**International Journal of Engineering, Science and Mathematics**
Vol. 7 Issue 8, August 2018,
ISSN: 2320-0294 Impact Factor: 6.765
Journal Homepage: http://www.ijesm.co.in, Email: ijesmj@gmail.com
Double-Blind Peer Reviewed Refereed Open Access International Journal - Included in the International Serial Directories Indexed & Listed at: Ulrich's Periodicals Directory ©, U.S.A., Open J-Gage as well as in Cabell's Directories of Publishing Opportunities, U.S.A

# BUSINESS ANALYTICS IN ONLINE CUSTOMER TEXTUAL REVIEWS

**Sunil Kumar Sharma[1], Dr. K.P. Yadav[2]**
**Department of Computer Science**
**[1,2]Shri Venkateshwara University, Gajraula (Uttar Pradesh)**

***Abstract***

Customers not only buy online, but they also publish evaluations about their experiences on social media, thanks to the rapid advancement of information technology. User-generated content (UGC) can be helpful in gaining insight into the shopping habits of current customers and in influencing the purchasing decisions of potential new ones. It is becoming increasingly common for companies to use business intelligence and analytics to assess customers' user-generated content (UGC) on social media and help their marketing efforts. Using textual customer evaluations from hotel websites, this study seeks to determine how satisfied or dissatisfied guests are with various aspects of their stay. Latent semantic analysis (LSA) is a text mining technique used to uncover the factors that influence customer happiness and discontent with hotel products and services. Our research also examines how customers' impressions of hotel product and service features are affected by factors such as travel purpose, hotel type, hotel star level, and editor recommendations. UGC (user-generated content) can help businesses better understand their customers' needs by bridging customer online textual reviews with customers' impressions..

**Keywords:**customer satisfaction;customer dissatisfaction;online textual reviews

## Introduction

Customers not only buy online, but they also publish evaluations about their experiences on social media, thanks to the rapid advancement of information technology. UGC is a term used to describe content that has been created by individuals using social media sites like Facebook and Twitter (UGC). Customers' perceptions of online shopping risk can be alleviated by the use of user-generated content (UGC)**(Ladhari & Michaud, 2015)**Obtaining client input and making improvements to products and services as a result of it is another important benefit of using user-generated content (UGC). UGC is rapidly being used to help businesses' marketing efforts, therefore the use of "business intelligence and analytics (BI & A)" has become increasingly popular. BI & A has grown increasingly critical to the profitability and operational efficiency of businesses.. Firms can cut marketing costs by gaining a better grasp of client preferences and executing market segmentation through the use of effective BI & A. "A shortfall of 140,000 to

190,000 professionals with advanced analytical skills" is expected in the United States alone by 2018. But user-generated content (UGC), such as textual reviews posted by customers online, is frequently in open form and unrestricted structurally. Firms face difficulties in applying BI & A to UGC, despite the fact that UGC contains more information about customers' purchase experiences than ratings. It's difficult to deal with unstructured texts in big data analytics**(Gandomi & Haider, 2015)**. In texts, customers' feelings can be revealed through the use of sentiment analysis. Our study focuses on customer contentment and dissatisfaction in order to better understand customer online reviews.

Customer satisfaction and discontent are frequently studied in the literature but few research focus on customer happiness and dissatisfaction based on online textual customer reviews. Descriptive approaches like "frequency analysis and content analysis" are frequently employed in these few investigations. Therefore, additional research is needed to better understand how to use UGC to acquire customer insights and bridge the gap between customer satisfaction/dissatisfaction and UGC. Due to the dearth of widely accepted ways for successfully analysing and using social media data, BI and A is still a challenge. As a result, the usefulness of social media data for marketing purposes has not been fully explored. Using text mining, we hope to fill this need by analysing social media data. Our method uses LSA and regression to connect consumers' online textual evaluations with customer perceptions, resulting in crucial marketing insights for business practitioners. Customers' happiness and dissatisfaction with specific product and service features is the subject of our research..

## 2. Methodology

Identifying and quantifying the degree to which a client is satisfied or unsatisfied with the most notable hotel features connected with customer satisfaction and dissatisfaction are two of the study's goals;gap analysis**(Parasuraman et al., 1991)**linear regression conjoint analysis**(Danaher, 1997)**and content analysis have been employed in the past to attain these goals. When it comes to collecting data, these approaches rely on self-completion questionnaires or client letters. They require researchers to define service qualities based on their own personal experiences with the product or service. Instead, the data in this study comes from a massive volume of online text reviews. LSA and text regression are used to examine the online review data. The next sections go into great detail on how we went about gathering and analysing the data..

### 2.1. Data collection

Our data came from "booking.com," which requires users to post both positive and negative ratings, albeit in separate sections. We may use this to detect factors of consumer contentment and unhappiness based on positive reviews and negative reviews. Positive and negative ratings are displayed side by side because we believe this will help customers overcome the tendency to read the bad ones first. Customer feedback can be both positive and negative as shown in the Fig. Booking.com also collects information about customers' travel goals (business vs. pleasure) as seen in Fig.

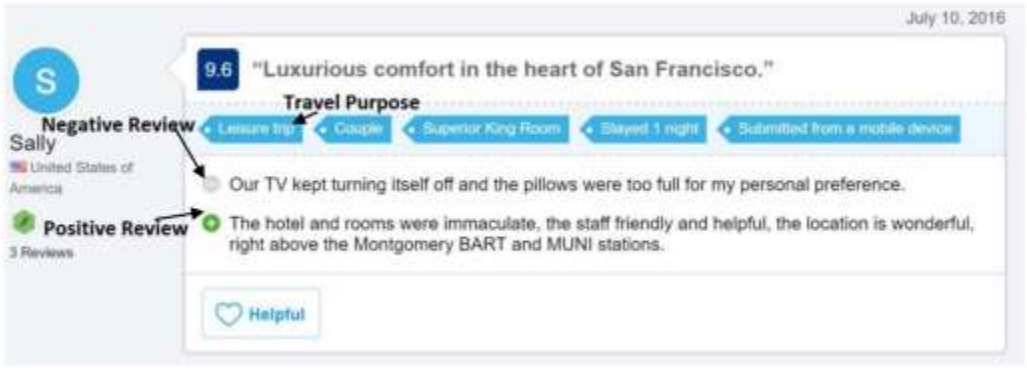


Figure customer online textual review webpage on Booking.com

Based on (Xiang et al., 2015) we collected online customer reviews from hotels in the 100 most populous US cities reported by the "US Census Bureau". Among the 96 cities on the list are "New York, Los Angeles, Chicago, and Houston". The hotels in each city were whittled down based on their star ratings, which we calculated (star level from 0 to 5). Then, for each star level, we created two random numbers, a and b, ranging from 1 to 20, and collected review samples a and b from hotels listed on booking.com based on their review position.. Using this method six times for each hotel's star level is recommended. It was impossible for us to add any reviews that were incomplete. In the end, we collected 3596 bits of data. As seen in this table, client demographics have been taken into account. The data reveals the country and purpose of the customers' journeys. To be included in our list of nationalities, we must have at least 10 consumers who gave us their opinions on the product or service. There are 20 countries represented in Table 1, with the majority of our samples from the United States. Customers were vacationing four times as often as they were going to work.

| Nationality | Number | Percentage | Nationality | Number | Percentage |
|---|---|---|---|---|---|
| United States | 2808 | 78.09% | Brazil | 14 | 0.39% |
| UK | 134 | 3.73% | Switzerland | 14 | 0.39% |
| Canada | 113 | 3.14% | Ireland | 13 | 0.36% |
| Australia | 77 | 2.14% | Hong Kong | 12 | 0.33% |
| Germany | 32 | 0.89% | China | 11 | 0.31% |
| Netherlands | 23 | 0.64% | France | 11 | 0.31% |
| Saudi Arabia | 23 | 0.64% | UAE | 10 | 0.28% |
| Italy | 20 | 0.56% | Other | 198 | 5.51% |
| New Zealand | 20 | 0.56% | | | |
| Japan | 18 | 0.50% | **Travel Purpose** | **Number** | **Percentage** |
| Mexico | 15 | 0.42% | Leisure | 2876 | 79.98% |
| Nigeria | 15 | 0.42% | Business | 720 | 20.02% |
| Spain | 15 | 0.42% | | | |
| | | | **Total Samples** | 3596 | 100% |

Table 1 Customer demographic information

### 3. Theoretical implications

Our research sheds light on how satisfied or dissatisfied hotel guests are with various aspects of their stay. We looked at client satisfaction and discontent separately, rather than just looking at overall satisfaction. Our research, which makes use of LSA, has been able to pinpoint the factors that influence customer happiness and discontent. A pleasant stay is greatly influenced by the following factors: helpful and friendly staff, a pleasant and convenient location, and a reasonable price. A bad experience is largely influenced by these factors as well: a poor value, an unpleasant and dirty room, an unhelpful and inefficient staff, problems with amenities and facilities, and operational problems. As a result, our study is able to determine the most essential product and service factors that influence consumer happiness and discontent in the hospitality industry and prevent ignoring important attributes. Customers' actual experiences are linked to their online textual reviews for our study. These findings add to the existing body of knowledge about customer happiness, and other research, such as surveys, may help to confirm them in the future..

## 4Conclusions

Customer satisfaction with hotel products and services is influenced by a variety of factors, including the type of vacation taken, the hotel's star level, and the suggestion of a travel editor. Customer happiness is influenced by a variety of factors, including the staff, the accommodation, the location, and the value. Client discontent, on the other hand, is influenced by all five of these crucial factors, which are listed above. Based on the hotel's rating, there are significant disparities in client satisfaction and dissatisfaction with many areas of the product and service offered. According to our research, by using text mining and regression to analyse online consumer textual assessments, business managers may be able to gain a better understanding of their consumers' perspectives.

Reference

Danaher, P. J. (1997). Using conjoint analysis to determine the relative importance of service attributes measured in customer satisfaction surveys. *Journal of Retailing*, *73*(2), 235–260. https://doi.org/https://doi.org/10.1016/S0022-4359(97)90005-1

Gandomi, A., & Haider, M. (2015). Beyond the hype: Big data concepts, methods, and analytics. *International Journal of Information Management*, *35*(2), 137–144.

Gössling, S., & Lane, B. (2015). Rural tourism and the development of Internet-based accommodation booking platforms: a study in the advantages, dangers and implications of innovation. *Journal of Sustainable Tourism*, *23*(8–9), 1386–1403. https://doi.org/10.1080/09669582.2014.909448

Ladhari, R., & Michaud, M. (2015). eWOM effects on hotel booking intentions, attitudes, trust, and website perceptions. *International Journal of Hospitality Management*, *46*, 36–45. https://doi.org/https://doi.org/10.1016/j.ijhm.2015.01.010

Parasuraman, A., Zeithaml, V. A., & Berry, L. L. (1991). Refinement and Reassement of the Servqual scale (pp. 420–450). *Estados Unidos*.

Xiang, Z., Schwartz, Z., Gerdes, 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/https://doi.org/10.1016/j.ijhm.2014.10.013