

# Assessment of Hotel Guest Satisfaction Using Sentiment Analysis: A Case Study of Maldives Hotels

Hauwa'u Uraifa Shehu<sup>1</sup>, A.F. Donfack Kana<sup>2</sup>, and Fatima Sulaiman<sup>3</sup>

<sup>1</sup>School of Science, Engineering and Environment, University of Salford  
Manchester, United Kingdom

<sup>2,3</sup>Department of Computer Science, Ahmadu Bello University, Zaria  
[hauwaurifa@gmail.com](mailto:hauwaurifa@gmail.com)<sup>1</sup>, [donfackkana@abu.edu.ng](mailto:donfackkana@abu.edu.ng)<sup>2</sup>, [fateemahsb57@gmail.com](mailto:fateemahsb57@gmail.com)<sup>3</sup>

## Abstract

*Nowadays online reviews by hotel customers greatly influence business as potential new consumers seek unbiased information while making their hotel booking decisions. Hotel management and marketers are more aware of the impact of online reviews on financial performance. This awareness arises from the universal consensus that internet consumer reviews have a significant impact on hotel business performance. Customers use social media to share information about products and services, and online reviews have a substantial influence on customer purchasing decisions. The goal of this study is to provide formative assessment feedback on Maldives hotels using word cloud technique. This include investigating the hotel that is mostly used by guests, finding out the percentage of positive and negative comments made about the hotel, and also assessing the type of comments the majority of customers give about the services rendered to them. Data from 104 distinct Maldives hotels were utilized in this case study to provide quick visual insight using a word cloud approach with R programming language. The result shows that, more than 80% of the comments are positive, implying that the vast majority of these hotels' customers are pleased with their accommodations and services.*

**Keywords:** *Text mining, Sentiment analysis, Guest Satisfaction*

## 1. Introduction

The rapid development of information technology over the last few decades has greatly influenced the way individuals seek and disseminate travel-related information. Nowadays, the large volume of emotive reviews are posted online. In tourism industry where consumers seek unbiased information while making their hotel booking decisions, previous guests' experiences expressed in the hotel reviews across social media generally form the basis for their decision. The importance of other people's opinions in decision-making process is well understood. Pang & Lee (2008) noted long before the widespread use of the World Wide Web that many people seek friends for recommendations. The process of business interaction with clients has evolved with social media. Customers are urged to use social networks to maintain ties with others through their current online business environment (Joo et al., 2011; Yue et al., 2019). Social media customers use social media to share product and service information, and online review information has a crucial impact on customer purchasing decisions (Camacho et al., 2020; Coulter & Roggeveen, 2012).

Satisfaction and improving the service process have been changed by social media. It raises the awareness of the clients of a hotel, inspires the buying conduct of customers and influences hotel performance overall (Anderson, 2012; Chittiprolu et al., 2021). Most tourists nowadays utilise the internet as one of their primary information source (Chittiprolu et al., 2021; Sparks & Browning, 2010). Some of these websites also offer hotel ratings in line with the opinions of past guests and propose that their hotels offer their online clients. An

increasing number of Internet users monitor and rely on customer reviews online (DeMers, 2015). Accordingly, Bright Local study shows that 88% of consumers trust internet evaluations as much as they trust personal recommendation (DeMers, 2015). Both researchers and practitioners have attached great importance to the evaluation of hotel customer satisfaction. Lots of efforts have been performed in that direction through mining reviews posted online by previous guests. Sentiment analysis has become a key task for modern business and society. Sentiment analysis has got a lot of attention from the tourism industry and academia as social media and social computing have grown in popularity (Escandon-Barbosa & Salas-Paramo, 2021; Gao et al., 2015). Sentiment analysis strives to classify texts into distinct categories, such as good, negative, and others, by identifying the sentiment present in them. Use of sentiment analysis in tourism can have a significant impact on the industry. Tourists can read other people's comments about sites, hotels, and restaurants on travel blogs before departing, study the opinions expressed on these blogs, and make decisions about which hotels to stay in or which restaurants to dine at. Also, hoteliers and managers can make continual adjustments based on visitor feedback in order to attract more guests.

Several research areas are now utilising sentiment analysis. Many studies used datasets acquired from social media to review movies online through employing sentiment analysis. Despite the complexity and difficulties of the problems, numerous organisations and academic institutions are introducing new techniques and web applications which might resolve some of the aforementioned difficulties (Gao et al., 2015; Lu & Zheng, 2021). Technically, pre-processing, extraction of features and classification of text are methods included in sentiment analysis. Hence, sentiment analysis is a text mining technique that discovers meanings, patterns, and structure hidden in unstructured data that is used to investigate content collections and individual document (Xie et al., 2019). However, text mining have really gone far to encompass integrated framework using techniques from natural language processing (NLP), data mining, machine learning, knowledge management, and information retrieval (IR) (Goslee & Urban, 2007; Liu, 2012; Pejic-Bach et al., 2019). It also integrates tools and techniques used to gain knowledge from unstructured data, including: analysis of feelings, information retrieval, content management, exploratory analysis, categorization, idea extraction, ontology management and description. In addition, the field of text mining is exploratory analysis, summarization and categorization. The text mining process follows few steps in respect to mine the information as shown in Figure 1:

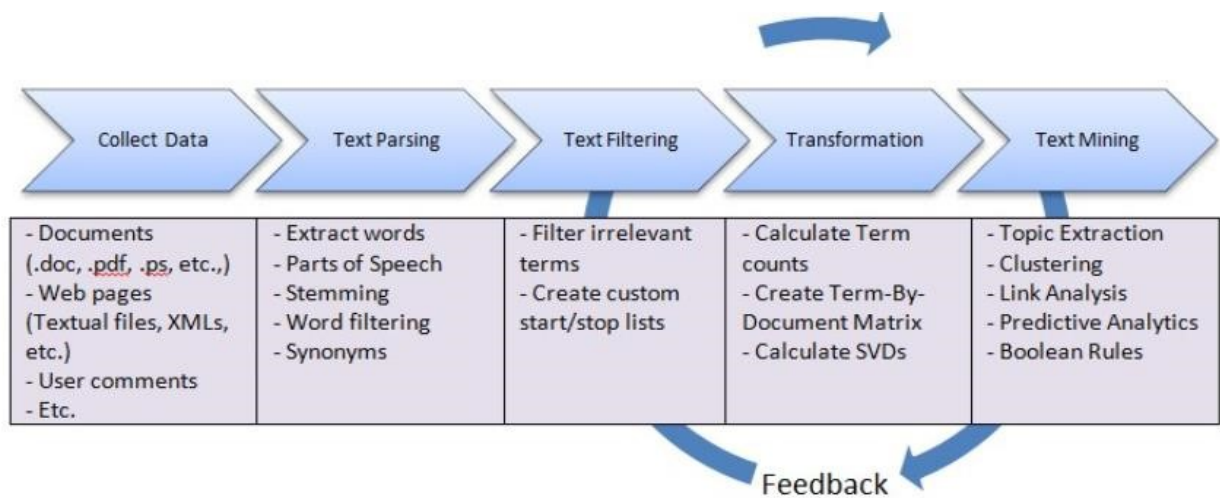


Figure 1: Text mining process (Chakraborty et al., 2013)

Considering the significance and emergence of sentiment analysis particularly in tourism and online services, this study looked at the structures and patterns of Maldives hotels in order to locate and analyse the opinions and feelings expressed in tourist reviews. Furthermore, the study analyses the hotels and established the most frequently used by guests and assess the types of comments made by the majority of customers about the services they received.

## 2. Related Works

### A- Hotel guest satisfaction

Guest happiness is regarded to be the ultimate goal of any hotel business (Chittiprolu et al., 2021; Torres et al., 2014). Guest satisfaction, according to (Xiang et al., 2015), is "the guest's evaluation of his/her experience through engagement with numerous service areas." Literature has been comprehensive about the factors of hotel client happiness. For example, (Li et al., 2013) employed text mining and content analysis to evaluate the hotels in Chinese cities. The data showed that space, travel comfort, closeness to tourist places, and value for money, are the primary characteristics of the happiness of guests. Additionally, Ding et al., (2021) explored the sources of satisfaction and dissatisfaction aspects of accommodations using both supervised learning and unsupervised learning. The outcome indicated numerous major determinants of guest satisfaction, including location, sleep disturbance, room size, communication, amenities, booking and refund among others. Similarly, (Ko, 2021) highlighted that hotel employees and their attitudes are classified as both satisfaction and dissatisfaction variables. Hence, in the current hotel industry, the utilization of raw data such as internet reviews in customer satisfaction understanding and monitoring has therefore become vital.

### B- Sentiment analysis

Sentiment analysis is concerned with categorizing views expressed in textual documents. These text units are often divided into multiple categories based on the views conveyed, such as positive, negative or neutral (Pang & Lee, 2008). For sentiment analysis, many researchers use machine learning approaches, which entail training a classifier on datasets and then using the taught model to classify fresh documents. A dictionary of word lexicons, for example, was offered by several authors as an alternative technique (Raut & Londhe, 2014). A great number of studies have been conducted on sentiment-based classification. Semantic orientation and Machine learning techniques are the two primary groups of methodologies used by researchers. The semantic orientation technique is an unsupervised method that assesses the general sentiment of a document based on the semantic orientation of the keywords included within it without previous training. Using reviews from the movie industry, (Chaovalit & Zhou, 2005) compared the two aforementioned techniques. While the unsupervised semantic orientation technique achieves poor accuracy, it is significantly more efficient in real-time applications, according to the findings. On the other hand, the supervised machine learning methodology produces more accurate classification results but has the disadvantage of taking a long time to train the classifier. Consequently, unsupervised techniques are commonly used to label a corpus, which is then used for supervised learning.

Hatzivassiloglou & McKeown, (1997) presented a method for determining the semantic orientation of keywords automatically. In order to acquire information indirectly from the corpus, the authors employ conjunctions between keywords. More specifically, when two keywords are connected by conjunctions like 'and', the keywords are most likely of the same orientation (e.g., reasonable and appropriate, reasonable and ruthless), whereas when the keywords are connected by 'but', or other similar conjunctions, they are most likely of different orientations (e.g., reasonable but ruthless). When these limitations were paired with other morphological criteria, the accuracy of predicting whether two conjoined keywords are of the same or different orientation was 82 percent. (Pang et al., 2002), were among the early users of supervised learning at the document level to solve the automatic sentiment classification problem. (Pang et al., 2002) compared the effectiveness of three machine learning methods [Naive Bayes (NB), and Support Vector Machines (SVMs), Maximum Entropy (MaxEnt)]

using a corpus of movie reviews with various characteristics such as unigrams, bigrams, part-of-speech information, and location information. Similarly, (Kennedy & Inkpen, 2006) suggested a way to figure out how people feel about a movie review by combining two different methods: a term-counting strategy and a machine-learning strategy that uses SVMs unigrams as features. However, a combination of both methods gave superior outcome (85.4 percent accuracy).

### 3.Methodology

This study seeks to discover the aspects that contribute to higher customer satisfaction in Maldives hotels based on online comments. The Maldives is a gorgeous Indian Ocean Island that is popular, dynamic, and a favorites holiday destination for tourists from all over the world. Given its beachy setting, it is known for its abundance of hotels and resorts (Camacho et al., 2020) Data obtained from 104 different Maldives hotels were used in this case study.

A word cloud was used to examine the characteristics of clients’ satisfaction and dissatisfaction at the hotel. The suggested research approach included a visualization tool for representing textual data. The benefit of this technique is that it can extract important patterns from thousands of tediously read comments and automatically group them together for easier comprehension.

Word cloud is a text mining approach that highlights the most often used words in any given context. To produce a Word cloud, text processing procedures such as tokenization, standardizing, cleaning, eliminating common words, stemming, and indexing are used. Tokenization is a technique for breaking down each sentence in a text into several words. This process is simplified by using white space as a word delimiter. Additionally, the process of standardisation entails altering all characters in all publications to lower case. Similarly, the cleaning process removes unique letters and symbols such as the punctuations, exclamation marks, question marks, and dollar signs. Stemming is the process of changing the ends of words to create a root form. The purpose of indexing is to create a list of all unique terms and to count the number of times each word appears in all texts.

Therefore, the procedure adopted in this paper is by reading text file representing the online comments about Maldives hotels, building the corpus, cleaning of input text (e.g., removing “commas, web links etc.”), term doc matrix (creating a matrix for how often each word appears since there are so many words in the document), result plot and finally analyse the results. A series of activities are performed to efficiently mine information in text mining as shown in Figure 2.

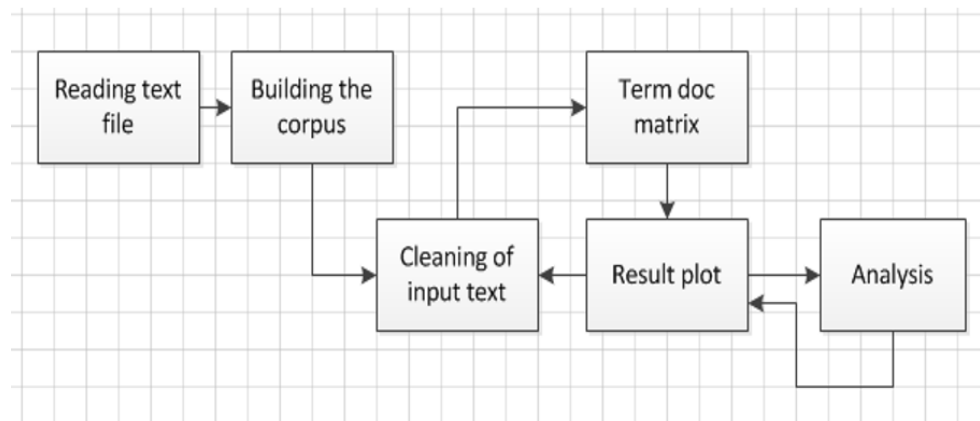


Figure 2: Text mining workflow

*Preparation of Dataset*

Unstructured text data called `maldive_hotel_reviews.csv` was downloaded and read into the data frame using the R Studio. The first step was to read the data file in the R environment and then the imported text is read in CSV. The dataset has 21,093 observations and 8 variables which are: `Review.id`, `Hotel.Name`, `Total.review.count`, `Review.Date`, `Review_viaMobile`, `Guest.Location`, `Review.Heading` and `Review` as shown in figure 3. The tail variables, which are the last six columns of the dataset are shown in figure 4.

```
> names(Maldives_Hotel)
[1] "Review.id"      "Hotel.Name"     "Total.review.count" "Review.Date"
[5] "Review_viaMobile" "Guest.Location" "Review.Heading"    "Review"
> head(Maldives_Hotel)
  Review.id      Hotel.Name Total.review.count  Review.Date
1 rn629306037 Holiday Inn Resort Kandooma Maldives      1679  October 29, 2018
2 rn623195151 Holiday Inn Resort Kandooma Maldives      1679  October 8, 2018
3 rn627907077 Holiday Inn Resort Kandooma Maldives      1679  October 24, 2018
4 rn623362603 Holiday Inn Resort Kandooma Maldives      1679  October 8, 2018
5 rn623255673 Holiday Inn Resort Kandooma Maldives      1679  October 8, 2018
6 rn615351040 Holiday Inn Resort Kandooma Maldives      1679  September 10, 2018
  Review_viaMobile      Guest.Location
1      via mobile      Waikiki, Hawaii
2      via mobile      Berkshire, United Kingdom
3      via mobile      Edinburgh, United Kingdom
4      <NA>      Melbourne, Australia
```

Figure 3: Description of the data set

```
> tail(Maldives_Hotel)
  Review.id      Hotel.Name Total.review.count  Review.Date  Review_viaMobile
21088 rn56122603 Fun Island Resort      906  February 16, 2010      <NA>
21089 rn55366842 Fun Island Resort      906  February 7, 2010      <NA>
21090 rn53563419 Fun Island Resort      906  January 15, 2010      <NA>
21091 rn53517375 Fun Island Resort      906  January 15, 2010      <NA>
21092 rn2964678 Fun Island Resort      906  December 29, 2004      <NA>
21093 rn2603322 Fun Island Resort      906  September 26, 2004      <NA>
  Guest.Location      Review.Heading
21088      London Paradise on Earth!!!! HIGHLY RECOMMENDED!!!
21089      Hampshire      Fantastico!
21090      NYC      fun island not so fun
21091      www13      ASBESTOS !!! Stay away !!!
21092      Surrey      Absolutely Fantastic!
21093 Macclesfield, UK      Stunning resort and 1st Class service
```

Figure 4: Tail variables



### Data Pre-Processing Steps

The data file was translated into a comprehensible format in this step. This is because real-world data is frequently partial, inconsistent, or insufficient in particular activities, and is likely to contain many errors. Therefore, data pre-processing is a verified and reliable method of addressing these issues of inconsistency and insufficiency. Hence, after loading the data file in the R environment and the imported text is read in CSV, the task utilizes the 'tm' package, which contains a number of predefined sources for handling directories, vectors that read each component as a document, and data frame-like formats like CSV files, including DirSource, VectorSource, and DataframeSource. The structure for the application of text mining is shown in Figure 5.

```
> Maldives_Hotel <- read.csv("maldives_hotel_reviews.csv", header = TRUE)
> install.packages("tm")
WARNING: Rtools is required to build R packages but is not currently installed. Please download
and install the appropriate version of Rtools before proceeding:

https://cran.rstudio.com/bin/windows/Rtools/
Installing package into 'C:/Users/HP/Documents/R/win-library/3.6'
(as 'lib' is unspecified)
trying URL 'https://cran.rstudio.com/bin/windows/contrib/3.6/tm_0.7-6.zip'
Content type 'application/zip' length 1370651 bytes (1.3 MB)
downloaded 1.3 MB

package 'tm' successfully unpacked and MD5 sums checked
```

Figure 5: the tm package

### Create Corpus

After pre-processing, the Corpus was created using the corpus function of the tm package. A corpus that represents a collection of texts (data), typically marked with text annotations is now created. The corpus is the main structure that tm uses for storing and manipulating text documents. However, two corpuses were created for the two dependent variables; the Review attribute and the Hotel.Name attribute. Hence, the text file is converted to corpus as shown in Figure 6.

```
> #converting the text file to corpus
> corpus_hotel.name <- Corpus(VectorSource(Maldives_Hotel$Hotel.Name))
> corpus_hotel.name
<<SimpleCorpus>>
Metadata: corpus specific: 1, document level (indexed): 0
Content: documents: 21093
> corpus_hotel.review <- Corpus(VectorSource(Maldives_Hotel$Review))
> corpus_hotel.review
<<SimpleCorpus>>
Metadata: corpus specific: 1, document level (indexed): 0
Content: documents: 21093
> |
```

Figure 6: the corpuses showing the dependent variables

*Normalization*

The next step is the data normalization, which is the method of converting a text into a standard form. It includes cleaning the data by converting to lowercases (this is a form of normalization to enable R program to carry out comparison activities properly), removing commas, punctuation marks and other text that are irrelevant making it into a more uniform form which is known as ‘stemming’ in text mining. Furthermore, the `getTransformations()` function is used to retrieve the list of predefined transformations (mappings) which can be used with `tm_map` function as shown in Figure 7.

```
> getTransformations()
[1] "removeNumbers" "removePunctuation" "removewords" "stemDocument"
[5] "stripwhitespace"
#convert stem document.
corpus_hotel.name <- tm_map(corpus_hotel.name,stemDocument)
corpus_hotel.review <- tm_map(corpus_hotel.review,stemDocument)
#remove numbers
corpus_hotel.name <- tm_map(corpus_hotel.name,removeNumbers)
inspect(corpus_hotel.name[8])
corpus_hotel.review <- tm_map(corpus_hotel.review,removeNumbers)
inspect(corpus_hotel.review)
#remove whitespaces
corpus_hotel.name <- tm_map(corpus_hotel.name,stripwhitespace)
corpus_hotel.review <- tm_map(corpus_hotel.review,stripwhitespace)
#remove punctuations
corpus_hotel.name <- tm_map(corpus_hotel.name,removePunctuation)
corpus_hotel.review <- tm_map(corpus_hotel.review,removePunctuation)
```

Figure 7: predefined transformations

Figure 8 shows the reviews of the converted text to lower case:

```
> corpus_hotel.review <- tm_map(corpus_hotel.review,tolower)
warning message:
In tm_map.simpleCorpus(corpus_hotel.review, tolower) :
  transformation drops documents
> inspect(corpus_hotel.review[8])
<<SimpleCorpus>>
Metadata: corpus specific: 1, document level (indexed): 0
Content: documents: 1

[1] we went for a holiday along with our month old kid service courteous staff and service was
impeccable they went out of the way to make meals for our kid as oer our requirements ambience
we stayed in beach villamore
> |
```

Figure 8: Text converter to lower case

The removal of stopwords aids accuracy and reduces redundancy, as well as filters out words that are not important. Removing all these stop words can be seen in Figure 9.

```
> #remove stop words from the document
> corpus_hotel.review <- tm_map(corpus_hotel.review, removeWords, stopwords("English"))
warning message:
In tm_map.simpleCorpus(corpus_hotel.review, removeWords, stopwords("English")) :
  transformation drops documents
> inspect(corpus_hotel.review[8])
<<SimpleCorpus>>
Metadata: corpus specific: 1, document level (indexed): 0
Content: documents: 1

[1] went holiday along month old kid service courteous staff service impeccable went
way make meals kid oer requirements ambience stayed beach villamore
```

Figure 9: Removal of stopwords

Next is to stem the words to their root on all reviews present in the corpus as shown in Figure 10.

```
> #convert stem document.
> corpus_hotel.review <- tm_map(corpus_hotel.review, stemDocument)
warning message:
In tm_map.simpleCorpus(corpus_hotel.review, stemDocument) :
  transformation drops documents
```

Figure 10: Stem document

*Document Term Matrix (DTM)*

DTM is a matrix that lists all terms that happen in the body. Records in DTM are represented by rows and columns represent the terms (or words). If a word appears n times in a specific document, the matrix entry for that row and column is n, if it does not occur at all, the entry is 0. As shown in figure 11, it helps to switch it from text words to numbers.

```
> dtm_hotel.review <- TermDocumentMatrix(corpus_hotel.review, control = list(minwordLength=c(1,
Inf),
+                                     bounds = list(global = c(40, Inf))))
> dtm_hotel.review
<<TermDocumentMatrix (terms: 1304, documents: 21093)>>
Non-/sparse entries: 398155/27107117
Sparsity           : 99%
Maximal term length: 14
weighting          : term frequency (tf)
> |
```

Figure 11: DTM

The findfreqterm function is used for finding the frequency of each word in the text word matrix as illustrated in figure 12. This indicates the frequent term in the hotel review in Maldives.



```
> #findFreqTerms function to find frequent terms in a document-term
> findFreqTerms(dtm_hotel.review, lowfreq = 2)
[1] "accommod" "bar" "class" "crew" "dive"
[6] "ever" "favorit" "maldiv" "need" "none"
[11] "one" "profession" "resort" "staff" "take"
[16] "visit" "world" "youmor" "absolut" "afford"
[21] "australia" "boat" "enjoy" "everi" "gorgeous"
[26] "home" "moment" "night" "still" "stop"
[31] "surround" "transfer" "unreal" "way" "wemor"
[36] "year" "airport" "anoth" "arriv" "beach"
[41] "breakfast" "comfort" "excel" "experi" "hotel"
[46] "hous" "ismor" "recept" "right" "servic"
[51] "speedboat" "start" "stay" "unfortun" "welcom"
```

Figure 12: findfreqterm function

*Identification of the Variables*

The identified dependent attributes are the Hotel.Name attribute and the Review attribute. These attributes form the basis of the review given to the hotel and its management.

- Hotel.Name; these may be among the 104 different Maldives hotels
- Review; the comment given by the customers

The independent variables are:

- Review\_id: the Identification of the reviews
- Review.Heading: this is a heading or title of the review
- Total review counts: the total review for a hotel in a day
- Review.Date: the day the review was made
- Review\_viaMobile: whether the review was done with a mobile or not
- Guest.Location: the location of the guest hotel

**4. Results and Discussion**

In an attempt to uncover hidden structures in the Maldives hotels reviews, it was realized that the csv file includes several duplicates, which needed to be removed using Microsoft Excel. The materials were divided into two sections using the R software, focusing on columns such as "Hotel name," and "Reviews." Many of the tourists who visited the hotels in the Maldives praised the hospitality and services provided by the hotel employees. The most frequent words from the word cloud of the Hotel.Name are Maldives, Resort, Holiday Inn, Kandooma, Island spa, Beach hotel (Figure 13). These are the most frequent words that show the most hotels visited. Also, the most frequent words from the word cloud of Reviews are stay, great, friendly, food, staff, good, great service and most of these words are related to positive reviews (Figure 14).



honeymoon. The food and services were praised, demonstrating that Maldivian hotels excel in hospitality. Another important finding is the Word clouds that provides a more visual picture. A term's appearance in the word cloud grows in proportion to its frequency of use. Certain keywords have been used more frequently in the case of hotels and reviews, as seen by the word clouds.

Many claims of positive reviews could be inferred from the word series created by the word cloud model and most of the customers have good things to say about the hotels and their facilities. However, more than 80% of the word cloud terms are associated with positive remarks, implying that the majority of these hotels' clients had positive things to say about the services they render as shown in Figure 15.

```
> sentiment(stem_hotel.review)
Review_Type Count
1    Postive 44762
2    Negitive 6756
[1] "Percentage of Positive Reviews: 86.89 %"
```

Figure 15: percentage of reviews

## 5. Conclusion

This paper employs a word cloud, which searches the internet for user sentiments expressed through reviews, comments, and other forms of content. Nowadays, social media platforms are frequently used to share user reviews on a wide range of topics, including movies, news, food, fashion, and politics, among others, however these reviews and opinions play an important role in determining a user's level of satisfaction with a particular business. Hence these reviews are then used to determine the polarity, i.e., positive, negative, and neutral. The study found that the most often used terms in the word cloud are excellent, friendly, food, staff, nice, wonderful service, and the majority of these phrases are associated with positive reviews. However, more than 80% of the words are related to positive comments and therefore could be concluded that most of the customers of these hotels had good reviews to say about the hotels and their services. Moreover, more prior reviews lead to more future reviews, demonstrating the powerful influence of social media reviews and the importance of knowing the distinctiveness of using social media marketing strategy in various industries. Thus, these could influence future customers' travel decisions.

## References

- Anderson, C. (2012). *The impact of social media on lodging performance*.
- Camacho, D., Panizo-Lledot, A., Bello-Orgaz, G., Gonzalez-Pardo, A., & Cambria, E. (2020). The four dimensions of social network analysis: An overview of research methods, applications, and software tools. *Information Fusion*, 63, 88–120.
- Chakraborty, G., Pagolu, M., & Garla, S. (2013). Text mining and analysis. In *Practical methods, examples, and case studies using SAS*. Cary: SAS Institute Inc.
- Chaovalit, P., & Zhou, L. (2005). Movie review mining: A comparison between supervised and unsupervised classification approaches. *Proceedings of the 38th Annual Hawaii International Conference on System Sciences*, 112c-112c.

- Chittiprolu, V., Samala, N., & Bellamkonda, R. S. (2021). Heritage hotels and customer experience: a text mining analysis of online reviews. *International Journal of Culture, Tourism and Hospitality Research*.
- Coulter, K. S., & Roggeveen, A. (2012). “Like it or not”: Consumer responses to word-of-mouth communication in on-line social networks. *Management Research Review*.
- DeMers, J. (2015). How important are customer reviews for online marketing? *Forbes*, [Http://Www. Forbes. Com/Sites/Jaysondemers/2015/12/28/How-Important-Are-Customer-Reviews-for-Online-Marketing](http://www.forbes.com/sites/jaysondemers/2015/12/28/how-important-are-customer-reviews-for-online-marketing).
- Ding, K., Choo, W. C., Ng, K. Y., Ng, S. I., & Song, P. (2021). Exploring sources of satisfaction and dissatisfaction in Airbnb accommodation using unsupervised and supervised topic modeling. *Frontiers in Psychology*, 12.
- Escandon-Barbosa, D., & Salas-Paramo, J. (2021). Tourism Amidst COVID-19: consumer experience in luxury hotels booked through digital platforms. *Tourism Recreation Research*, 1–6.
- Gao, S., Hao, J., & Fu, Y. (2015). The application and comparison of web services for sentiment analysis in tourism. *2015 12th International Conference on Service Systems and Service Management (ICSSSM)*, 1–6.
- Goslee, S. C., & Urban, D. L. (2007). The ecodist package for dissimilarity-based analysis of ecological data. *Journal of Statistical Software*, 22(7), 1–19.
- Hatzivassiloglou, V., & McKeown, K. (1997). Predicting the semantic orientation of adjectives. *35th Annual Meeting of the Association for Computational Linguistics and 8th Conference of the European Chapter of the Association for Computational Linguistics*, 174–181.
- Joo, Y.-H., Kim, Y., & Yang, S.-J. (2011). Valuing customers for social network services. *Journal of Business Research*, 64(11), 1239–1244.
- Kennedy, A., & Inkpen, D. (2006). Sentiment classification of movie reviews using contextual valence shifters. *Computational Intelligence*, 22(2), 110–125.
- Ko, C.-H. (2021). Exploring Social Media Management On Hotel Performance. *International Journal of Organizational Innovation*, 14(2).
- Li, H., Ye, Q., & 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.
- Liu, B. (2012). Sentiment analysis and opinion mining. *Synthesis Lectures on Human Language Technologies*, 5(1), 1–167.
- Lu, Y., & Zheng, Q. (2021). Twitter public sentiment dynamics on cruise tourism during the COVID-19 pandemic. *Current Issues in Tourism*, 24(7), 892–898.
- Pang, B., & Lee, L. (2008). Opinion mining and sentiment analysis. *Foundations and Trends® in Information Retrieval*, 2(1–2), 1–135.
- Pang, B., Lee, L., & Vaithyanathan, S. (2002). Thumbs up? Sentiment classification using machine learning techniques. *ArXiv Preprint Cs/0205070*.
- Pejic-Bach, M., Bertoncel, T., Meško, M., & Krstić, Ž. (2019). Text mining of industry 4.0 job advertisements. *International Journal of Information Management*.  
<https://doi.org/https://doi.org/10.1016/j.ijinfomgt.2019.07.014>
- Raut, V. B., & Londhe, D. D. (2014). Opinion mining and summarization of hotel reviews. *2014 International Conference on Computational Intelligence and Communication Networks*, 556–559.
- Sparks, B. A., & Browning, V. (2010). Complaining in cyberspace: The motives and forms of hotel guests’ complaints online. *Journal of Hospitality Marketing & Management*, 19(7), 797–818.
- Torres, E. N., Fu, X., & Lehto, X. (2014). Examining key drivers of customer delight in a hotel experience: A cross-cultural perspective. *International Journal of Hospitality Management*, 36, 255–262.
- Venkatachalam, S. (2017). An exploratory study on the building information modeling adoption in United Arab Emirates municipal projects- current status and challenges. *MATEC Web of Conferences*, 120, 02015.  
<https://doi.org/10.1051/mateconf/201712002015>
- 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.
- Xie, X., Fu, Y., Jin, H., Zhao, Y., & Cao, W. (2019). A novel text mining approach for scholar information extraction from web content in Chinese. *Future Generation Computer Systems*.  
<https://doi.org/https://doi.org/10.1016/j.future.2019.08.033>
- Yue, L., Chen, W., Li, X., Zuo, W., & Yin, M. (2019). A survey of sentiment analysis in social media. *Knowledge and Information Systems*, 60(2), 617–663.