DOI:
10.1016/j.tourman.2016.09.009

Document Version
Peer reviewed version

Link to publication record in King's Research Portal

Citation for published version (APA):
Guo, Y., Barnes, S. J., & Jia, Q. (2017). Mining meaning from online ratings and reviews: Tourist satisfaction analysis using latent dirichlet allocation. *TOURISM MANAGEMENT*, *59*, 467-483.
https://doi.org/10.1016/j.tourman.2016.09.009

Mining Meaning from Online Ratings and Reviews: Tourist Satisfaction Analysis Using Latent Dirichlet Allocation

Yue Guo
Hohai University
Nanjing City, Jiang Su Province
China 21110
Tel:+86-15295781861
Email. yueggcn@aliyun.com

Stuart J. Barnes*
School of Management and Business
King's College London
Franklin-Wilkins Building, 150 Stamford Street
London SE1 9NH, United Kingdom
Tel: +44 (0)20 7848 3327; Fax: +44 (0)20 7848 4639
Email: stuart.barnes@kcl.ac.uk

Qiong Jia
Hohai University
Nanjing City, Jiang Su Province
China 21110
Tel:+86-15268514342
E-mail: jiaqionghit@163.com

(*) corresponding author.

This paper is a post-print (final draft post-refereeing) of:
Guo, Y., Barnes, S.J., and Jia, Q. (2016). Mining meaning from online ratings and reviews: Tourist satisfaction analysis using latent dirichlet allocation, *Tourism Management*, 59, 467-483.

The publisher's version is available at:
http://www.sciencedirect.com/science/article/pii/S0261517716301698

Mining Meaning from Online Ratings and Reviews: Tourist Satisfaction Analysis Using Latent Dirichlet Allocation

Abstract

Consumer-generated content has provided an important new information medium for tourists, throughout the purchasing lifecycle, transforming the way that visitors evaluate, select and share experiences about tourism. Research in this area has largely focused on quantitative ratings provided on websites. However, advanced techniques for linguistic analysis provide the opportunity to extract meaning from the valuable comments provided by visitors. In this paper, we identify the key dimensions of customer service voiced by hotel visitors use a data mining approach, latent dirichlet analysis (LDA). The big data set includes 266,544 online reviews for 25,670 hotels located in 16 countries. LDA uncovers 19 controllable dimensions that are key for hotels to manage their interactions with visitors. We also find differences according to demographic segments. Perceptual mapping further identifies the most important dimensions according to the star-rating of hotels. We conclude with the implications of our study for future research and practice.

Keywords: Online reviews; visitor satisfaction; data mining; latent dirichlet analysis; perceptual mapping.

1.0. Introduction

Prior studies in marketing and consumer behavior have defined customer satisfaction as a customer's subjective evaluation of a service or product provided based on expectations and actual performance (Anderson et al 1994; Oliver 1980; Woodruff et al.1983). Customers evaluate the degree of satisfaction based on their perceptions of the attributes of hotels that they deem most important. In other words, these attributes of hotels represent dimensions of satisfaction. Numerous studies have proposed that customer satisfaction plays an important role in motivating customers' behavioral loyalty, such as giving positive reviews, returning, or making a recommendation (e.g., Hallowell 1996; Kim et al. 2009a; Hui et al. 2007). Multiple factors contribute to the formation of consumer satisfaction (e.g., price, service quality, and product quality), and it is thus a multidimensional construct consisting of different aspects or sub-constructs, in a similar manner to service quality (Klein and Leffler 1981; Mitra and Golder 2006; Tellis and Johnson 2007). Prior studies typically rely on traditional qualitative, quantitative or mixed methods (e.g., questionnaire survey and focus groups) to identify the dimensions of satisfaction, and subsequently develop empirical measurement scales. These traditional research methods require researchers to seek an effective trade-off between the cost of sample collection and estimation performance. Existing studies on satisfaction are empirically examined based on limited samples during a specific period. Moreover, initial measurement items and survey questions tend to be developed based on the knowledge of researchers on related industries (e.g., the hospitality industry). At a result, inconsistent measurement items and questions are often created and used in prior studies (Barsky 1992; Cronin et al. 2000; Danaher and Haddrell 1996; Fornell 1992).

The Internet has fostered a rapid rise in user-generated content (UGC), particularly alongside the widespread diffusion of Web 2.0 technologies. Tourists can now share their experiences and give specific suggestions to others for hotels, restaurants and attractions (e.g., via comments on customer service, car parking and cleanliness) (Sotiriadis and Van Zyl 2013, Sparks and Browning 2011, Vermeulen and Seegers 2009, Ye et al. 2009a, Ye et al. 2009b,). Extant literature has shown that online customer reviews can be used as a major information source for researchers and practitioners that can help in correctly understanding consumer preferences and demand: for example, to predict financial performance or attempt to increase sales (e.g., Chevalier and Mayzlin 2006; Clemons et al. 2006; Liu 2006; Chau and Xu 2012; Ghose, and Ipeirotis 2011; Ye et al. 2011). Online customer reviews can empower individuals to bypass unclear and inaccurate product or service descriptions and rely directly on the first-hand usage experiences of other consumers, particularly in the case of high involvement products (e.g., vehicles). Moreover, some firms actively encourage their customers to submit online opinions about their products or services, e.g., by offering vouchers or discounts.

UGC may be considered as spontaneous, insightful and passionate feedback provided by consumers that is widely available, free or low cost, and easily accessible anywhere, anytime. Large volumes of data, as represented in the continuous stream of UGC over time, provides practical input (i.e., know-how, know-what) to augment traditional research methods for identifying important issues. Latent dimensions, such as social status, are variables that consumers may not explicitly mention, but that capture or represent a large number of attributes, often indirectly from other indicators (e.g., income and occupation). As a consequence, in the past decade, there have been an increasing number of studies examining the phenomenon of online consumer reviews (e.g., Clemons et al. 2006; Dellarocas et al. 2007; Ho-Dac et al., 2013). Specifically, as for the hospitality and tourism field, we propose that UGC provides a rich source of data to extract the dimensions of customer satisfaction. Hundreds of thousands of community members may contribute to creating online content, thereby creating the “wisdom of crowds” (Surowiecki 2005). Thus, UGC can serve as a useful source of information for enterprises that care about consumers’ demands, particularly in the hospitality industry (e.g. hotels and restaurants). Appendix A provides a brief summary of the recent empirical literature on UGC from multiple academic disciplines, including travel and tourism, marketing, and information systems, since research on UGC in tourism is a recent phenomenon (e.g., Li et al. 2013; Liu and Park 2015; Park and Nicolau 2015). Our review includes influential studies and is intended to be representative rather than exhaustive. Despite the growing significance of online reviews, and in the face of concerns voiced about them, we argue that the extant literature mainly focuses upon examining the impact of online ratings; thus, the literature largely ignores online reviews, which we consider potentially more valuable to academics and practitioners. For example, most prior tourism studies have principally focused on the impact of online ratings on hotel sales (e.g., Xie et al., 2011; Ye et al., 2009; Ye et al., 2011). Compared to online ratings, which are numerical and easily understood, online reviews are text-based and often comprise of large information repositories beyond the analytical capabilities of traditional econometric and statistical methods. Thus far, there is limited empirical evidence from large-scale online reviews to help in understanding consumer satisfaction and its antecedents.

An emerging stream of research in the hospitality industry has attempted to understand customer satisfaction from the content of online reviews. For example, Li et al. (2013) exploit 42,886 online reviews of 774 star-rated hotels in Beijing and use content analysis to identity the determinants of satisfaction. Lu and Stepchenkova (2012) analyzed 373 reviews extracted from TripAdvisor based on content analysis to identity satisfaction attributes. Levy et al. (2003) carry out a content analysis of complaints within one-star online hotel reviews from ten popular review websites to understand customer satisfaction. They content analyzed 225 managerial responses to these one-star reviews. Relative to these prior studies, the proposed framework in the current study differs in three major ways. First, in this research, we attempt investigate online reviews and online ratings together. On the one hand, more effort has to be put into the analysis of large-scale online review contents in order to explore fully the potential of the data to identify the antecedents of satisfaction. The study uses topic modelling: advanced software and mathematical techniques developed in the fields of natural language processing and data mining. On the other hand, a stepwise regression analysis of a large volume of numerical ratings is carried out to verify the validity of important dimensions proposed by prior empirical studies on satisfaction of hotel customers. These studies have principally used questionnaire surveys or experiments to understand the antecedents of satisfaction (e.g., Min et al., 2015; Wu and Liang, 2009). Such a combined analysis of online ratings and online reviews enables us to further link specific dimensions extracted from large textual contents with abstract-level factors identified by prior studies, such as service quality. Second, LDA uses an unsupervised Bayesian learning algorithm to capture effectively context-specific dimensions and does not make any assumption about the distribution of online reviews or grammatical attributes of language. Consequently, relative to prior methods for text analysis, LDA can complete many steps of the textual analysis with little human intervention, even labeling dimensions, and is more suitable for dealing with large and unstructured online reviews, thus creating meanings that are more realistic. Finally, this study demonstrates the method on a relatively broad sample of more than 200,000 online reviews of hotels located in more than 100 cities in 16 countries, which enables us to make more reliable generalizations than prior studies.

In sum, an important contribution of this research is that we empirically develop and identify the dimensions of satisfaction based on big data from UGC including numerical and textual information. Thus, the dimensions from the data provide a genuine "voice of the customer" (Griffin and Hauser 1993), generally understood as:

> "a complete set of customer wants and needs; expressed in the customer's own language; organized the way the customer thinks about, uses and interacts with the product and service; and prioritized by the customer in terms of both importance and performance – in other words current satisfaction with existing alternatives" (Katz 2011, p. 34).

In this study, we attempt to mine the sensitive and important factors influencing consumer satisfaction through UGC. By extracting value from UGC, we believe we will be able to hear the voice of the customer more correctly and effectively, providing practical help for business owners and investors.

Specifically, the goal of this study is to answer the following questions:

1. What are the key dimensions of customer satisfaction expressed in UGC?
2. How valid are the new dimensions?
3 What is the heterogeneity of perceptions for different groups of consumers across these dimensions?
4 What are the most important aspects influencing consumer satisfaction based on regression analysis using the large sample?
5 How do these dimensions vary across star-rated hotels?

The paper is organized as follows. We first present the method used in our study, including sampling and text mining techniques. The findings are then discussed, followed by the perceptual mapping of results by hotel star rating. Finally, the paper rounds off with a discussion of limitations and implications for research and practice.

2.0. Method

We provide a unified framework to extract latent dimensions from rich user-generated data, as shown in Figure 1. This summarizes the process used in our study. This provides a practical guideline for researchers to extract the latent dimensions of multidimensional constructs in other social science studies.

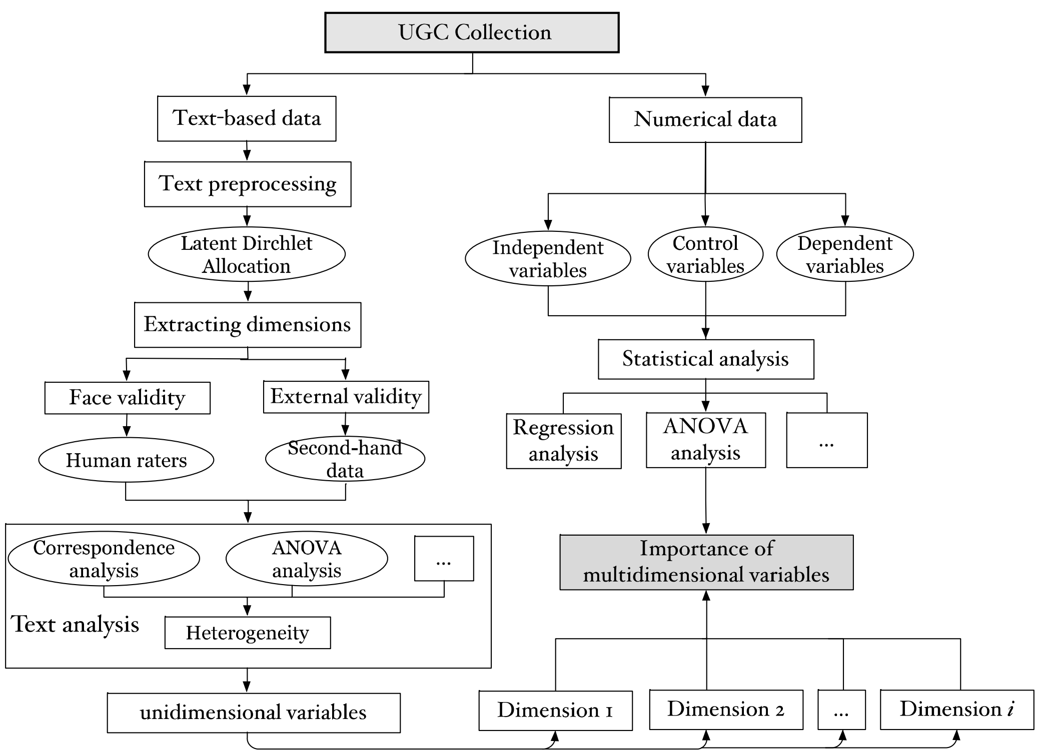

Figure 1. Framework for extracting latent dimensions based on UGC

2.1. Sampling

In this study, our target population is consumers within the hospitality industry, particularly in the hotel sector. The hospitality industry has a strong focus upon achieving customer satisfaction, including in sectors such as hotels and resorts, cruise lines, airlines and other various forms of travel, tourism, special event planning, and restaurants. Our empirical setting is the consumer review website, TripAdvisor, which is one of the largest online review communities of travel consumers and an early adopter of UGC. TripAdvisor began in 2000, and contains unbiased traveler reviews for hotels, restaurants and attractions. For a more comprehensive description of TripAdvisor, see O'Connor (2008). We utilized automated techniques for data collection and analysis. Reviews within TripAdvisor are collected at a disaggregate level (i.e., individual reviews) and it was thus possible to parse individual reviews of the website and aggregate them for our analysis. We developed a web crawler to collect data from TripAdvisor periodically. The data collection process ran until we had in excess of 250,000 reviews, lasting 3 months. In addition to text-based review contents, we also extracted and stored numerical data, such as ratings, in numeric format. Our Web crawler first visits TripAdvisor and automatically creates a list of hotel URLs to visit, called the seeds. As the crawler visits these URLs, it identifies and collects all the required information in each hotel introduction page. In total, 266,544 hotel reviews created by 39,287 unique reviewers were downloaded in June 2015 and saved in CSV file format.

Reviewers in TripAdvisor may choose whether to disclose publicly their demographic information (see Table 1). Around 30% of reviewers disclosed their age, split into the following bands: 13-17, 18-24, 25-34, 35-49, 50-64, and 65+. The median disclosed age was in the 35-49 band. Some reviewers also choose to display their gender, which accounted for around 32% of the total sample – 16% each of male and female. In addition, TripAdvisor encourages users to evaluate the "helpfulness" of reviews and sorts reviews using helpfulness votes as a default setting. As part of the review process, reviewers are able to allocated star ratings to hotels (one to five stars) in five specific aspects including hotel location, cleanliness, hotel room experience, service quality, value of money, as well as an overall rating of satisfaction (see Table 2). Overall, the reviews collected for this study involved 25,670 hotels located in 16 countries.

	Frequency	***%***
Gender		
Male	42541	15.96%
Female	43322	16.25%
Undisclosed	180681	67.79%
Age		
13-17	59	0.02%
18-24	2958	1.11%
25-34	16516	6.20%
35-49	28429	10.67%
50-64	26218	9.84%
65	6147	2.30%
Undisclosed	186217	69.86%

	Frequency	***%***
Helpfulness votes		
<=20	198499	74.47%
21-40	35295	13.24%
41-60	14208	5.33%
61-80	7006	2.63%
81-100	3758	1.41%
101+	7778	2.92%
Rating number		
0	990	0.37%
1-20	188599	70.76%

21-40	34860	13.08%	81-100	5305	1.99%
41-60	15674	5.88%	101+	12051	4.52%
61-80	9065	3.4%			

Table 1. Characteristics of reviewers in the sample (n=266,544)

2.2. Text pre-processing

Text pre-processing used steps very similar to those adopted in prior studies (e.g., Tirunillai and Tellis 2012; Lee and Bradlow 2006), including eliminating non-English characters and words, word text tokenization, part-of-speech tagging (POS tagging or POST), replacing common negative words, word stemming, and removing low frequency words (below 2%). For example, an original review appeared as:

"I stayed at the Residence Inn on the weekend of March 13-15. The overall look of the property and room was very nice. However, housekeeping did not do a good job at cleaning our room, as soon as I checked in my bathroom shower had someone else's hair, and the floor was dirty. Also, the microwave had crumbs."

After pre-processing of the text, the review became:

"*stay property room nice do not do a good job cleaning room check bathroom shower have someone hair, floor dirty microwave have crumbs.*"

We implemented text pre-processing by using modules of the Natural Language Toolkit (www.nltk.org) in the Python programming environment.

Rating score	Overall	Cleanliness	Service	Location	Room	Value
1 star	12907 (4.84%)	6389 (2.40%)	12431 (4.67%)	2796 (1.05%)	7350 (2.76%)	8062 (3.02%)
2 stars	12474 (4.68%)	5084 (1.91%)	9227 (3.46%)	3076 (1.15%)	6997 (2.63%)	6892 (2.59%)
3 stars	29337 (11.01%)	13490 (5.06%)	25228 (9.46%)	14374 (5.39%)	20772 (7.79%)	18501 (6.94%)
4 stars	73584 (%27.61)	33122 (12.43%)	53648 (20.13%)	36527 (13.70%)	39629 (14.87%)	38793 (14.55%)
5 stars	138241 (51.86%)	93464 (35.12%)	162772 (61.07%)	93202 (34.97%)	73452 (27.55%)	78509 (29.45%)
No stars	1 (≈0%)	114995 (43.08%)	3238 (1.21%)	116566 (43.74%)	118344 (44.40%)	115787 (43.45%)

Table 2. Frequencies for review ratings (n=266,544)

2.3. Dimension extraction

Based on the large sample of online reviews collected, the primary contribution of this research is the effective extraction of potential dimensions influencing customer satisfaction. The basic principle of extracting dimensions from online reviews is analogous to dimensionality reduction methods (e.g., principal component analysis) that shrink the number of random variables under consideration: such traditional methods are unfeasible in the big data analysis context owing to several reasons. First, online reviews consist of large numbers of different words voiced by different customers; the text is therefore highly characteristic of “the long tail” (Anderson 2008), as shown in Figure 2. In total, 9384 words appear in the reviews. Moreover, consumers tend to give specific comments based on their personal hotel experiences and preferences. As a result, consumers tend to review only the few dimensions that they are most concerned about. Consequently, the process of dimension extraction is quite challenging, since useful information on the dimensions of customer satisfaction appears extremely scant. Thus, standard dimension reduction methods used in social science are not applicable.

This research utilizes recent developments in topic model technologies and techniques, as used in the fields of machine learning and natural language processing, to effectively extract dimensions of customer satisfaction from a large corpus of text data. A topic model is a type of probability model for discovering the abstract "topics" that occur in a collection of documents. Latent dirichlet allocation (LDA) (Blei et al. 2003) is the most common method for topic modelling and is a generalization of probabilistic latent semantic indexing (PLSI) (Hofmann 1999). In essence, LDA assumes that the words in each review are independently drawn from a mixture of baskets, each containing a set of words. Each basket contains words taken from the vocabulary, a topic-word distribution. Topics can potentially be shared by all reviews. Every review will have its own mixing proportion of topics. Topic modelling using LDA enables the discovery of underlying topics from massive volumes of unstructured text data – big data. By relying on the LDA approach, we can quickly discover a mixture of topics (i.e., aspects influencing hotel customers’ satisfaction) from huge numbers of documents (i.e., reviews), and a series of specific purposes can be achieved, including identifying an optimum number of dimensions, labelling the dimensions, and assessing the heterogeneity and relative importance of dimensions according to different reviewer characteristics (e.g., gender and age).

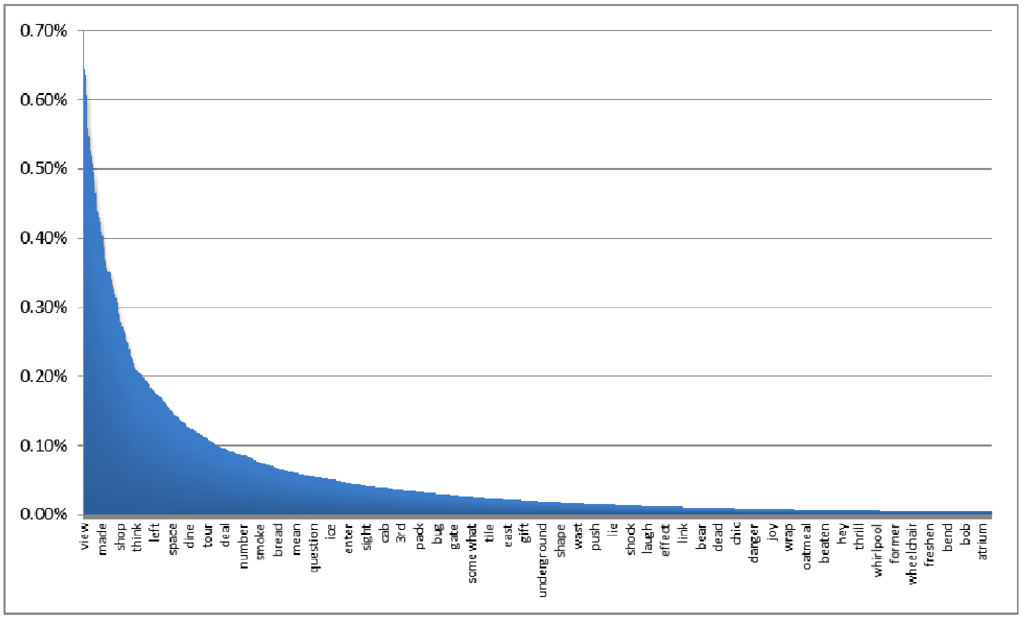

Figure 2. Word distribution (9,384 words) for hotel reviews (n=266,544)

As an unsupervised method, LDA is highly efficient, because it can be adapted to handling both big data and highly disaggregated time periods with sparse data (Blei et al. 2003). LDA is utilized to extract the dimensions of customer satisfaction, the importance of different dimensions, and the words related to the dimensions based on preprocessed comments. Following Tirunillai and Tellis (2014), we define a "dimension" as a latent construct distributed over a vocabulary of words that consumers use to describe their hotel experience, also referred to as a "topic" in the LDA literature. It is assumed that a sequence of N words constitutes a review, which is referred to as a "document" in the literature, $w=(w_1,w_2,\ldots,w_N)$, whilst M reviews form a corpus, $D=\{w_1,w_2,\ldots,w_M\}$. We also assume that there are K number of dimensions across the corpus comprising of all the M reviews in a given time period. LDA is a generative probabilistic model of a corpus, and reviews are represented as random mixtures over K latent dimensions, where each dimension is characterized by a distribution of words. In other words, consumers' reviews represent the K different dimensions with probabilities. For instance, a consumer may give his or her textual review based on personal opinions signifying 30% for inexpensive, 35% for cleanliness, and 35% for car parking space.

Similar to a hierarchical Bayesian model, the LDA model consists of three hierarchies, as shown in Figure 3 (with plate notation). The shadow circle w represents observable variables, while non-shadow circles, z and θ refer to latent variables. The boxes depicted as plates represent replications. The outer plate refers to documents, and the inner plate refers to repeatedly choosing latent topics and words within a document. The parameters α and β are hyper-parameters at the corpus level – which are assumed to be sampled once.

Based on these definitions, LDA assumes the following generative process for each review in a corpus, *D* (Blei et al. 2003):

1) Choose $N \sim$ Poisson (ζ), N represents the length of documents;
2) Choose $\theta \sim$ Dir (α), where α is the parameter of the Dirichlet prior on the per-review topic distributions; and
3) For each of N words w_n:
    (a) Choose a topic $z_n \sim$ Multinomial (θ); and
    (b) Choose a word w_n from $p(w_n|z_n, \beta)$, a multinomial probability conditioned on the topic z_n.

The Poisson distribution in step 1 shows the length of the reviews distributed in each document. In step 2, the k^{th} dimension's importance, as perceived by consumers, which is the probability that the dimension (topic) occurs, can be represented as a k-dimensional dirichlet random variable θ in a given review according to the probability density function as follows:

$$p(\theta|\alpha) = \frac{\Gamma(\sum_{i=1}^{k} \alpha_i)}{\prod_{i=1}^{k} \Gamma(\alpha_i)} \prod_{i=1}^{K} {\theta_i}^{\alpha_i - 1} \tag{1}$$

Equation (1) follows the probability density function of Dirichlet distribution, which is a conjugate prior for the categorical and multinomial distribution so that Equation (2) can be operated. Here the variable θ lies in the (k-1)-simplex $\theta_1, \ldots. \theta_{i\ 1} > 0$, $\sum_{i=1}^{K-1} \theta_i < 1$ $\sum_{i=1}^{K} \theta_i = 1$, where parameter α is a k-vector with components α_i>0. $\Gamma(x)$ is the Gamma function. In step 3, parameter β is a k-dimensional dirichlet random variable in a given topic. The joint distribution of a topic mixture θ, a set of N topics z, and a set of N words w can be obtained by:

$$p(\theta, z, w|\alpha, \beta) = p(\theta|\alpha) \prod_{n=1}^{N} p(z_n|\theta) p(w_n|z_n, \beta) \tag{2}$$

In equation (2), $p(\theta|\alpha)$ corresponds to review-level (document-level) parameters, whilst the other terms correspond to word-level parameters. Here, $p(z_n|\theta)$ is θ_i for the unique i where $z_n^i = 1$. We can attain the marginal distribution of a review, which is the likelihood of a review, via equation (3):

$$p(w|\alpha, \beta) = \int p(\theta|\alpha) \left(\prod_{n=1}^{N} \sum_{z_n} p(z_n|\theta) p(w_n|z_n, \beta) \right) d\theta \tag{3}$$

Subsequently, we can obtain the probability of a corpus by taking the product of the marginal probabilities of single review:

$$p(D|\alpha,\beta) = \prod_{d=1}^{M} \int p(\theta_d|\alpha) \left(\prod_{n=1}^{N_d} \sum_{z_{d_n}} p(z_{d_n}|\theta_d) p(w_{d_n}|z_{d_n},\beta) \right) d\theta_d \tag{4}$$

In the LDA model, the parameters α and β need to be inferred by learning using methods such as the variation expectation maximization algorithm (Blei et al. 2003), Gibbs sampling methods (Griffiths and Steyvers 2004), or the maximum likelihood estimate method (Asuncion et al. 2009). The term θ_d is a document-level variable assumed to be sampled once per review. The variables z_{d_n} and w_{d_n} are word-level variables, sampled once for each word in each document or review. Following the above mentioned LDA processes, we can extract the final topics that customers are most concerned about and the probability of each topic occurring in customer reviews.

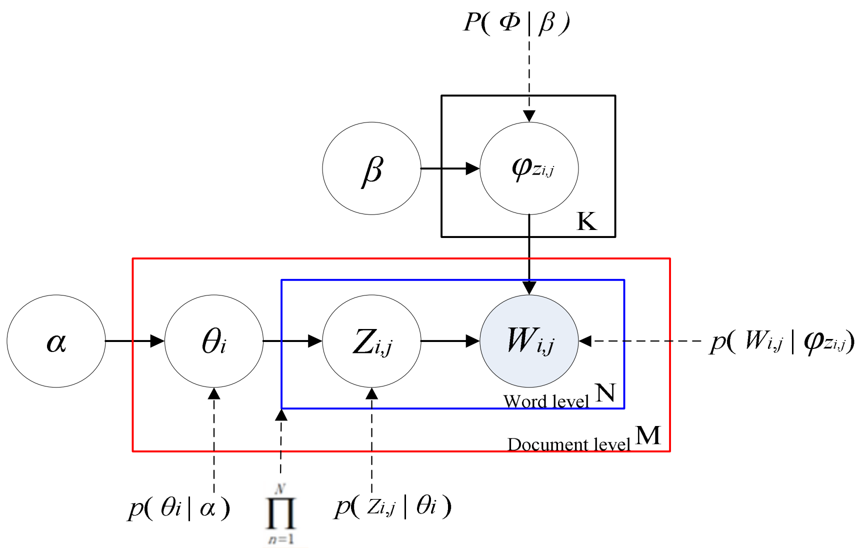

Figure 3. LDA model with plate notation

The LDA model does not make any assumption about the structure of text or the syntactical or grammatical properties of the language. We adopt the LDA model instead of other techniques of text analysis found in the literature based on the following grounds. First, the LDA model excels at efficiently analyzing large-scale data at a highly granular level and thus enables us to explore the heterogeneity of dimensions within different customer groups (e.g., male vs. female). In addition, LDA helps us to compute the practical frequency of occurrence of each extracted dimension based on its intensity in hotel reviews. For instance, tourists select words from their own vocabulary to express personal opinions on different aspects of hotels, such as location, facilities, price, and so on. These topics representing the important aspects related to traveler satisfaction have a distribution across the reviews depending on their frequency of occurrence associated with consumers' own hotel experiences. Latent dirichlet allocation is one of the core models in the "topic models" family (Blei 2012) and is flexible enough to enable us to undertake rich analyses. This research used the Stanford Topic Modelling Toolbox version 0.4.0 to extract dimensions using LDA (Ramage and Rosen 2011).

3.0. Results

In this section, we summarize the results of the extraction of the dimensions of satisfaction. We then examine the validity of these dimensions by comparing the results with prior studies on hotel satisfaction and service quality. Subsequently, we assess the heterogeneity of the dimensions across consumers (reviews) based on the demographic characteristics given in Table 1. A sensitivity analysis of the dimensions of satisfaction is then provided using a stepwise regression model, followed by perceptual mapping of hotel reviews according to dimensions and star-ratings.

3.1. Dimensions of customer satisfaction

We apply the LDA to extract and label the dimensions of customer satisfaction across all collected hotel reviews in our sample. The LDA identified 30 topics and within each topic showed the top-20 words and their relative weight. The naming of dimensions was first conducted by one researcher and then confirmed by a second researcher. Naming was based on the identification of a logical connection between the most frequent words for a topic. For example, in Table 3, the topic name "Car Parking" is based on the word 'park', weighted 14.5%, and 'car', weighted 4.4%, both of which appear at the top of the list. Once identified a candidate topic name was then further tested via logical connection to other words in the top-20 distribution list. If a connection was found, the topic name was retained. If a word was found that did not fit the topic name, the naming process restarted using this information to inform it.

Figure 2 presents the 30 most important dimensions (topics) extracted from 266,544 online reviews for 25,670 hotels located in 16 countries. Two of the dimensions represent consumers' overall perceptions of the level of hotel experience: high standards and satisficing. High standards imply that a consumer perceives his or her hotel experience as superior – the selected hotel maintains an elevated standard. Satisficing, on the other hand, is viewed as a decision-making or cognitive status where a customer's acceptability threshold is just met. Three dimensions represent consumers' responses that are significantly determined by their degree of satisfaction or dissatisfaction, including giving poor reviews, making a recommendation, and return visits. The remaining dimensions show 25 specific aspects of hotel quality (e.g., bathroom problems, poor communication, and dining). We organize these dimensions into three basic categories: controlled, partially controlled and uncontrolled (see Table 4). First, we view a dimension as controlled if it can be significantly improved through the management of a hotel, such as available car parking space and the service quality of hotel staff. Second, uncontrolled factors are those that are considered impossible for hotel managers to improve, such as weather and location. Third, partially uncontrolled aspects occur somewhere in the middle of the continuum, and can be partly improved or solved through effective management approaches, often when creating a hotel, e.g. hotel location and public transport. For example, a hotel may consider offering a pick-up and drop-off service if its location is not convenient for available public transport. We suggest that although the 25 dimensions significantly influence customers' satisfaction, their relative importance will tend to be different for other stakeholders such as hotel operators, owners, and investors. Specifically, hotel operators and management should place an emphasis on leveraging controlled and some partially controlled dimensions. However, hotel owners and investors should carefully consider uncontrolled and partly controlled aspects when planning, designing, and building a new hotel.

Topic	Relative weight	%	Topic	Relative weight	%
Topic 1: Car Parking			***Topic 2: Bathroom***		
park	30702.62	14.5%	shower	22219.07	8.0%
car	9930.66	4.7%	water	13710.02	4.9%
street	8373.58	3.9%	bathroom	12487.22	4.5%
free	8115.33	3.8%	hot	8254.09	3.0%
downtown	6510.97	3.1%	air	5988.69	2.1%
lot	5881.83	2.8%	towel	5766.55	2.1%
across	5356.02	2.5%	work	5337.27	1.9%
close	5050.80	2.4%	small	4795.33	1.7%
right	4790.92	2.3%	use	4675.67	1.7%
center	4351.29	2.1%	bath	4662.39	1.7%
drive	4204.18	2.0%	toilet	3914.35	1.4%
block	4124.77	1.9%	cold	3885.34	1.4%
away	3768.79	1.8%	heat	3852.58	1.4%
shop	3134.07	1.5%	light	3316.73	1.2%
access	2978.80	1.4%	sink	2914.07	1.0%
within	2661.61	1.3%	window	2865.75	1.0%
quiet	2566.23	1.2%	open	2647.17	1.0%
next	2269.43	1.1%	need	2596.46	0.9%
road	2261.28	1.1%	two	2481.19	0.9%
side	2013.30	0.9%	con	2325.11	0.8%

Table 3. An example of identifying dimension labels

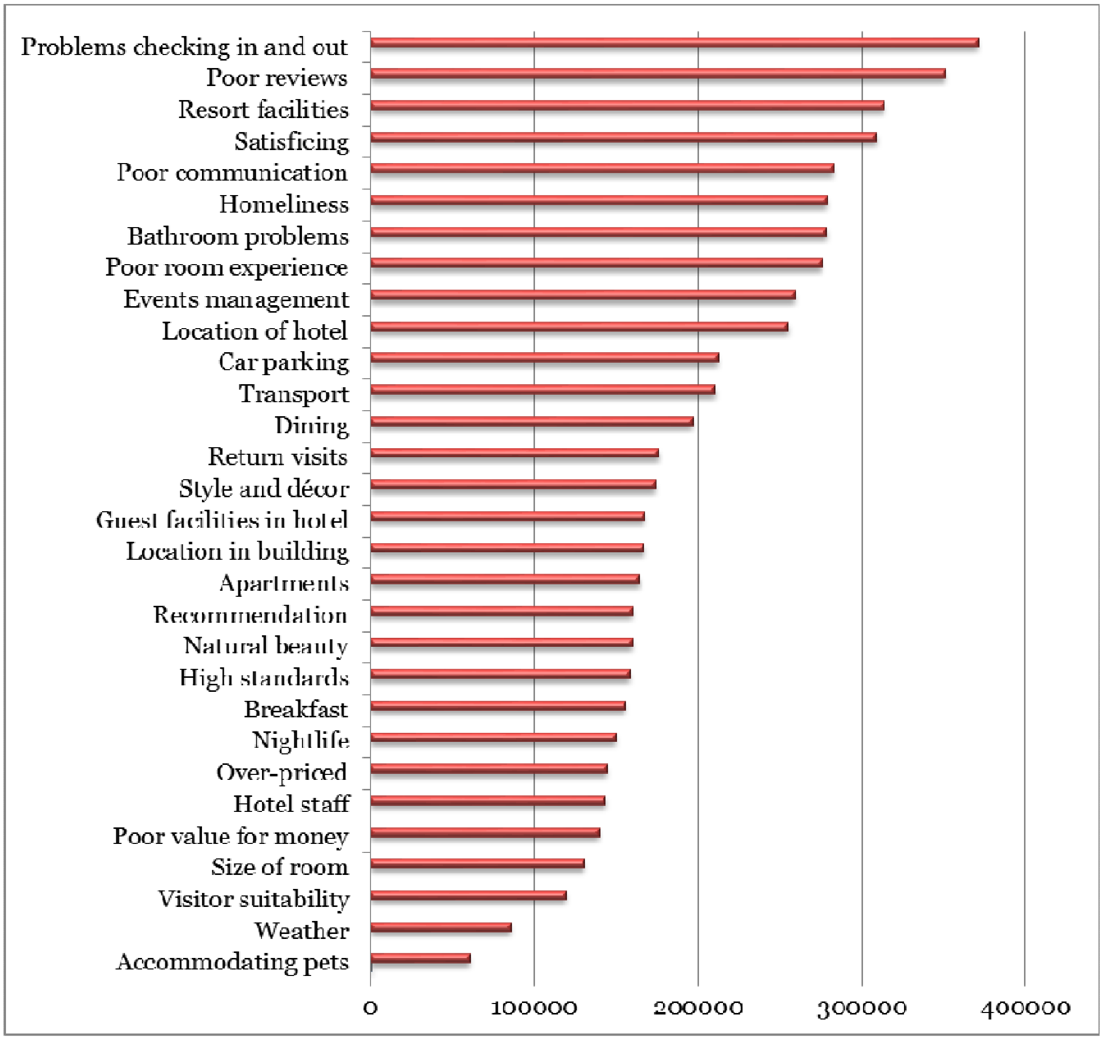

Figure 4. Extracted dimensions

3.2. Comparison of dimensions

We compared the results of the automated analysis with the dimensions derived from prior studies on hotel satisfaction. As mentioned above, prior studies mainly utilize cross-sectional questionnaire surveys and focus groups to identify sets of dimensions that influence the degree of consumer satisfaction, typically limited by the number of respondents (i.e., travelers in this case) due to the cost and effort involved. Online reviews contributed by consumers do not suffer from these restrictions. We assessed the overlap of the dimensions extracted from our automated analysis with the dimensions used in prior studies. The Jaccard coefficient is use to test the degree of overlap between the dimensions proposed by prior research and those extracted in our automated analysis.

If N (Dim_{lda}) represents the set of dimensions derived from online reviews using the LDA model and N(Dim_{pr}) represents the set of dimensions mentioned in prior studies, we calculate the Jaccard coefficient as:

$$JC = \frac{|N(Dim_{lda} \cap Dim_{CR}|}{|N(Dim_{lda} \cup Dim_{CR}|} \quad (5)$$

Controlled dimensions (19)	**Partially controlled dimensions (3)**	**Uncontrolled dimensions (3)**
Checking in and Out	Transport	Weather
Resort Facilities	Location	Natural beauty
Communication	Visitor suitability	Nightlife
Homeliness		
Bathroom		
Room experience		
Events Management		
Car Parking		
Style and Decoration		
Guest facilities in hotel		
Location in building (e.g., first floor)		
Breakfast		
Value for money		
Price		
Staff service		
Room size		
Apartment		
Dining		
Accommodating Pets		

Table 4. A typology of extracted dimensions

The higher the Jaccard coefficient's value, the higher the degree of overlap between the two sets of dimensions. As shown in Table 5, the Jaccard coefficient is 0.58. If we consider only the controlled dimensions, the value of Jaccard coefficient is further increased to 0.65. This implies that our LDA analysis derives new information from the online reviews and identifies new dimensions that have been overlooked by prior studies. We argue that our LDA findings are more reliable for generalization (i.e., external validity) due to a large sample across different regions: as noted earlier, consumers opinions on specific aspects contributing to their satisfaction were obtained from 266,544 online reviews of 25,670 hotels located in 16 countries, which is much larger than the sample size of any prior studies using traditional research methods. It is noted that in some cases, the dimensions proposed by prior studies are more detailed than the results of LDA, and vice versa. In these cases, we use the dimension that is at the most general level for comparison, covering both the details involved in the automated analysis and prior studies. Some dimensions used in prior studies were ignored since they did not relate directly to consumer hotel experience.

Dimensions	LDA analysis	Prior studies
Checking in and Out	√	√
Resort Facilities	√	√
Communication	√	√
Homeliness	√	X
Bathroom	√	√
Room experience	√	√
Events Management	√	X
Car Parking	√	√
Style and Decoration	√	X
Guest facilities	√	√
Location in building (e.g., first floor)	√	X
Breakfast	√	√
Price and value for money	√	√
Staff service	√	√
Room size	√	√
Apartment	√	√
Dining	√	√
Accommodating Pets	√	X
Transport	√	X
Location	√	√
Visitor suitability	√	X
Weather	√	X
Natural beauty	√	X
Nightlife	√	√
Electronic key card	X	√
Security	X	√

Notes: ✓ = included; X = not included. Jaccard coefficient: 0.58.
We compared the dimensions derived from LDA with the following studies: Al Khattab and Aldehayyat (2011), Blesic et al. (2011), Chang (2008), Chi and Qu (2009), Clemens et al. (2010), Emir and Kozak (2011), Fawzy (2010), Gagnon and Roh (2007), Gill et al. (2006), Gunderson et al. (1996), Heung (2000), Kim et al. (2009b), Hsieh and Tsai (2009), Kandampully and Suhartanto (2000), Kang et al. (2004), Kuo (2007), Lau et al. (2005), Law and Yip (2010), Markovic et al.(2010), Mohsin et al.(2011), Mohsin and Lockyer (2010), Nadiri and Hussain (2005), Prayukvong et al. (2007), Ramanathan (2012), Ramanathan and Ramanathan (2011), Ryan and Huimin (2007), Sanchez-Gutierrez et al. (2011), Sim et al. (2006), Skogland and Siguaw (2004), Wang et al. (2008), Alegre and Garau (2010), Yilmaz (2009), and Weng et al. (2012).

Table 5. A comparison of dimensions between LDA analysis and prior studies

We further examined the face validity of our extracted dimensions by comparing the results of our analysis with that of human analysis (see Table 6). Two independent data-mining researchers who have good knowledge of natural language processing and textual analysis were invited to read collected online reviews and then to identify the dimensions mentioned in these reviews. Both researchers were from the same University. They each randomly selected 150 reviews for this purpose.

A t-test of the demographic characteristics of reviewers found no significant differences between the 300 selected reviews and the main sample, including gender ($t = 0.855$, $p > 0.10$), age ($t = 0.124$, $p > 0.10$), and star-rating ($t = 0.397$, $p > 0.10$). Moreover, the first 150 random reviews related to hotels located in 12 countries; the second 150 hotel reviews related to 10 countries. Overall, the 300 random reviews covered all 16 countries. Thus, the face validation reviews should be representative of our main sample. We compared the dimensions derived from the LDA analysis with the dimensions identified by the two researchers to calculate the reliability of the LDA result. The Jaccard coefficient is 0.66 and 0.76 between the automated analysis and the two researchers, A and B respectively. Given the nature of the task and the level of ambiguity the raters faced while identifying topics from reviews, these figures indicate that the LDA approach is feasible and reliable for extracting latent dimensions from online reviews.

Dimensions	**LDA analysis**	**Researcher A**	**Researcher B**
Checking in and Out	√	√	√
Resort Facilities	√	√	√
Communication	√	√	√
Homeliness	√	X	√
Bathroom	√	√	√
Room experience	√	√	√
Events Management	√	X	X
Car Parking	√	√	√
Style and Decoration	√	√	√
Guest facilities	√	√	√
Location in building (e.g., first floor)	√	√	√
Breakfast	√	√	√
Price and value for money	√	√	√
Staff service	√	√	√
Room size	√	√	√
Apartment	√	√	√
Dining	√	√	√
Accommodating Pets	√	X	X
Transport	√	√	√
Location	√	√	√
Visitor suitability	√	X	X
Weather	√	√	X
Natural beauty	√	√	√
Nightlife	√	X	√
Housekeeping	X	√	X
Elevator	X	√	√
Security	X	√	√
Room facilities	X	√	√
External environment	X	√	X
Notes: ✓ = included; X = not included. The Jaccard coefficient is 0.66 between LDA analysis and researcher A. The Jaccard coefficient is 0.76 between LDA analysis and researcher B.			

Table 6. A comparison of dimensions between LDA analysis and human analysis

3.2. The relative importance of dimensions

This section illustrates the relative importance of the dimensions of satisfaction according to consumers' profile characteristics, including gender and age. The importance level is calculated based on the practical frequency of occurrence of each extracted dimension from the reviews of each gender or age group in the hotel review sample. We first assess the heterogeneity of dimensions in terms of the probability distribution of dimensions on any given consumer group (e.g., male versus female). Figure 5 illustrates the different impacts of consumer satisfaction according to gender (where indicated by a reviewer). This is the result of the LDA slicing routine for words associated with a topic, performed in the Stanford Topic Modeling Toolbox (TMT) using the Scala programming language. This analyses the distribution of topics within documents for each gender (slice). Figure 5 combines the results of both slices into a single graph in descending order of distribution for males (to aid interpretation). A further t-test reveals where the significant differences in topics occurs between males and females (based on the combined distribution of topics by respondent). Females had significantly greater mention of homeliness ($F=30.305$, $p<.001$), resort facilities ($F=18.257$, $p<.001$), room experience ($F=7.639$, $p=.006$), guest facilities in hotel ($F=6.726$, $p=.010$), and checking in and out ($F=4.329$, $p=.037$). Males had significantly greater mention of location ($F=8.497$, $p=.004$), communication ($F=5.405$, $p=.020$), and value for money ($F=5.080$, $p=0.024$).

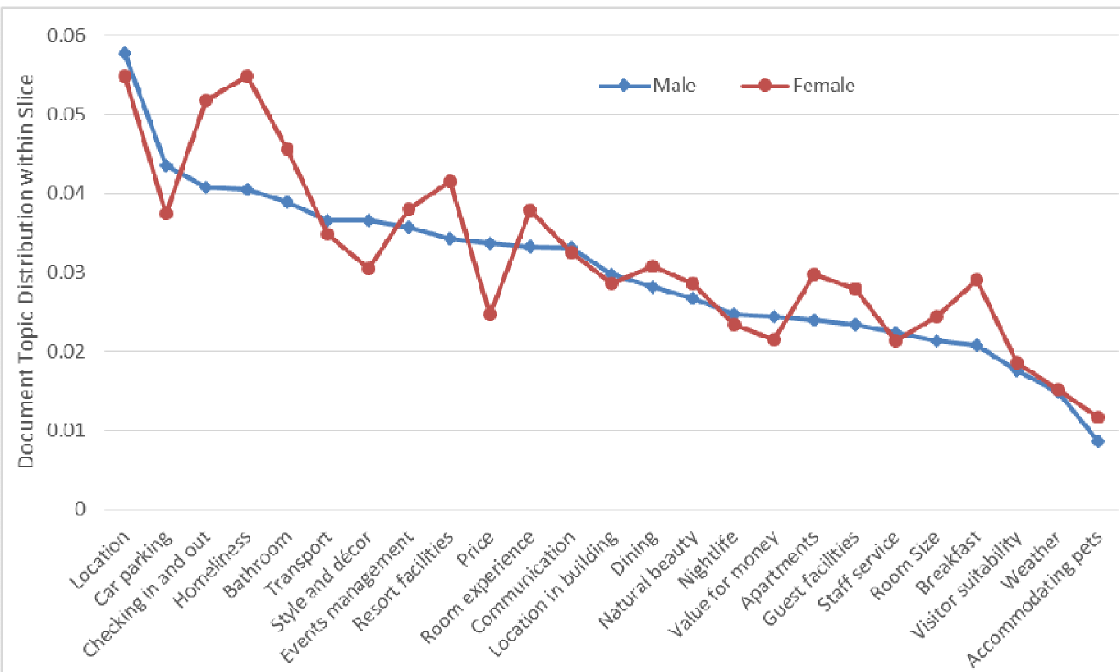

Figure 5. The relative importance of dimensions by gender

Figure 6 examines the profile of consumer satisfaction dimensions according to age (where given by a reviewer). This figure is the result of an LDA slicing routine by document topic distribution within age groups, combining the distribution of topics within each age group into a single graph. The demand for hotel homeliness appears to increase sharply with increasing age. However, room experience appears to decrease according to age group. Differences between document topic distributions by age group were further evaluated using the Kruskall-Wallis test (since the age data is ordinal not interval). This uses the document topic distribution data by respondent rather than by slice. The results show a significant difference in mentions of homeliness (H=17.233, df=5, p=.004), resort facilities (H=21.813, df=5, p=.001), weather (H=13.480, df=5, p=.019), checking in and out (H=12.826, df=5, p=.025), and guest facilities (H=11.693, df=5, p=.039).

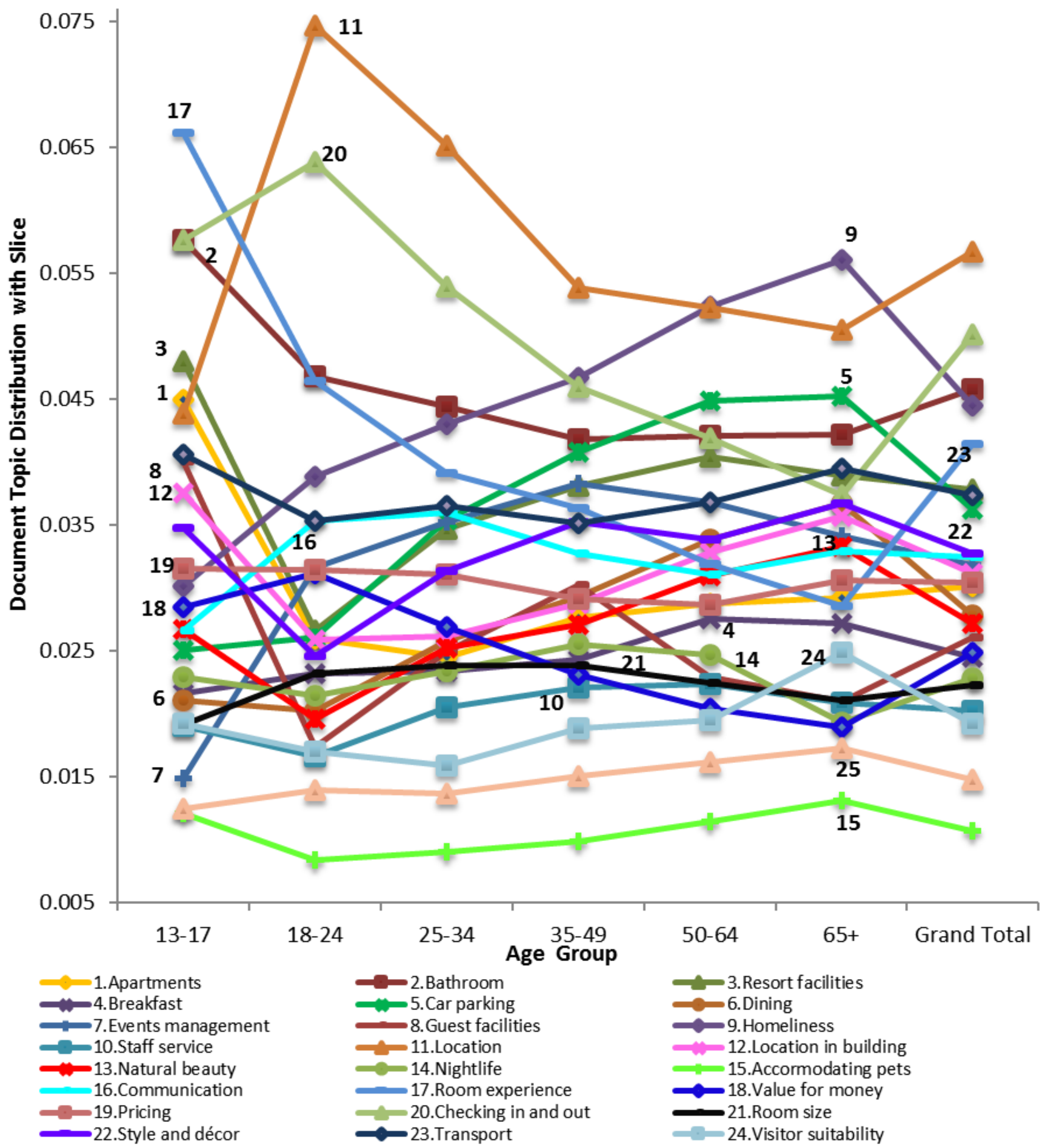

Figure 6. The relative important of dimensions perceived by different age groups

3.3. Stepwise regression analysis

As shown in Table 2, TripAdvisor enables consumers to rate their overall satisfaction level and satisfaction from five specific aspects – hotel location, cleanliness, room experience, service quality, and value for money – based on a 5-point Likert scale where 5 indicates 'the highest degree of satisfaction' and 1 indicates 'the lowest degree of satisfaction'. In this section, we run a stepwise regression analysis based on Equation 6. We use the regression results to compute the impact of the five ratings on the degree of consumer satisfaction. After discarding ratings with missing data, we had a final sample of n=73,203 for our regression analysis. Ratings on the five dimensions (hotel location, cleanliness, room experience, service quality, and value for money) were taken as independent variables and the overall satisfaction rating was used as the dependent variable.

$$y_i = \hat{y} + e_i = b_1 x_{1i} + b_2 x_{2i} + b_3 x_{3i} + b_4 x_{4i} + b_5 x_{5i} + e_i \quad (6)$$

Where,

y_i is the overall satisfaction rating of customer *i;*

x_{1i} is the rating of customer *i* for value for money;

x_{2i} is the rating of customer *i* for hotel location;

x_{3i} is the rating of customer *i* for room experience;

x_{4i} is the rating of customer *i* for cleanliness;

x_{5i} is the rating of customer *i* for service quality; and

e_i is a residual error assumed to follow a normal distribution $N(0, \sigma^2)$.

To analyze input-output relationships over the range of parameter variations and model outcomes, we operated a stepwise rank regression analysis to identify those input parameters that have the greatest influence on the uncertainty of a probabilistic model (Helton 1993). As shown in Table 7, column 3 shows the coefficient of determination (R^2) of the regression model at each step, and columns 5 and 6 show the significance of models. Columns 7 and 8 show the variables entered into the regression model with each successive step. Columns 9, 10 and 11 show the uncertainty importance measures, i.e., the standardized regression coefficient (*SRC*) (Helton et al. 1991), corresponding loss in the coefficient of determination (R^2_{loss}) (RamaRao et al. 1998), and partial correlation coefficient (*PCC*) (Draper and Smith 1981) associated with each of these variables, defined as follows:

$$SRC = \frac{\beta_j \sigma(x_j)}{\sigma(y)} \quad (2)$$

$$(R^2_{loss})_k = \frac{(1-R^2)}{\frac{1}{PCC_k^2}-1} \quad (3)$$

$$SCC[y, x_k] = \frac{\sum_j (x_{k,j} - \overline{x_k})(y_j - \bar{y})^2}{[\sum_j (x_{k,j} - \overline{x_k})^2 \sum_j (y_j - \bar{y})^2]^{1/2}} \qquad (4)$$

$$PCC[y, x_p] = SCC[y - y_{p\_fit}, x - x_{p\_fit}] \qquad (5)$$

Model	R	R^2	Adjusted R^2	F	Sig.	Rank	Variable	SRC	R^2_{loss}	PCC
1 [a]	.778	.605	.605	64388.512	<.001	1	Room experience	.765	.605	.778
2 [b]	.820	.672	.672	35746.147	<.001	1	Room experience	.463	.091	.467
						2	Service quality	.415	.067	.411
3 [c]	.826	.683	.683	24215.445	<.001	1	Room experience	.346	.033	.309
						2	Service experience	.332	.041	.338
						3	Cleanliness	.188	.011	.182
4 [d]	.828	.686	.686	18253.207	<.001	1	Room experience	.326	.029	.289
						2	Service quality	.314	.035	.318
						3	Cleanliness	.176	.009	.170
						4	Location	.060	.003	.104
5 [e]	.828	.686	.686	14603.282	<.001	1	Room experience	.325	.028	.288
						2	Service quality	.314	.035	.318
						3	Cleanliness	.176	.009	.170
						4	Location	.060	.003	.104
						5	Value for money	.033	<.001	.010

Notes:

a. Predictors: (Constant), Room experience;

b. Predictors: (Constant), Rooms experience, service quality.

c. Predictors: (Constant), Rooms experience, service quality, cleanliness

d. Predictors: (Constant), Rooms experience, service quality, location.

e. Predictors: (Constant), Rooms experience, service quality, location, value for money

f. Dependent Variable: the overall rating of satisfaction

Table 7. Stepwise regression models and sensitivity analysis (N=73,203)

	Unstandardized		Standardized coefficient	t	Sig.	Collinearity Diagnostics	
	B	Std. Error	Beta			Tolerance	VIF
(Constant)	.057	.012		4.567	.000		
Room experience	.336	.004	.330	81.364	.000	.260	3.843
Service quality	.325	.004	.319	90.633	.000	.345	2.896
Cleanliness	.202	.004	.188	46.800	.000	.265	3.773
Location	.098	.003	.076	28.333	.000	.596	1.678
Value for money	.001	.000	.006	2.803	.005	.980	1.020

Table 8. Coefficients for Model 5

Tables 8 and 9 provide empirical evidence of a strong relationship between the combination of five independent variables and the dependent variable at the 0.1% level of significance; the five specific dimensions explain 68.6% of the variance in overall satisfaction. The Durbin-Watson test statistic (1.924) was close to 2, confirming that residual error values were independent.

Subsequently, based on the above analysis, the fitted regression model is as follows:

Overall customer rating ($\hat{y}$) = 0.057 + 0.330*(*Room experience*) + 0.319*(*Service quality*) + 0.188*(*Cleanliness*) + 0.076*(*Location*) + 0.006*(*Value for money*)

The findings show that room experience has the strongest impact on consumer satisfaction, followed by service quality. However, the dimension of value for money appears to be the least important factor influencing consumer satisfaction. To further test for multicollinearity, we computed variance inflation factors (VIFs). All VIFs were found to be less than the conservative threshold of 5, suggesting that multicollinearity is not a major issue in our study. Stepwise regression analysis and LDA may be regarded as a complementary analyses of data: the former focuses on online ratings (i.e., numerical data) while the latter examines online text content (textual data). For example, the research findings of stepwise regression analysis may identify that room experience is the most important dimension of satisfaction for hotel customers, while the findings of LDA may shed light on how hotel owners can improve room experience through the identification of dimensions such as homeliness.

We also examined the relative importance of variables via the uncertainty importance factor (*UIF*) and the ratio of R^2_{loss} to R^2, the results of which are summarized in Figure 7. Service quality has the largest R^2_{loss} and PCC, although room experience has a larger coefficient in the regression model than service quality. This indicates that most of the uncertainty of the overall evaluation is influenced by service quality. The variance in the overall rating would be reduced by approximately 5.13% if service quality was to be held fixed at its nominal value. The uncertainty of value for money has the least influence on consumers' overall evaluation: consumers appear more inclined to value their living experiences (e.g., service quality and room experience).

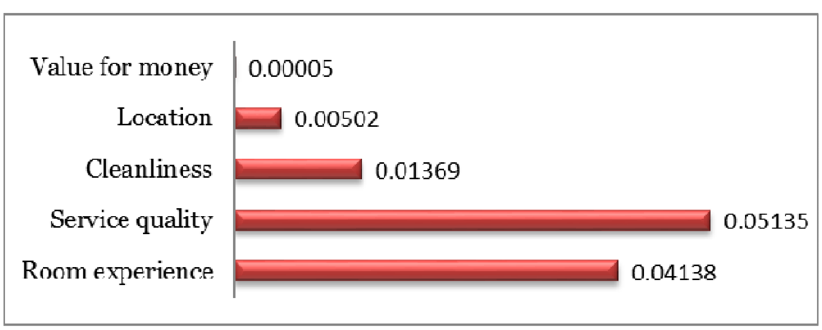

Figure 7. Uncertainty importance factor of R^2_{loss} / R^2

3.4. Perceptual mapping of hotels

The section presents the perceptual mapping of hotels on satisfaction attributes. Perceptual or positioning maps (Dev et al. 1995; DeSarbo et al. 2008) are tools used by marketers and researchers to capture consumer evaluations of competing hotel products using two or more attributes, represented in a two-dimensional map. The attributes used to map competitors are typically dimensions that are important to consumers (see Table 4). TripAdvisor categorizes all hotels into nine levels from one-star to five-star, in increments of 0.5. We determine the different marketing positions of the nine hotel classifications by mapping the "distance" between the different hotels on a given dimension. A hotel is associated with each online review and is given star rating. Using the automatically generated dimension clusters, we generate a positioning matrix for hotel star ratings, counting the number of hotel star occurrences for each attribute. Then, to turn this into a visual "map," we use correspondence analysis (CA), a technique for analyzing two-way, two-mode frequency data (Everitt and Dunn 2001), making it more appropriate for this task than the continuously-scaled multidimensional scaling procedures commonly used for market structure maps. The CA approach is designed to help generate hypotheses by representing the data in a reduced space as determined by consulting the eigenvalues and the corresponding scree plot (Greenacre 1992). To help interpret the dimensions in the reduced space, we use brand mapping for the coordinates (x, y) of each hotel star rating on the derived dimensions. To ensure stability of the regression results, we use a rotationally invariant, asymmetric CA (Bond and Michailidis 1997).

In a correspondence analysis, we try to simplify dimensions by collapsing variables into as few factors as possible and examining them in two-dimensional space. We do this by examining inertia (Greenacre 1992). Geometrically, inertia is the weighted average of the Chi-squared distances from the centroid. Topics which contribute highly to a principal axis largely determine its orientation and identity. As shown in Table 9, F1 with one dimension has principal inertia of 80.4%. When combined with F2, the two factors explain 96.7% of the accumulated inertia ratio –the total information from all dimensions. We will examine star ratings and dimensions according to these two factors, comprehensively summarizing the data set.

$F_{\text{No. of dimensions}}$	Singular value	Inertia	Inertia ratio	
			Explained	Accumulated
F1	.413	**.171**	**.804**	**.804**
F2	.186	**.035**	**.164**	**.967**
F3	.074	.005	.026	.993
F4	.031	.001	.005	.997
F5	.018	.000	.001	.999
F6	.013	.000	.001	1.000
F7	.008	.000	.000	1.000
F8	.005	.000	.000	1.000
Total		.213	1.000	1.000

Note: The square of singular value is the inertia value.

Table 9. Summary of CA

As shown in Figure 8, regressing the CA coordinates (F1 and F2) on the extracted dimensions defines F1 as a combination of comments about the following dimensions: apartments, breakfast, homeliness, communication, room experience, value for money, checking in and out, room size, style and decor, events management, resort facilities, bathroom, accommodating pets, guest facilities, dining, and pricing. F2 is a combination of car parking, staff service, and location in building.

We now assess the relative importance of dimensions derived from online reviews for different hotel classifications, shown on a hotel perceptual map (see Figure 8). As explained above, the 19 controlled dimensions are factors that consumers frequently mentioned in reviews, captured through our LDA analysis. The position of star-rated hotels with respect to specific attributes represents consumers' perceived ratings in relative terms. Referring to Figure 8, for one-star hotels, there are acute angles within close proximity to room experience and communication, suggesting that they are prevalent for this subgroup. On the contrary, five-star hotels appear to impress customers with their homeliness, strong events management capability, and pet friendliness. Customers of 4 and 4.5 star hotels are inclined to consider the quality from the perspectives of hotel and resort facilities, food quality and room size and decoration. Consumers of mid-range hotels (i.e., two- or three-stars) appear to focus on hotel performance from the perspective of more basic aspects, including car parking, checking in and out, hotel staff service, bathroom facilities, and price. The further the distance from the origin, the more different profiles are from the average. As shown in Figure 8, 5-star hotels are furthest from the origin; the profiles of 5-star hotels are most different to the average profile of other hotels as a result of their differentiation through luxury. In obverse, one-star hotels are further from the origin than the mid-range hotels owing to their below average performance.

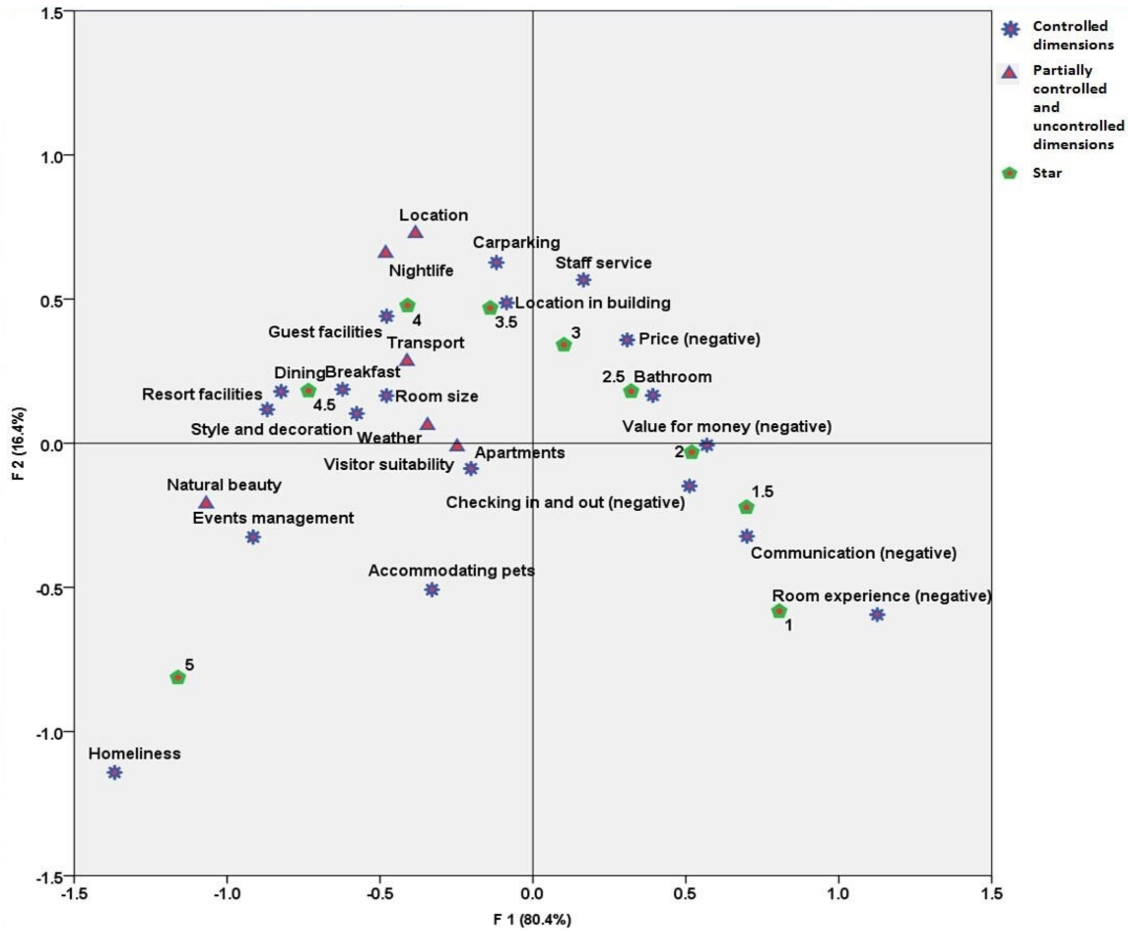

Figure 8. Perceptual mapping of hotels from customer reviews - F1 versus F2

4. Discussion and conclusions

This study proposes a novel approach to extract latent dimensions of consumer satisfaction from rich online customer reviews. For dimension extraction, the LDA analysis of customer reviews reveals meaningful dimensions that are not found via traditional means. The relative importance of the extracted dimensions is identified according to the intensity of the conversations for each. We also estimate the heterogeneity of perceptions across different demographic profiles of consumers using the dimensions. The study enjoys a relatively broad sample of 25,670 hotels located in 16 countries, enabling us to make more reliable generalizations than prior studies using traditional research methods. This research further provides a stepwise regression and sensitivity analysis for TripAdvisor's five consumer ratings for hotels and overall consumer satisfaction. Room experience and service quality are identified as the most important dimensions in our analysis. This supports the findings of prior studies that have proposed room experience and service quality as two key factors influencing customer satisfaction, typically via small data samples based on questionnaire surveys (e.g., Skogland and Siguaw, 2004; Choi and Chu, 2001).

In principle, the more data that is available, the more accurate will be the generalizations made. Our statistical analysis was calculated based on 266,544 reviews created by 39,287 unique reviewers; our regression results should therefore be more reliable and accurate than prior statistical results based on limited sample data. Moreover, many online rating indicators represent multidimensional variables, such as room experience and service quality, and these may involve different sub-dimensions. For example, poor efficiency during check-in and during event management can both lead to poor service quality perceived by customers. Our method can identity the most prevalent factors perceived by customers when they express their opinions and share hotel experiences online.

This study has many valuable implications for managerial practice. First, it enables hotel managers and investors to ascertain the importance and heterogeneity of latent dimensions of consumer satisfaction from user-generated data. Second, perceptual mapping of hotels through the voice-of-the customer reveals which dimensions are salient for influencing consumer satisfaction and how consumer perceptions vary among different hotel classifications. In this way, dimensions may be used to help identify unique submarkets. For example, the economy or budget hotel sector has rapidly developed in the past decade in China (Mohsin, and Lengler, 2015; Ren et al., 2016). Budget hotels usually offer a limited hotel service and their rates are 25-30% cheaper than average market rates (Gilbert and Lockwood, 1990). Most hotels in this segment are considered zero to three-star (Ruetz and Marvel, 2011), but, apart from price, there is a lack of understanding about the factors influencing consumer behavior in this area of the market (Ren et al., 2016). Based on our online review analysis, several important non-price dimensions are identified by customers of two- to three-star hotels, including "bathroom" and "checking in and out". In addition, our research findings suggest that 5-star hotel managers and owners should focus upon the feeling homeliness for customers; it is one of the most important dimensions influencing customer satisfaction for 5-star hotels, particularly compared to other lower-level hotels. More generally, Figure 7 provides a guide as to the important elements defining each segment of the hotel market according to star ratings. Clearly, hotels need to fulfil basic lower level dimensions and focus on more complex and financially expensive dimensions in order to move further up the star rating scale. Dimensions will differ according to demography and this is clearly shown in our study. For example, our LDA analysis shows that men are more sensitive to price than women. As a result, hotel managers might consider offering some special discounts for single male customers, particularly during low peak periods. Similarly, older hotel customers value homeliness more than younger customers. Thus, for example, advertising and promotions for older target customer segments should emphasize homeliness. Other findings in our demographic investigation also provide potential avenues for customer segmentation, targeting, and positioning of hotel marketing.

This study has some important limitations. First, the models used to extract the latent dimensions of quality are computationally intensive. However, with the advances in computing and largescale computing techniques, this limitation will dissipate over time. Second, this study focuses only on hotel reviews, but it could be extended to online reviews for other tourism and travel-related service industries (e.g., restaurants and attractions) to reveal or confirm dimensions of customer satisfaction. Third, we do not analyze rare or infrequent words in the long tail of the distribution; these words could reflect emerging consumer preferences that could be very helpful in developing new hotel marketing space. Each of the above limitations could be rich avenues for further research.

References

Alegre J. and Garau, J. (2010). Tourist satisfaction and dissatisfaction, A*nnals of Tourism Research,* 37(1)*,* 52-73.

Al Khattab, S. and Aldehayyat, J. S. (2011). Perception of service quality in Jordanian hotels. *International Journal of Business Management, 6*(7), 226-233.

Amblee, N. (2015). Research Note -- The impact of eWOM density on sales of travel insurance, A*nnals of Tourism Research*, in press.

Anderson E.W., Fornell, C., Lehman, D.R. (1994) Customer satisfaction, market share, and profitability: Findings from Sweden, *Journal of Marketing*, 58(3), 53-6

Anderson, C (2008). *The Long Tail: Why the Future of Business Is Selling Less of More*. New York: Hyperion Books.

Asuncion, A., Welling, M., Smyth, P. and Teh. Y.W. (2009). On smoothing and inference for topic models, *in Proceedings of the UAI*, 27–34.

Barsky, J. D. (1992). Customer satisfaction in the hotel industry meaning and measurement. *Journal of Hospitality & Tourism Research*, 16(1), 51-73.

Blesic, I., Ivkov-Dzigurski, A., Stankov, V., Stamekovic, I, and Bradic, M. (2011). Research of expected and perceived service quality in hotel management. *Journal of Tourism,* 11, 5-13.

Blei, D.M., Andrew Y. N.G., and Michael I.J. 2003. Latent Dirichlet Allocation, *Journal of Machine Learning Research*, 3, 993–1022.

Blei, D.M. (2012), Probabilistic topic models, *Communications of the ACM*, 55 (4), 77–84.

Bond, Jason and George Michailidis (1997), Interactive correspondence analysis in a dynamic object-oriented environment, *Journal of Statistical Software*, 2(8)), 1-31.

Chang, H. (2008). Increasing hotel customer value through service quality cues in Taiwan. *The Service Industries Journal, 28*(1), 73-84.

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

Chintagunta, P.K., Gopinath, S., and Venkataraman, S. (2010). The effects of online user reviews on movie box office performance: accounting for sequential rollout and aggregation across local markets, *Marketing Science*, 29(5), 944-957

Chi, C.G. and Qu, H. (2009). Examining the relationship between tourists' attribute satisfaction and overall satisfaction. *Journal of Hospitality Marketing & Management, 18*, 4-25.

Chen, P.Y. Wu S.Y., and Yoon J. (2004). The impact of online recommendations and consumer feedback on sales, *in Proceedings of the International Conference on Information Systems* , 711–24.

Chevalier, J and Mayzlin, D.A. (2006). The Effect of Word of Mouth on Sales: Online Book Reviews, *Journal of Marketing Research*, 43 (August), 345–54

Choi, T. Y., and Chu, R. (2001). Determinants of hotel guests' satisfaction and repeat patronage in the Hong Kong hotel industry. *International Journal of Hospitality Management*, 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.

Clemens, M. D., Gan, C., and Ren, M. (2010). Synthesizing the effects of service quality, value, and customer satisfaction on behavioral intentions in the motel industry: An empirical analysis. *Journal of Hospitality & Tourism Research, 35*(4), 530-568.

Cohen, J (1960). A coefficient of agreement for nominal scales. *Educational and Psychological Measurement* 20 (1): 37

Cronin, J.J and Taylor, S.A. (1992) Measuring Service Quality: A Reexamination and Extension, *Journal of Marketing*, 56 (July), 55-68.

Danaher, P. J., & Haddrell, V. (1996). A comparison of question scales used for measuring customer satisfaction. *International Journal of Service Industry Management*, 7(4), 4-26.

Dellarocas, C., and Narayan, R. (2006). A statistical measure of a population's propensity to engage in post-purchase online word-of-mouth, *Statistical Science* 21(2), 277-285.

Dellarocas, C, Zhang, X and Awad, N.F. (2007), Exploring the value of online product reviews in forecasting sales: the case of motion pictures, *Journal of Interactive Marketing*, 21 (4), 23-5

Dellarocas, C., Gao, G., and Narayan, R. (2010). Are consumers more likely to contribute online reviews for hit or niche products?, *Journal of Management Information Systems*, 27(2),127-157.

Dev, C., Morgan, M.S. and Shoemaker, S. (1995). A positioning analysis of hotel brands C*ornell Hotel and Restaurant Administration Quarterly*, 36 (6), 48-55.

DeSarbo, W.S., Grewal, R. and Scott, C.J. (2008). A clusterwise bilinear multidimensional scaling methodology for simultaneous segmentation and positioning analyses, *Journal of Marketing Research*, 45 (June), 280–92.

Draper, N.R., and Smith, H. (1981). *Applied Regression Analysis*, 2nd ed. New York: John Wiley.

Duan, W, Gu, B and Whinston, A.B. (2008). Do online reviews matter? An empirical investigation of panel data, *Decision Support Systems*, 45 (4), 1007–1016.

Emir, O. and Kozak, M. (2011). Perceived importance of attributes on hotel guests' repeat visit intentions. *Tourism, 59*(2), 131-143.

Everitt, B.S. and Dunn, G. (2001). *Applied Multivariate Data Analysis*. New York: Oxford University Press.

Fawzy, A. (2010). Business travelers' accommodation selection: A comparative study of two international hotels in Cairo. *International Journal of Hospitality & Tourism Administration, 11*, 138-156.

Fleiss, J.L. (1971), Measuring nominal scale agreement among many raters, *Psychological Bulletin*, 76 (5), 378–82.

Forman, C., Ghose, A., and Wiesenfeld, B. (2008). Examining the relationship between reviews and sales: the role of reviewer identity disclosure in electronic markets, *Information Systems Research*,19(3), 291-313.

Fornell, C. (1992). A national customer satisfaction barometer: the Swedish experience. *The Journal of Marketing*, 6-21.

Gagnon, G.B. and Roh, Y. S. (2008). The impact of customization and reliability on customer satisfaction in the U.S. lodging industry. *Journal of Quality Assurance in Hospitality & Tourism, 8*(3), 60-78.

Gao, G., Greenwood, B.N. and Ritu, A. (2015). Vocal minority and silent majority: how do online ratings reflect population perceptions of quality? *MIS Quarterly*, 39(3), 565-589.

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.

Gill, S. M., Hudson, S., and Quintana, T. A. (2006). The influence of service recovery and loyalty on perceived service quality: A study of hotel customers in Spain. *Journal of Hospitality & Leisure Marketing, 14*(2), 47-68.

Gilbert, D.C., and Lockwood A. (1990). Budget hotels-the USA, France and UK compared. Travel & Tourism Analyst, 3(1), 19-35.

Greenacre, M.J. (1992), Correspondence Analysis in Medical Research, *Statistical Methods in Medical Research*, 1 (1), 97–117.

Godes, D and Mayzlin D. (2004). Using online conversations to study word-of-mouth communication, *Marketing Science*, 23 (4), 545–60.

Griffin, A., & Hauser, J. R. (1993). The voice of the customer. *Marketing Science*, 12(1), 1–27.

Griffiths, T.L. and Steyvers, M. (2004). Finding Scientific Topics, in *Proceedings of the National Academy of Sciences of the United States of America*, 101 (Supplement 1), 5228–35

Gunderson, M. G., Heide, M., and Olsson, U.H. (1996). Hotel guest satisfaction among business travelers. *Cornell Hospitality Quarterly, 37*(2), 72.

Gu, B., Park, J., and Konana, P. (2012). Research note—the impact of external word-of-mouth sources on retailer sales of high-involvement products, *Information Systems Research*, 23(1), 182-196.

Hansen, F. (1976), Psychological theories of consumer choice, *Journal of Consumer Research*, 3 (3), 117–42.

Hallowell, R. (1996). The relationships of customer satisfaction, customer loyalty, and profitability: an empirical study. *International journal of service industry management*, 7(4), 27-42.

Heung, V. C. S. (2000). Satisfaction levels of Mainland Chinese travelers with Hong Kong hotel services. *International Journal of Contemporary Hospitality Management, 12*(5), 308-315.

Helton, J.C. (1993). Uncertainty and sensitivity analysis techniques for use in performance assessment for radioactive waste disposal. *Reliability Engineering & System Safety* 42(2–3), 327–367

Helton, J.C., Garner, J.W., McCurley, R.D. and Rudeen, D.K. (1991). Sensitivity analysis techniques and results for performance assessment at the Waste Isolation Pilot Plant. SAND 90-7103. Albuquerque, New Mexico: Sandia National Laboratories

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.

Hofmann T. (1999). Probabilistic latent semantic indexing, in Proceedings of the 22nd annual international ACM SIGIR conference on research and development in information retrieval, 50-57

Hsieh, A. and Tsai, C. (2009). Does national culture really matter? Hotel service perceptions by Taiwan and American tourists. *International Journal of Culture Tourism and Hospitality Research*, 3(1), 54-69.

Hui, T. K., Wan, D., & Ho, A. (2007). Tourists' satisfaction, recommendation and revisiting Singapore. *Tourism management*, 28(4), 965-975.

Kang, S., Okamoto, N., & Donovan, H. A. (2004). Service quality and its effect on customer satisfaction and customer behavioral intentions: Hotel and ryokan guests in Japan. *Asia Pacific Journal of Tourism Research*, 9(2), 189-202.

Katz, G. (2011). Nine myths about the voice of the customer. *Visions* (October), 24–25.

Kandampully, J. and Suhartanto, D. (2000). Customer loyalty in the hotel industry: The role of customer satisfaction and image. *International Journal of Contemporary Hospitality Management*, 12(6), 346-351.

Kim, W.G., Ng, C.Y.N., and Kim, Y.S. (2009a). Influence of institutional DINESERV on customer satisfaction, return intention, and word-of-mouth. *International Journal of Hospitality Management*, 28(1), 10-17.

Kim, D. J., Kim, W. G., & Way, K. A. (2009b). Adapting the customer satisfaction index to the lodging industry: Foreign customers' evaluations. *Florida International University Review, 27*(1), 18-36.

Klein, B. and Leffler, K.B. (1981), The role of market forces in assuring contractual performance, *Journal of Political Economy*, 89 (4), 615–41.

Kuo, C. (2007). The importance of hotel employee service attitude and satisfaction of international tourists. *The Service Industries Journal, 27*(8), 1073-1085.

Lau, P. M., Akbar, A. K., and Fie, D. Y. G. (2005). Service quality: A study of the luxury hotels in Malaysia. *Journal of American Academy of Business, 12*(2), 46-55.

Law, R. & Yip, R. (2010). A study of satisfaction level of Hong Kong tourists with hot springs hotels and resorts in Guangdong, China. *Florida International University Review, 28*(1), 83-107.

Lee, T., and Bradlow, E. (2006). Automatic construction of conjoint attributes and levels from online customer reviews. *Working Paper*, Wharton School

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, X., and Hitt, L.M. (2008). Self-selection and information role of online product reviews, *Information Systems Research*, 19(4), 456-474.

Li, H.Y., 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.

Liu, Y. 2006. Word of mouth for movies: Its dynamics and impact on box office revenue. *Journal of Marketing* 70 (3): 74-89.

Liu, Z.W., Park, S.W. (2015). What makes a useful online review? Implication for travel product websites. *Tourism Management*, 47, 140-151.

Lu, W., and Stepchenkova, S. (2012). Ecotourism experiences reported online: Classification of satisfaction attributes. *Tourism Management*, 33(3), 702-712.

Markovic, S., Raspor, S., and Segaric, K. (2010). Customer satisfaction and customer loyalty measurement in hotel settings: An empirical analysis. *Tourism & Hospitality Management*, 125-137.

Mitra, D and Golder, P.N. (2006) How does objective quality affect perceived quality? short-term effects, long-term effects, and asymmetries, *Marketing Science*, 25(3), 230–47.

Min, H. N., Lim, Y. M. Magnini, V. P. (2015) Factors affecting customer satisfaction in responses to negative online hotel reviews: The impact of empathy, paraphrasing, and speed, *Cornell Hospitality Quarterly*, 56(2), 223-231

Moe, W.W., and Trusov, M. (2011). The value of social dynamics in online product ratings forums, *Journal of Marketing Research,* 48(3), 444-456.

Mohsin, A., Hussain, I., & Rizwan, M. (2011). Exploring service quality in luxury hotels: Case of Lahore, Pakistan. *Journal of American Academy of Business, 16*(2), 296- 303.

Mohsin, A. & Lockyer, T. (2010). Customer perceptions of service quality in luxury hotels in New Delhi, India: An exploratory study. *International Journal of Contemporary Hospitality Management, 22*(2), 160-173.

Mohsin, A., and Lengler, J. (2015). Service experience through the eyes of budget hotel guests: do factors of importance influence performance dimensions? *Journal of Hospitality and Tourism Management*, 23, 23-34.

Mudambi, S.M., and Schuff, D. (2010). What makes a helpful online review? A study of customer reviews on amazon.com, *MIS Quarterly*, 34(1), 185-200

Nadiri, H. and Hussain, K. (2005). Perceptions of service quality in North Cyprus hotels. *International Journal of Contemporary Hospitality Management, 17*(7), 469-480.

O'Connor, P. (2008). User-generated content and travel: A case study on Tripadvisor.com. *Information and Communication Technologies in Tourism*, 47-58.

Oliver, R.L. (1980). A cognitive model of the antecedents and consequences of satisfaction decisions. *Journal of Marketing Research*, 17(4), 460-469.

Park, S. and Nicolau, J.L. (2015) Asymmetric effects of online consumer reviews, A*nnals of Tourism Research*, 50, 67-83

Prayukvong, W., Sophon, J., Hongpukdee, S., and Charapas, T. (2007). Customer's satisfaction with hotel guestrooms: A case study in Ubon Rachathari province, Thailand. *Asia Pacific Journal of Tourism Research, 12*(2), 119-126.

Ramage, D., and Rosen, E. (2011). *Stanford Topic Modeling Toolbox Version 0.4.0*. Available online: http://nlp.stanford.edu/software/tmt/tmt-0.4/

Ramanathan, R. (2012). An exploratory study of marketing, physical and people related performance criteria in hotels. *International Journal of Contemporary Hospitality Management, 24*(1), 44-61.

Ramanathan, U. & Ramanathan, R. (2011). Guests' perceptions on factors influencing customer loyalty. *International Journal of Contemporary Hospitality Management, 23*(1), 7-25.

RamaRao, B.S., Mishra, S., Sevougian, S.D. and Andrews, R.W. (1998). Uncertainty importance of correlated variables in a probabilistic performance assessment. In *Proceedings of SAMO'98, Second International Symposium on Sensitivity Analysis for Model Output*, Venice, Italy, April 19-22.

Resnick, P and Zeckhauser R. (2002), Trust among strangers in internet transactions: empirical analysis of ebay's reputation system," in The Economics of the Internet and E-Commerce, Advances in Applied Microeconomics, 11, Michael R. Bay, ed. Amsterdam: Elsevier Science, 127–57.

Ren, L., Qiu, H., Wang, P., and Lin, P. M. (2016). Exploring customer experience with budget hotels: Dimensionality and satisfaction. *International Journal of Hospitality Management*, 52, 13-23.

Ruetz, D., and Marvel, M. (2011). Budget hotels: Low cost concepts in the US, Europe and Asia. In R. Conrady, M. Buck (Eds.), Trends and Issues in Global Tourism 2011, Springer, Verlag Berlin

Ryan, C. and Huimin, G. (2007). Perceptions of Chinese hotels. *Cornell Hospitality Quarterly, 48*(4), 380-391.

Sanchez-Gutierrez, J., Gonzalez-Uribe, E. G., & Coton, S. G. H. (2011). Customer satisfaction in the hospitality industry in Guadalajara, Mexico. *Advances in Competitiveness Research, 19*(4), 17-30.

Senecal, S. and Nantel, J.(2004). The influence of online product recommendations on consumers' online choices, *Journal of Retailing*, 80 (2), 159–69

Sim, J., Mak, B., and Jones, D. (2006). A model of customer satisfaction and retention for hotels. *Journal of Quality Assurance in Hospitality & Tourism, 7*(3), 1 -23.

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

Sotiriadis, M. D., & van Zyl, C. (2013). Electronic word-of-mouth and online reviews in tourism services: the use of twitter by tourists. *Electronic Commerce Research*, 13(1), 103-124.

Sparks, B. A., & Browning, V. (2011). The impact of online reviews on hotel booking intentions and perception of trust. *Tourism Management* 32(6), 1310-1323.

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

Tellis, G.J. and Johnson, J. (2007), The value of quality, *Marketing Science*, 26 (6), 758–73.

Tirunillai, S., and Tellis, G.J. (2014). Mining marketing meaning from online chatter: analysis of big data using latent dirichlet allocation, *Journal of Marketing Research*, (August), 463-479

Wang, Y., Vela, M. R., & Tyler, K. (2008). Cultural perceptions: Chinese perceptions of UK hotel service quality. *International Journal of Culture, Tourism and Hospitality Research, 2*(4), 312-339.

Weng, M., Ha, J., Wang, Y., and Tsai, C. (2012). A study of the relationship among service innovation, customer value, and customer satisfaction. *The International Journal of Organizational Innovation*, 4(3), 98-112.

Woodruff, R. B., Cadotte, E. R., and Jenkins, R. L. (1983) Modeling consumer satisfaction processes using experience-based norms, *Journal of Marketing Research*, 20(3), 296-304

Wu, C.H.J., and Liang, R.D. (2009). Effect of experiential value on customer satisfaction with service encounters in luxury-hotel restaurants. *International Journal of Hospitality Management*, 28(4), 586–593.

Vermeulen, I. E., & Seegers, D. (2009). Tried and tested: The impact of online hotel reviews on consumer consideration. *Tourism Management*, 30(1), 123-127.

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

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

Ye, Q., Zhang, Z., & Law, R. (2009b). Sentiment classification of online reviews to travel destinations by supervised machine learning approaches. *Expert Systems with Applications* 36(3), 6527-6535.

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

Yilmaz, I. (2009). Measurement of service quality in the hotel industry. *International Journal of Tourism & Hospitality Research, 20*(2), 375-386.

Zhu, F., and Zhang, X. (2010). Impact of online consumer reviews on sales: The moderating role of product and consumer characteristics, *Journal of Marketing*, 74(2), 133-148.

Appendix A. Previous Empirical Research Related to Online Reviews.

Study	**Variables Analyzed**	**Research Context, Method and Sample Size**	**Findings**
Resnick and Zeckhauser (2002)	The number of positive, neutral and negative feedback reviews; the number of problematic transactions.	• eBay marketplaces. • Multiple regression. • 1580 negative feedback reviews.	Sellers with better reputations are more likely to sell their items, but they enjoy no boost in price.
Godes and Mayzlin (2004)	The number of posts; past ratings.	• Television shows from Usenet newsgroups. • Multiple regression. • 169 groups and 2398 posts evaluated.	The dispersion of online conversations may be used to measure ratings.
Chen et al. (2004)	The numbers of reviews; the number of recommendations; sales rank.	• Book data from Amazon.com. • Multiple regression. • 610 observations, with 58566 total reviews.	Consumer ratings are not found to be related to sales, but recommendations are highly significant.
Senecal and Nantel (2004)	Categorical independent variables: exposition to a recommendation; type of product; type of website; recommendation source. Dichotomous dependent variable: selection or non-selection of a recommended product.	• Computer mice, calculators and wine. • Generalized estimating equations. • 3 (websites) × 4 (recommendation sources) × 2 (products). • Online experiment was conducted with 487 subjects.	Online recommender systems are more influential than expert reviews when determining customer product choice.
Liu (2006)	The number of positive, neutral and negative word-of-mouth comments.	• Movie box office: Yahoo Movies. • Pooled regression. • 40 movies, 12136 WOM messages.	Posted reviews are highly correlated with box office sales.
Chevalier and Mayzlin (2006)	The number of reviews per book; average stars.	• Books on Amazon and BN. • Differences-indifferences method. • Over three time points, 1636-2387 observations, with 12.79-68.31 reviews per book. • A total of 134904-176112 reviews on each time point.	An improvement in a book's reviews leads to an increase in relative sales.
Dellarocas	Number of movies; number of	• Film data from Yahoo!	The density of online ratings can

and Narayan (2006)	user ratings; number of critic ratings; number of unique users.	Movies and BoxOfficeMojo. • A differences-indifferences approach to eliminate book- and site-specific effects. • 104 movies with total number of 63,889 user ratings and 1392 critic reviews.	be used as a proxy of a population's propensity to engage in post-purchase online.
Dellarocas et al. (2007)	Number of movies; number of user ratings; number of critic ratings; number of unique users.	• Movie box-office: Yahoo! Movies, BoxOfficeMojo. • Hazard rate model. • 80 movies, with 55156 total user ratings and 1040 total critic ratings.	Online movie ratings can be used as a proxy for word-of-mouth.
Duan et al. (2008)	User review rating; number of user postings.	• Movie box-office: Yahoo! Movies, Variety.Com, BoxOfficeMojo.com. • A three-stage least-squares (3SLS) procedure. • 71 movies with mean of 1350.24 users' posts, 95867 total user posts.	The rating of online reviews has no significant impact on movies' box office revenues.
Forman et al. (2008)	Number of reviews per book; average stars.	• Books from Amazon.com. • Ordinary least squares with product-level fixed effects. • 786 books with 175714 reviews.	Reviewer disclosure of identity may be used as a measurable proxy for both future sales and future geographic sales.
Li and Hitt (2008)	Average rating of all reviews.	• Books from Amazon.com. • Exponential Model. • 136,802 single review observations for 2,651 books.	Word-of-mouth is not an unbiased indicator of quality and will affect sales.
Mudambi and Schuff (2010)	Star rating of the reviewer; total number of votes on each review's helpfulness; word count of the review.	• Amazon.com. • Tobit regression. • 1,587 reviews of 6 products.	Review depth is correlated with helpfulness, but review extremity is less helpful for experience goods.
Chintagunta et al. (2010)	Valence, volume and precision of online reviews.	• Film data from Yahoo! Movies website. • Logistic regression, generalized method of moments (GMM) procedure. • 148 movies, 253 markets	Online user reviews are correlated with film box office performance.

		with average volume of 474.82 reviews, 70273 total reviews.	
Dellarocas et al. (2010)	Number of user ratings; number of movies.	• Films data from Yahoo! Movies. • Ordinary least squares (OLS). • 2002 data set contains 104 movies and 63,889 reviews; • 2007–8 data set contains 143 movies and 95,443 reviews.	Products that are less available and less successful in the market are less likely to receive online reviews.
Zhu and Zhang (2010)	Average rating, coefficient of variation of ratings; the total number of reviews posted.	• Video game industry (single-purchase products). • Psychological choice model, differences-in-differences approach. • 220 games, 4292 reviews.	Online reviews are more influential for less popular games and games whose players have greater Internet experience.
Moe and Trusov (2011)	Average of all ratings; variance across all ratings; total number of ratings.	• Bath, fragrance and beauty products. • Exponential hazard function; two-stage estimation approach. • 500 products, 3801 ratings.	Word-of-mouth affects sales and is subject to social dynamics in that ratings will affect future rating behavior.
Gu et al. (2012)	Average customer rating; number of ratings.	• Amazon, Cnet, DpReview, and Epinions. • Logistic regression. • 31522 reviews of 148 digital cameras.	WOM on external review websites is a more effective indicator of sales for high-involvement products.
Li et al. (2013)	Type of reviews; contents of reviews.	• TripAdvisor. • Content analysis, ICTCLAS tool. 42,886 reviews of 774 star-rated hotels in Beijing.	Determinants of customer satisfaction in hospitality venues can be identified through an analysis of online reviews.
Gao et al. (2015)	Average online rating; overall physician quality measure.	• Care physicians (professional services). • Logistic regression. • Dataset comprises 1,425 physicians of which 794 have been rated online.	Physicians who are rated lower in quality by the patient population are less likely to be rated online: online ratings affect underlying consumer perceived quality.
Liu and Park (2015)	Review star rating; review length; perceived review readability.	• Restaurants from Yelp.com. • Logistic regression. • 5090 online reviews.	Qualitative aspects of reviews were identified as the most influential factors that make travel reviews useful.
Park and	Number of "useful" votes	• Restaurants (tourism and	Valence of online reviews has a

Nicolau (2015)	awarded to the review; ratings; number of “funny” votes that were given to the review.	hospitality products). • Negative Binomial distribution. • 5090 reviews of 45 restaurants in London and New York.	U-shaped effect on usefulness and enjoyment. Negative ratings of reviews are more useful than positive reviews. Positive ratings are associated with higher enjoyment than negative reviews.
Amblee (2015)	Density of negative reviews.	• SquareMouth.com. • Pooled OLS regression. • Over 21,000 reviews of insurance providers.	When the density of negative reviews is high, sales are lower, and vice versa.