia region by text mining and to determine the satisfaction level of customers from hotels. For this purpose, text mining methods were used and 10 263 comments were downloaded and analyzed.

In the research, firstly, the literature study has been mentioned and then methods used including sampling and text mining techniques were explained. The results were discussed and then the results were presented. Finally, the article was completed with a discussion of limitations and recommendations for research and implementation.

Literature Review

The first method used in online studies to reveal customer satisfaction is the manual coding method (Manickas and Shea, 1997; Levy et al., 2013; Zhou et al., 2014). Another method is the word frequency analysis of computer-based text analysis methods (Lee and Hu, 2005, Berezina et al., 2016; Li et al., 2013). The automated text analysis method is also a powerful model that analyzes topics in satisfaction research (Xu and Li, 2016; Büschken and Allenby, 2016; Mankad et al., 2016; Guo et al., 2017). In the field of tourism, in order to examine the customer experiences of workers in hotel products and services, frequency analysis (Xiang et al., 2015), content analysis (Li et al., 2013), text-link analysis (Berezina et al., 2016) and hidden meaning analysis (Xu and Li, 2016), including text mining techniques are used.

Nowadays, while tourists are choosing destinations, they have difficulty in making decisions among many alternatives. With the development of the Internet, tourists are communicating their experiences on online platforms that enable them to express themselves in the best, fastest and most accessible way. Thus, the information that tourists share about their experiences on these online platforms becomes common to use (Chung and Koo, 2015). The information shared on these platforms creates a positive effect because of the richness and reliability of the tourists who have written the experience in the decision making of tourists, which reflects their true feelings (O'Connor, 2008).

The emergence of new sources of information on online platforms, the increase in areas where tourists can write their experiences and interest in studies in this area have increased (Leung et al., 2013). There are studies on customer satisfaction (Levy et al. 2013; Lu and Stepchenkova 2012), about customer perceptions and attitudes (Vermeulen and Seegers, 2009) and about likes (Leung et al. 2013) on online platforms. Online reviews help hotel businesses predict about hotel performance (Blal and Sturman, 2014), customer repetition rates (Ludwig et al., 2013) and competitiveness power (Gao et al., 2018).

Recent studies have examined the relationships between text mining and online customer assessments of hotel businesses. Geetha et al. (2017) examined the relationship between emotionality and customer scores, on the other hand, He et al. (2017) and Qu et al. (2008) found a high correlation between the general evaluation score of the customers for the hotel and the content and title sensitivity score obtained through text mining. Kim, et al. (2015) stated that customer scores are the most critical factor in determining hotel performance.

Tian et al. (2016), in addition to the existing studies, argued that reviewing the interpretations made by customers who have different language and culture can achieve different results. Assessing cultural differences, Zhang et al. (2016) examined the difference between Chinese customers and Western customers and found that Chinese customers had a less emotional response as a result of text reviews and ratings. Xiang et al. (2017) stated that the differences in the technical capacities of the websites that are included in the online platform cause different results in the level of information quality. Zhang et al. (2010) examined the demographic characteristics of the customers who commented on these pages and tried to understand the power of persuasion of the comments on the hotel businesses and managers and customers who made the purchase decision. The study contributes to this research flow by determining the dimensions of text mining and customer satisfaction in tourism enterprises.

REALIZATION OF THE APPLICATION

DATA COLLECTION:

In this study, data from the web site which is one of the most frequently used websites worldwide, were used. The research focuses on 5 and 4-star hotel reviews of the Cappadocia region in the English language version. For this study, 10263 comments were obtained from 28 different hotels. The names, the number of stars, and the number of comments of each hotel where the comments were downloaded are given in Table 1.

Table 1. *The hotels that comments have been downloaded and number of comments*

Hotel Name	Comment Number	Hotel Star
Aydinli_Cave_Hotel	1003	4
Kayakapi_Premium_Caves_Cappadocia	917	5
Gamirasu_Cave_Hotel	812	4
Sultan_Cave_Suites	785	4
Cappadocia_Cave_Suites	742	5
MDC_Hotel	695	4
Argos_in_Cappadocia	550	4
Museum_Hotel	520	5
Goreme_Inn_Hotel	517	4
Kismet_Cave_House	390	4
Hezen_Cave_Hotel	379	4
Fresco_Cave_Suites_Cappadocia	335	4
Doors_Of_Cappadocia_Hotel	317	4
Taskonaklar	311	4
Mithra_Cave_Hotel	308	4
Cappadocia_Cave_Resort_Spa	290	4
Grand_Cave_Suites	267	5
DoubleTree_by_Hilton_Avanos_Cappadocia	199	5
Dinler_Hotels_Urgup	167	5
Goreme_Kaya_Hotel	148	4
Dere_Suites	148	4
Uchisar_Kaya_Hotel	126	4
Suhan_Cappadocia_Hotel_Spa	116	4
Cappadocia_Inn_Hotel	71	4
Perissia_Hotel_Convention_Center	54	5
Caldera_Hotel	33	4
Avrasya_Hotel	33	5
The_Cappadocia_Hotel	30	4

As can be seen in Table 1. There are 8 5-star hotels and 20 4 star hotels of Cappadocia on travel web site. In addition, most of the comments are made to the Aydinli Cave Hotel and the Cappadocia Hotel has been commented to the least.

The data has been downloaded from the HTML code source of the web interface, which contains the comments of the hotels and open to everyone's use. ASP language, one of the dynamic web programming languages, was used for downloading. With the HTTP Request method, special software has been produced that enables the reading and recording of the comments on the website. The data read is stored in the MS ACCESS database. These data were transferred to the MS EXCEL spreadsheet program and converted into a data set that can be used in the data mining application. Table 2 shows an example of a three-line data set of Hotel comments.

Table 2. *Hotel Comments Sample Data Set*

ID	Hotels	Commentator	Comment Date	Hotel Class	Title	Comments
1	Aydinli Cave Hotel	xxxx	Dec,4,2018	4	Promise fulfilled	"Though there are now many cave hotels to choose from in this towns remarkable geography, it unlikely that you will find a family more willing to serve than the one that serves as host and hotelier at the Aydinli. From generous breakfast fare offering flavors..."
2	Argos in Cappadocia	yyyy	Sep,18,2017	5	Best Hotel	"This hotel is a world-standard when it comes to hospitality, and food. Most lavish breakfast spread overlooking breathtaking scenery. The hotel staff went out of their way to make sure we had a great trip. We would definitely recommend this hotel to any..."
3	Gamirasu Cave Hotel	zzzz	Oct 24, 2015	5	Perfect	"What a wonderful experience to have, my husband and I have nothing but the best to say about the hotel and the service. We have traveled the world and nowhere have we found such friendly and amazing staff. From the lovely cleaning ladies, too..."

APPLICATION:

After the data downloading phase, the text and data mining phase were started. RapidMiner program was used for this phase. RapidMiner is an easy-to-use, open-source software that has been used for many academic studies because it is a data mining software.

The RapidMiner program, developed by using Java language by scientists at Yale University, is used to obtain meaningful information from large-scale data (Albayrak, 2015: 756). Software that provides an integrated environment for data mining, text mining, predictive analytics, and business analytics; It supports all steps of data mining processes including data preprocessing, result visualization, verification, and optimization (Arunadevi et al., 2018: 15983).

Data Preprocessing:

The excel file created in the context of data acquisition was added to the RapidMiner data mining program. In many different file formats, the program can also read files with the .xlsx extension. When adding the Excel file to the program, only the header and comment columns have been transferred according to the research purpose. During the transfer, each column is selected as an attribute. Then, a model was created for the individual separation of the words mentioned in the headlines and comments. The aim is to make the word frequencies by determining the words in the comments as attributes, and also to make them suitable for the association analysis.

The design appearance of the model created in the data preprocessing stage is shown in Figure 1.
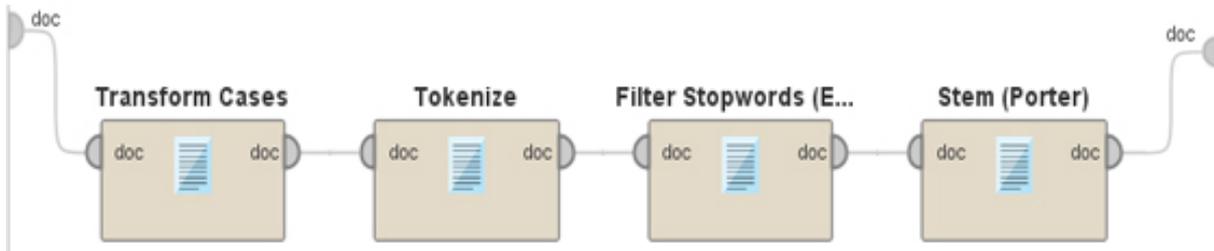
Figure 1. View of Data Pre-Processing Design

As seen in Figure 1. Transform Cases, Tokenize, Filter Stopwords, and Stem operators were used to obtain word frequencies, respectively.

As a parameter of transform cases' operator, lower case is preferred and all the words in the comments are converted to lower case. Thus, the problem of upper/lowercase letters in word groups has been eliminated. Then the words in the comments are divided into individual parts by the Tokenize operator. Filter Stopwords (English) operator with the broken words, such as the connector, pronoun, prepositional word structures was filtered. With the help of Stem (Porter) operator, the words from the same root in English are combined in a single word root. So the words like "amazing, amazingly" are transformed into one word: "amazing".

When the model shown in Figure 1 is executed, word frequencies of the words mentioned in the comments are obtained. The frequency of the word frequencies obtained from large to small (<500) is seen in Table 3.

Table 3. *Word Frequencies Regarding Obtained Comments*

Word	Total Occurrences	Document Occurrences
hotel	9985.0	6802.0
room	7218.0	5408.0
cave	4265.0	3571.0
staff	3681.0	3482.0
view	2435.0	2277.0
night	2304.0	2156.0
cappadocia	1980.0	1834.0
breakfast	1929.0	1862.0
service	1590.0	1501.0
suit room	1510.0	1348.0
tour	1079.0	955.0
food	951.0	917.0
restaurant	951.0	917.0
Turkey	794.0	760.0
balloon	791.0	739.0
town	735.0	684.0
trip	726.0	676.0
travel	681.0	632.0
bathroom	615.0	605.0

When Table 3 is examined, it is seen that words of the most comments are about hotels (6802 comments), rooms (5408 reviews), caves (3571 comments), staff (3482 comments) and views (2477 comments). The least common words are bath (605 comments), travel (632 comments), towns (684 comments) and balloons (739 comments). The comments on breakfast, service, and Cappadocia have also been remarkably repeated. The fact that the hotel rooms are in the form of a cave and have rooms with views may have influenced tourists.

Table 4. *Word Frequencies for Adjectives*

Word	Total Occurrences	Document Occurrences
great	2357.0	1931.0
help	2078.0	1904.0
good	1680.0	1379.0
amazing	1615.0	1472.0
beautiful	1603.0	1482.0
friendly	1569.0	1544.0
love	1515.0	1340.0
nice	1411.0	1200.0
experience	1312.0	1236.0
comfort	1175.0	1133.0
excel	1163.0	1055.0
wonder	1140.0	1061.0
clean	1101.0	1075.0
family	895.0	803.0
recommend	867.0	841.0
unique	753.0	729.0
perfect	697.0	653.0
spacious	693.0	688.0
enjoy	635.0	607.0

When Table 4 is examined, it is seen that no negative words are used in the comments made for the hotels. It was found that the word "great" (1931 comments) is mostly used and the word "enjoy" is the least used. In general, when looked at the positive comments "good, amazing, beautiful, nice, comfort, excellent, wonder, clean, unique, perfect" words can be said by using their satisfaction. The "help and friendly" words used by the hotel staff can mean that the tourists are happy with their behavior.

Association Analysis:

The association analysis, which identifies common relationships between objects found in the data set and explains the connection of each object with another object, is often referred to as the market basket analysis.

The word roots obtained in the data preprocessing stage were subjected to unity analysis and the rules of association were found frequently in the comments. Thus, the frequency of the words used in the comments of the hotels in the Cappadocia region was determined by the confidence parameters and information about the satisfaction or dissatisfaction of the visitors were obtained.

The model used for association analysis is shown in Figure 2.
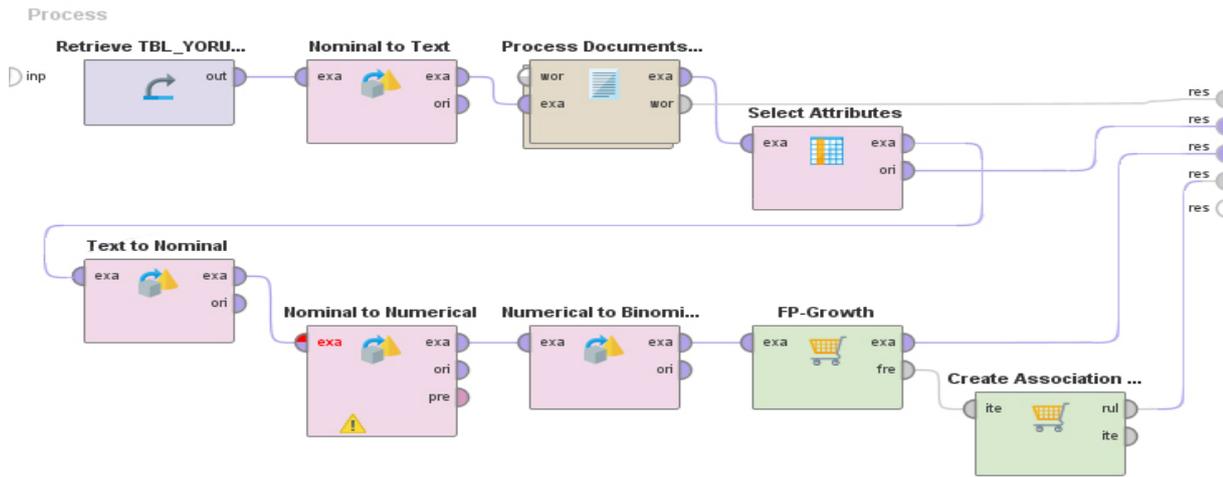
Figure 2. *Research Model*

According to Figure 2, the word frequencies obtained in the Process Documents folder are converted to Attributes one by one. It is necessary to know whether the word has passed in a comment for the extraction of the rules of association between words. For this purpose, the word matrix was created firstly. In order to form a word matrix, word matrix was obtained by passing through the Attributes Text to Nominal, Numerical, Numerical to Binominal stages. Table 4 shows the sample word matrix.

Table 4 *Sample Word Matrix*

Row No	absolute	accommod	air	airport	amaz	balloon	bath	beauti	bed
1	false	true	false	true	false	true	true	false	true
2	true	false	true	false	true	false	false	true	false
3	true	true	false	true	true	false	true	true	true
4	true	true	true	true	false	true	false	false	false
5	false	true	true	false	true	true	true	true	true

As seen in Table 4, if a word in a comment is passed, it is true and if doesn't pass, it is false. Then, the obtained word matrix was subjected to unity analysis to obtain the association rules. The preferred algorithm in rule formation is the Fp-Growth algorithm as seen in Figure 2. The Fp-Growth algorithm is an algorithm used to determine the most commonly used elements in a data set. This algorithm, which uses the divide-manage strategy, scans the items in the data set and calculates the support values and creates rules with the data tree (Ikhwan et al., 2018: 1663). Confidence and Support criteria are used to form rules for association analysis. According to the installed model, the Confidence value is 0.751 and the support value is 0.1. In this context, 37 pieces of rules were obtained. These rules are listed in Table 4.

Table 4. *Table of Association Rules for Hotel Rooms*

Premises	Conclusion	Confidence
hotel, spacious	room	0.904
spacious	room	0.901
breakfast, clean	room	0.822
huge	room	0.816
view, clean	room	0.811
view, comfort	room	0.803
clean	room	0.796
cave, clean	room	0.796
staff, clean	room	0.794
hotel, clean	room	0.788
staff, comfort	room	0.764
cave, nice	room	0.754
big	room	0.776

When Table 4 is examined, it is 0, 904 according to the confidence value of room word is passed with hotel and spacious words. The lowest confidence value is the value of the combination of big and room words (0,776). When the confidence values were examined together, the comments about the rooms were in line with the comments that the hotel and the cave rooms were large, large, clean, comfortable and pleasant.

Table 5. *Table of Association Rules for Hotel Personnel*

Premises	Conclusion	Confidence
room, help, friendly	staff	0.839
room, cave, friendly	staff	0.839
stay, help, friendly	staff	0.834
hotel, helpful, friendly	staff	0.828
location, friendly	staff	0.821
help, friendly	staff	0.816
hotel, room, friendly	staff	0.794
friendly, clean	staff	0.793
breakfast, friendly	staff	0.785
great, friendly	staff	0.775
hotel, friendly	staff	0.775
help, amazing	staff	0.757
help, beautiful	staff	0.752
help, breakfast	staff	0.751

When Table 5 is examined, the frequency of the staff, passed with the words staff and friendly words are 0, 839 according to the confidence value. The lowest confidence value is the value of the combination of help and breakfast words (0,751). When the rules are examined, it is concluded that the hotel staff is "helpful, friendly, clean, magnificent and beautiful" in the comments about the staff.

Table 6. *Table of Association Rules for Hotels*

Premises	Conclusion	Confidence
cave, beautiful	hotel	0.756
beautiful	hotel	0.756
staff, beautiful	hotel	0.755
room, cave, beautiful	hotel	0.753
staff, friendly	hotel	0.753
cappadocia, beautiful	hotel	0.753
cave, clean	hotel	0.753
help, beautiful	hotel	0.752
cave, unique	hotel	0.751

When Table 6 is examined, the "cave and beautiful" words passed with the word "hotel" together with the confidence frequency of 0, 756 according to the value. The lowest confidence value is the value of the "cave and unique" words together (0,751). When the rules are reviewed, in the comments on the hotels, the hotels are beautiful, the rooms are nice and the staff is nice too, the staff is friendly and helpful, the rooms are clean and unique.

Table 7. Table of Association Rules for Hotels

Premises	Conclusion	Confidence
highly	recommend	0.863

When Table 7 is examined, the frequency of the word "advise" is passed together with the word "a high degree" is 0,863 according to the confidence value. Taking this rule into consideration, it can be said that commentators tend to recommend a high level of hotels in Cappadocia.

Discussion

This study aimed to determine the satisfaction levels of the customers who visited the 4 and 5-star hotels in Cappadocia. This study, in which text mining analysis was used, leading to more generalizable results than the data collected with traditional methods. The study was conducted on a total of 10263 reviews among 28 hotels. This suggests that a wide sampling has been reached. As a result of the data collected about the hotels via travel web sites, room experience, staff service and hotel reviews, in general, were reached and analysis of the union was conducted and the satisfaction levels of the customers were determined by the words expressing their satisfaction with these three dimensions. These dimensions (room experience, staff service, and hotel) that affect satisfaction, support the results of previous studies (Choi and Chu, 2001, Skogland and Siguaw, 2004). The fact that the data obtained by the method of text mining has facilitated generalizations. This method can be defined as the most common factors that customers perceive when expressing their opinions and sharing their experiences online. In this respect, they have directly received the service quality, physical condition, etc., they expressed their comments about many factors as easily as they felt. It is thought that providing online environments with such facilities is important in obtaining reliable and valid information.

Theoretical and Managerial Contributions

There are some theoretical contributions off the study. This article has reviewed 10263 reviews to analyze the business value of online textual reviews in the hospitality industry and show how to use large data. Online comment reviews provide a roadmap for analyzing the impact of eWOM (E-Word of Mouth) with the specifications of online text reviews.

This study has many valuable effects in terms of managerial practices. The first impact helps hotel managers and entrepreneurs to easily achieve the rich content that affects the satisfaction of customers and to create satisfaction-oriented activities. It is also stated that new professions will emerge in the coming years for the analysis of this data. In this respect, managers and investors need to invest in these areas. The second effect is that it helps to match the comments of the tourists and the perceptions of the hotels. It also supports the determination of factors affecting the satisfaction of customers. The results of the study have an impact on the target market planning, market segmentation and product positioning activities of the hotel's marketing managers. For example, it can be determined by examining the comments of the tourists in which part of the hotel should be differentiated in order to create competition other than price. The planning of studies related to dissatisfaction with negative comments can be realized. The third effect is that it makes tourists' comments easier to measure e-WOM impact. The E-WOM effect is a powerful element that affects tourists' re-booking decisions.

Conclusion

This study aims to determine the satisfaction levels of customers by reviewing their online comments. For this purpose, 10623 comments were analyzed, factors affecting satisfaction and satisfaction levels were determined. It was determined that the tourists made comments on the room experience, staff and hotel dimensions. They have stated that the rooms were great, clean, gorgeous, amazing, unique and spacious. The hotel staff was found to be helpful and friendly. It is concluded that hotels are generally clean, beautiful and unique. It can be said that there is a high level of satisfaction from the hotels. Because no negative comments have been found. One of the most interesting comment is the use of words "in a high way" and "advice to others" together. It is possible to conclude that tourists will recommend the hotels at a high level.

Limitations and Future Studies

There are some limitations to the study. The study was limited only to the hotels operating in the Cappadocia region. Hotel comments of different regions can also be reviewed. Instead of hotels, other tourism businesses (restaurants, tourist destinations, etc.) comments can be subject of study that can be done. It was also carried out on only one online review platform. In future studies, different platforms can be addressed.

References

Albayrak, M. (2015). Bilimsel araştırmalarda veri madenciliği kullanımı. International Journal of Social Sciences and Education Research, 3(3), 751-760.

Arunadevi, J., Ramya, S., and Raja, M. R. (2018). A study of classification algorithms using Rapidminer. International Journal of Pure and Applied Mathematics, 119(12), 15977-15988.

Berezina, K., Bilgihan, A., Cobanoglu, C., and Okumus, F. (2016). Understanding satisfied and dissatisfied hotel customers: Text mining of online hotel reviews. Journal of Hospitality Marketing and Management, 25(1), 1–24.

Blal, I., and Sturman, M.C., (2014). The differential effects of the quality and quantity of online reviews on hotel room sales. Cornell Hospitality Q. 55 (4), 365–375.

Büschken, J., and Allenby, G. M. (2016). Sentence-based text analysis for customer reviews. Marketing Science, 35(6), 953–975.

Chau, M., and Xu, J. (2012). Business intelligence in blogs: Understanding consumer interactions and communities. MIS quarterly, 36(4), 1189-1216.

Chung, N. and Koo, C. (2015), "The use of social media in travel information search", Telematics and Informatics, Vol. 32 No. 2, pp. 215-229.

Choi, T. Y., and Chu, R. (2001). Determinants of hotel guests' satisfaction and repeat patronage in the Hong Kong hotel industry. International Journal of HospitalityManagement, 20(3), 277-297.

Clemons, E., Gao, G., and Hitt, L. (2006). When online reviews meet hyperdifferentiation: A study of the craft beer industry. Journal of Management Information Systems, 23(2), 149-171.

Gao, B., Hu, N., and Bose, I. (2017). Follow the herd or be myself? An analysis of consistency in the behavior of reviewers and helpfulness of their reviews. Decision Support Systems, 95, 1–11.

Gao, S., Tang, O., Wang, H., and Yin, P., (2018). Identifying competitors through comparative relation mining of online reviews in the restaurant industry. Int. J. Hospitality Manag. 71, 19–32.

Geetha, M., Singha, P., and Sinha, S., (2017). Relationship between customer sentiment and online customer ratings for hotels-An empirical analysis. Tourism Manag. 61, 43–54.

Ghose, A., and Ipeirotis, P. (2011). Estimating the helpfulness and economic impact of product reviews: Mining text and reviewer characteristics. IEEE Transactions on Knowledge and Data Engineering (TKDE), 23(10), 1498-1512.

Guo, Y., Barnes, S. J., and Jia, Q. (2017). Mining meaning from online ratings and reviews: Tourist satisfaction analysis using latent Dirichlet allocation. Tourism Management, 59, 467–483.

He, W., Tian, X., Tao, R., Zhang, W., Yan, G., and Akula, V., (2017). Application of social media analytics: a case of analyzing online hotel reviews. Online Inf. Rev. 41 (7), 921–935.

Ho-Dac, N. N., Carson, S. J., and Moore, W. L. (2013). The effects of positive and negative online customer reviews: Do brand strength and category maturity matter? Journal of Marketing, 77(6), 37-53.

Ikhwan, A., Yetri, M., Syahra, Y., Halim, J., Siahaan, A. P. U., Aryza, S., and Yacob, Y. M. (2018). A Novelty of Data Mining for Promoting Education based on FP-Growth Algorithm. International Journal of Civil Engineering and Technology, 9(7), 1660-1669.

Kim, W.G., Lim, H.,and Brymer, R.A., (2015). The effectiveness of managing social media on hotel performance. Int. J. Hospitality Manag. 44, 165–171.

Lee, C. C., and Hu, C. (2005). Analyzing Hotel customers' E-complaints from an internet complaint forum. Journal of Travel and Tourism Marketing, 17(2–3), 167–181.

Leung, D., Law, R., Van Hoof, H. and Buhalis, D. (2013), "Social media in tourism and hospitality: a literature review", Journal of Travel and Tourism Marketing, Vol. 30 Nos 1/2, pp. 3-22.

Levy, S. E., Duan, W., and Boo, S. (2013). An analysis of one-star online reviews and responses in the Washington, DC, lodging market. Cornell Hospitality Quarterly, 54(1), 49–63.

Li, H., Ye, Q., and Law, R. (2013). Determinants of customer satisfaction in the hotel industry: An application of online review analysis. Asia Pacific Journal of Tourism Research, 18(7), 784–802.

Lu, W., and Stepchenkova, S. (2012), Ecotourism experiences reported online: classification of satisfaction attributes, Tourism Management, Vol. 33 No. 3, pp. 702-712.

Ludwig, S., De Ruyter, K., Friedman, M., Brüggen, E.C., Wetzels, M., and Pfann, G., (2013). More than words: the influence of effective content and linguistic style matches in online reviews on conversion rates. J. Marketing 77 (1), 87–103.

Manickas, P. A., and Shea, L. J. (1997). Hotel complaint behavior and resolution: A content analysis. Journal of Travel Research, 36(2), 68–73.

Mankad, S., Han, H. S., Goh, J., and Gavirneni, S. (2016). Understanding online hotel reviews through automated text analysis. Service Science, 8(2), 124–138.

O'Connor, P. (2008), "User-generated content and travel: a case study on Tripadvisor.com", Proceedings of the International Conference in Innsbruck Information and Communication Technologies in Tourism, Austria, Springer, Vienna, pp. 47-58.

Pantano, E., Priporas, C.-V., and Stylos, N. (2017). 'You will like it!' using open data to predict tourists' response to a tourist attraction. Tourism Management, 60, 430–438.

Qu, Z., Zhang, H., and Li, H., (2008). Determinants of online merchant rating: a content analysis of consumer comments about Yahoo merchants. Decis. Support Syst. 46 (1), 440–449.

Skogland, I., and Siguaw, J. A. (2004). Are your satisfied customers loyal? Cornell hospitality Quarterly, 45(3), 221-234.

Surowiecki. (2005). The wisdom of crowds. Anchor.

Tian, X., He, W., Tao, R., and Akula, V., (2016). Mining online hotel reviews: a case study from hotels in China. In: Proceeding of Twenty-second Americas Conference on Information Systems. San Diego, CA, USA.

Vermeulen, I.E. and Seegers, D. (2009), "Tried and tested: the impact of online hotel reviews on consumer consideration", Tourism Management, Vol. 30 No. 1, pp. 123-127.

Xiang, Z., Du, Q., Ma, Y., and Fan, W., (2017). A comparative analysis of major online review platforms: implications for social media analytics in hospitality and tourism. Tourism Manag. 58, 51–65.

Xiang, Z., Schwartz, Z., Gerdes, J.H.J., and Uysal, M., (2015). What can big data and text analytics tell us about hotel guest experience and satisfaction? Int. J. Hospitality Manag. 44, 120–130.

Xie, H., Miao, L., Kuo, P. J., and Lee, B. Y. (2011). Consumers' responses to ambivalent online hotel reviews: The role of perceived source credibility and pre-decisional disposition. International Journal of Hospitality Management, 30(1), 178-183.

Xu, X., and Li, Y. (2016). The antecedents of customer satisfaction and dissatisfaction toward various types of hotels: A text mining approach. International Journal of Hospitality Management, 55, 57–69.

Ye, Q., Law, R., and Gu, B. (2009a). The impact of online user reviews on hotel room sales. International Journal of Hospitality Management, 28(1), 180-182.

Ye, Q., Law, R., Li, S., and Li, Y. (2011). Feature extraction of travel destinations from online Chinese-language customer reviews. International Journal of Services Technology and Management, 15(1), 106-118.

Zhang, D., Zhou, L., Kehoe, J.L., and Kilic, I.Y., (2016a). What online reviewer behaviors really matter? Effects of verbal and nonverbal behaviors on detection of fake online reviews. J. Manag. Inf. Syst. 33 (2), 456–481.

Zhang, Z., Ye, Q., Law, R., and Li, Y., (2010). The impact of e-word-of-mouth on the online popularity of restaurants: a comparison of consumer reviews and editor reviews. Int. J. Hospitality Manag. 29 (4), 694–700.

Zhou, L. Q., Ye, S., Pearce, P. L., and Wu, M. Y. (2014). Refreshing hotel satisfaction studies by reconfiguring customer review data. International Journal of Hospitality Management, 38(38), 1–10.