Understanding Important Hotel Attributes from the Consumer Perspective over Time

Sungha Jang
Tian Liu
Ji Hye Kang
University of Rhode Island, jh Kang @ uri.edu
Huichen Yang

Follow this and additional works at: https://digitalcommons.uri.edu/tmd_facpubs

Citation/Publisher Attribution
Available at: https://doi.org/10.1016/j.ausmj.2018.02.001

This Article is brought to you by the University of Rhode Island. It has been accepted for inclusion in Textiles, Fashion Merchandising and Design Faculty Publications by an authorized administrator of DigitalCommons@URI. For more information, please contact digitalcommons-group@uri.edu. For permission to reuse copyrighted content, contact the author directly.
Understanding Important Hotel Attributes from the Consumer Perspective over Time

The University of Rhode Island Faculty have made this article openly available. Please let us know how Open Access to this research benefits you.

This is a pre-publication author manuscript of the final, published article.

Terms of Use
This article is made available under the terms and conditions applicable towards Open Access Policy Articles, as set forth in our Terms of Use.

This article is available at DigitalCommons@URI: https://digitalcommons.uri.edu/tmd_facpubs/72
Understanding Important Hotel Attributes from the Consumer Perspective Over Time

Abstract

Consumers consider various product attributes when they evaluate products. Researchers and practitioners have used multi-attribute models to understand which product attributes are important for consumers. However, in those models, product attributes are limited and are determined by researchers at the time of the inquiry. In this study, using a longitudinal study of hotel reviews over six years, the top 30 important hotel attributes from the perspective of consumers are identified and examined as to how the importance of these hotel attributes has changed over time. Our findings show that staff is the most important attribute with a positive effect on ratings at all times and that other attributes show consistent positive (negative) effects with small changes of importance over the years. Our study provides managerial implications of what attributes hotel managers need to maintain or improve for customer satisfaction.
Understanding Important Hotel Attributes from the Consumer Perspective Over Time

1. Introduction

When consumers make purchase decisions, they evaluate products by considering many product attributes. For example, when consumers book a hotel, they may consider price, hotel stars, location, etc. To promote consumer purchase and enhance satisfaction, it is important to determine what product attributes consumers evaluate and how important to satisfaction each of these attributes is. However, American Customer Satisfaction Index recently reported that the average of customer satisfaction in the hotel industry is still below the superior customer satisfaction level (Vanamburg, 2016). It indicates that the hotel industry needs to investigate more accurate evaluation on attributes that influence consumers’ satisfaction. The investigation is critical because there has been a gap between the perceptions of hotel service providers and consumers (Nasution and Mavondo, 2008; Tsang and Qu, 2000). Therefore, it is worth identifying attributes and relationships between attributes and consumer satisfaction to improve customer satisfaction.

Product reviews can provide a wealth of information to hotel managers regarding important elements of customers’ evaluations of their experiences (Zhang and Verma, 2017). According to the TripAdvisor research, 77% of travelers usually or always refer to online reviews before choosing a hotel and 53% would not make a hotel reservation until they read online reviews (TripAdvisor, 2014). However, research on product reviews has used numeric ratings only (e.g., Xie et al. 2014) or has not fully explored text reviews. For example, Zhang and Verma (2017) identified words frequently mentioned in positive and negative reviews but did not investigate how much these words were related to review ratings regarding consumer experiences.

In addition, it is possible that consumer needs and interests change over time. For example, free WiFi has become more important as consumers constantly use their mobile devices. Competition in the hotel industry is becoming fierce, and thus hotels should perhaps shift their focus on attributes over time to cater to consumers’ changing needs and to compete with other hotel providers (Min et al., 2002). However, research on the trend of product attributes in reviews is even less explored in this dynamically changing environment. These gaps
of utilizing product reviews limit our understanding of consumer evaluations and possible appropriate actions for managers.

This study aims to investigate hotel attributes in relation to consumer satisfaction and the trend of the relationship over years. We have two research questions to address:

RQ1. *From the perspective of consumers, what are important hotel attributes that determine review ratings?*

RQ2. *Does the importance of the hotel attributes that impact review ratings change over time?*

To address these questions, we utilize online hotel reviews from 2011 to 2016 on the TripAdvisor website, a third party review provider in which customers who stayed at the hotels left numeric ratings and text reviews. First, we identify all the attributes customers mentioned in the text reviews and then choose the top 30 attributes that affect customer ratings the most. Last, by examining the relationships between customer ratings and these attributes by year, we determine whether the importance of these attributes has changed.

Our paper contributes in two ways. First, methodologically, we show how researchers and firms can identify attributes and determine their importance on ratings by utilizing product reviews. This approach can determine numerous product attributes from the perspective of customers, and the longitudinal data enable tracking the trend of important attributes over time. Second, the substantive contribution is that we provide hotel managers with attribute management guidelines. Hotel managers can maintain the hotel attributes with which consumers are satisfied, such as view, space, service, modern style, and quietness, and improve the hotel attributes with which consumers are dissatisfied, such as Internet/WiFi, wall, reservation, and room, to enhance service quality and customer satisfaction.

2. Literature Review

2.1. Hotel attributes

Once consumers take into consideration particular hotels, hotel attributes will determine their final choices. Thus, it is essential for hotel managers to understand hotel attributes that are closely related to customer satisfaction. Researchers have determined that decisive hotel attributes for consumers to evaluate quality include price, hotel environment, cleanliness, service, attractiveness, relaxation opportunities, loyalty program, guest experience, promotions, amenities,
conventions, green program, hotel image, and hotel reputation (Ananth et al., 1992; Cadotte and Turgeon, 1988; King et al., 2014; Teixeira and Kornfeld, 2013). Tsaur and Tzeng (1996) emphasized that when customers seek lodging, they first consider accommodation conditions (e.g., a comfortable and quiet room) and then service quality factors; the researchers also ranked the attributes that are related to customer satisfaction with hotel service quality.

Despite the importance of understanding focal attributes, substantial previous studies found that the attributes hotel managers focused on were different from the attributes that consumers actually considered important. Coyle and Dale (1993) found that decoration, facilities, and cleanliness, were less important to consumers than what hotel managers thought. Furthermore, competitive staff was a more important attribute to hotel managers than to consumers. Tsang and Qu (2000) found that managers overestimated consumers’ expectations in terms of some attributes regarding hotel service delivery, such as cleanliness/quietness/security of room, attractive decoration/furnishings of room/lobby, and comfortable mattress and pillow. On the other hand, managers underestimated consumers’ expectations in terms of attributes such as availability of year-round swimming pool and hotel staff with multilingual skills. Possible reasons for these discrepancies are self-assured hotel management (Coyle and Dale, 1993; Saleh and Ryan, 1991; Tsang and Qu, 2000), insufficient contact with consumers and inaccurate assessment of consumers’ needs (Coyle and Dale, 1993), and poor communication (Nasution and Mavondo, 2008). Even though the average of hotel industry customers’ satisfaction has been increased over time, it is still under the superior customer satisfaction level (Vanamburg, 2016), suggesting improvement for customer satisfaction toward hotels. Accurate assessment of consumers’ needs will contribute to hotel managers’ better understanding of their customers, resulting in improvement of customer satisfaction and hotel performance.

2.2. **Attribute evaluation**

To evaluate important product/service attributes that determine consumer evaluation, literature has suggested two main approaches: compositional approaches (e.g., Fishbein’s (1976) expectancy-value theory) and decompositional approaches (e.g., conjoint analysis method). Unlike compositional approaches, which are common in analyzing an existing decision process, decompositional approaches are useful to redesign the consumer decision process for reassessing important product attributes (Sweeney et al., 1978). Conjoint analysis, popularized in marketing
literature (Anderson and Bettencourt, 1993; Green and Srinivasan, 1990), is a technique to identify the degree of importance of multi-attributes of a product or service by gathering consumers’ preferences for a set of attribute combination profiles and to predict consumers’ preferences for new profiles (Oppewal and Vriens, 2000). Our approach, similar to conjoint analysis, examines the relationship between ratings that represent consumer satisfaction and hotel attributes. The merits of our approach, compared to conjoint analysis, are that we deal with consumer evaluation based on a large volume of actual transactions and the attributes are determined by consumers, not by researchers.

Another important topic in evaluating attributes is the possible change in attribute importance over time. Flint et al. (1997) proposed that changes in consumers’ evaluations of the value received and/or changes in what they expected from attributes could influence the level of disconfirmation consumers perceive and, thus, change satisfaction levels. They believed that it would be helpful for firms to understand how consumers’ perceptions of value change, which would improve competition and customer retention. Bolton and Drew (1991) found that a service change could influence individual customers’ ratings of the components of service quality, and, eventually, individual customer ratings could be used to predict the effects of the service change in the long run. Min and Min (2006) claimed that understanding the dynamic trend of important hotel attributes is a fundamental way to maintain consumer retention in the long run. Min et al. (2002) conducted a study to capture changes from the consumer perspective in terms of important hotel service attributes over time. For example, the degree of importance of courtesy of employees, handling of complaints, and convenience of reservation were much greater in 1995 compared to what they were in 2000, but price and hotel/tour guide were more important in 2000 than they were in 1995. We examine the change of attribute importance to understand consumer needs over time, another merit of our approach compared to conjoint analysis that only considers data at the time of the research.

2.3. Online reviews

Online reviews are accessible instantly for a great number of consumers who search for others’ opinions about products and services (King et al., 2014). Moreover, a strong stream of online reviews provides real-time information for people to continuously collect, which is not possible for questionnaires or surveys (Zhang and Verma, 2017). Research has found that online
reviews have an impact on customer attitudes such as willingness-to-pay (Pavlou and Dimoka, 2006) and trust and loyalty to a product (Awad and Ragowsky, 2008); determine a customer’s purchase decisions (Book et al., 2015; Ladhari and Michaud, 2015); and eventually influence a firm’s financial performance, such as box office revenue (Liu, 2006), the sale of entertainment goods (Dellarocas et al., 2007), and stock prices (Tirunillai and Tellis, 2012). Because of these characteristics, consumers actively use online reviews to exchange ideas and make decisions, and firms take online reviews as important references to understand consumer search and satisfaction.

Product reviews typically consist of numeric ratings and text opinions. Numeric ratings are much easier to read and utilize (Viglia et al., 2016) and help readers reduce the interpretation issues regarding the valence of consumer reviews (Chevalier and Mayzlin, 2006). However, numeric ratings cannot articulate consumers’ detailed experiences nor accurately predict outcomes such as sales (King et al., 2014). Kasper and Vela (2011) pointed out that numeric ratings might not indicate customers’ positive or negative attitudes toward specific attributes well.

In contrast, text reviews provide abundant information. Many researchers have focused on text reviews in terms of sentiment. Moe and Schweidel (2012) found that products with more positive reviews than negative reviews were more likely to attract consumers. A drawback of text reviews is that reading the overload of information is tough, and simple signaling guidance is needed (Zhao et al., 2015). Thus, it is necessary to simplify text reviews and utilize summarized information for better management. Cadotte and Turgeon (1988) indicated that hotel attributes mentioned in customer reviews could be used to identify key factors that satisfied or dissatisfied consumers. That is, attributes in reviews can reveal the reasons for consumer satisfaction or dissatisfaction. Our research uses a text mining technique to identify hotel attributes as summarized information from text reviews.

3. Data and Measurement

3.1. Online hotel review data

Firms have engaged and utilized third party review providers to collect consumers’ feedback and interact with consumers. Hotel managers use online reviews provided by TripAdvisor.com, Yelp.com, and Hotels.com to deal with consumers’ problems and increase hotel performance (King et al., 2014). This study chose TripAdvisor, which is one of the most visited travel websites and provides millions of reviews on hotels. We selected hotels in Chicago,
IL, where there are over 168 hotels in the different hotel classes. In 2016, Chicago attracted more than 54 million travelers, domestic and international, which awarded it one of the top visited cities in the U.S. (Koziarz, 2017). As a result, we were able to gather a sufficient number of reviews from the consumers that stayed at the different levels of hotels.

We used Java programming language and Jsoup as a library of Java for automatically collecting hotel reviews. Jsoup is a publicly accessible project and provides an application programming interface (API) for parsing HTML from a specific URL for extracting data. We first made a list of links to hotels in Chicago by searching for “Chicago” on the website. Each link is connected to hotel information and consumer reviews of a specific hotel. Then, we wrote Java codes to call the Jsoup library to parse hotel links, as well as all of the elements for each hotel. We parsed the elements with HTML labels such as <div>, <p>, and <li> from the website through Jsoup. The elements we parsed and extracted included hotel brands, dates consumers stayed at the hotel, review titles, text reviews, and review ratings. In total, we obtained 175,268 reviews from 149 hotels in Chicago from 2011 to 2016.

3.2. Attribute measurement

The major independent variables were hotel attributes. We used a text mining module in Stata (i.e., txttool). Specifically, we removed punctuation marks and stopping words, reduced all words to their stems, and finally displayed the review sentences as a bag of words. Because a substantial number of reviewers might give no response to some attributes (Zhang and Verma, 2017), we selected words mentioned at least 500 times, as an initial cut-off point, and obtained 474 words. Then, we chose 126 words that were highly related to hotel attributes, such as rooms and staff. Next, we regressed review ratings on these 126 hotel attributes and selected the top 30 significant hotel attributes in the whole year data using the stepwise selection option. We chose 30 attributes to provide hotel managers a wide but reasonable range of attributes. Finally, we regressed review ratings on these 30 hotel attributes by year and determined their importance using standardized coefficients. Figure 1 explains the process of extracting hotel attributes from text reviews.

< Figure 1 >

3.3. Descriptive statistics and preliminary correspondence analysis
Table 1 presents the descriptive statistics of ratings over time. The mean of review ratings in each year was around 4.25 out of 5 with standard deviations of around 1, which remained the same over the years. The number of reviews increased gradually over the six years (i.e., there were 13,678 reviews in 2011 and 40,648 in 2016).

Table 1

Figure 2 depicts the results of corresponding analysis, commonly used in text mining analysis (e.g., Lee and Bradlow, 2011), based on the cross-tabulation of years and attributes. The first two dimensions explained 77% and 8.6% of the total information, respectively. In the year plot, the main x-axis from the right to the left side shows the timeline of years from 2011 to 2016 and the minor y-axis from the top to the bottom side represents how far the years are from the middle of the observation years. In the attribute plot, the main x-axis from the right to the left side represents services- vs. room-related attributes while the minor y-axis from the top to the bottom side roughly represents external vs. internal attributes. Years and attributes in the same quadrant in each plot represent that they are correlated. Years 2011 and 2012 (the 1st quadrant) are correlated with room, river, corner, price, decoration, suite, reservation, and treatment, as consumers frequently mentioned these attributes. In 2013 and 2014 (the 4th quadrant), consumers mentioned water, services, lobby, shop, free breakfast/parking, luxury, smile, and Internet/WiFi more frequently. In 2015 (the 3rd quadrant), park, bathroom, walk, space, staff, quiet, door, renovation, noise, and amenity were attributes that consumers mentioned frequently. In 2016 (the 2nd quadrant), consumers more frequently talked about wall, clean, view, and modern. The results show that years and attributes can be correlated, implying that the importance of attributes could vary with years.

Figure 2

4. Model

By examining the effects of hotel attributes on review ratings, we can understand which attributes are related to review ratings. Later, we determine the importance of each attribute on ratings with standardized coefficients. To begin, we set up a panel regression model as follows.

$$\text{Rating}_{ij} = \beta_0 + \sum_{k=1}^{K} \beta_k \text{Attribute}_{ijk} + \sum_{l=1}^{L} \beta_l \text{Month}_{ijl} + \varepsilon_{ij},$$

where hotel $i=1,\ldots,I$ and review $j=1,\ldots,J$. 
The dependent variable \textit{Rating} is the numeric rating of a specific review. The independent variable \textit{Attribute} indicates the frequency a specific attribute is mentioned in the review. \textit{Month} is a monthly dummy variable that controls monthly effects. The parameter $\beta_0$ represents the intercept of the model. The parameter $\beta_k$ represents the effects of the frequency of attributes on ratings. The parameter $\beta_l$ means the monthly effects. Finally, $\varepsilon_{ij}$ is the error term. Due to the intragroup correlation, the error terms $\varepsilon_{ij}$ and $\varepsilon_{ij'}$ (where $j \neq j'$ for hotel $i$) are not independent. Thus, we use robust cluster standard errors, in which hotels are clusters. After identifying the top 30 attributes, we run the regression model by year.

5. Results

We present the estimation results calculated based on the accumulated data for six years regarding top 30 attributes in Table 2, which shows the importance ranks, the top 30 variables, the words co-occurring with each of the top 30 variables, and the estimation coefficients along with robust cluster standard errors and standardized coefficients, Beta. We decide the ranks of variables based on the absolute values of standardized coefficients from the entire data. The coefficients mean that the frequencies of attributes are associated with review ratings. One or two co-occurring words are selected based on the highly positive correlation coefficients between each attribute and all other attributes, which makes multiple appearances of a certain word possible. The co-occurring words are useful to interpret the meaning of the coefficients in the condition that we do not know the valence of each word. The coefficient of determination ($R^2$) is 9.2%.

\begin{center}< Table 2 >\end{center}

The top 30 variables address the first research question regarding hotel attributes from the perspective of consumers. The most important attribute is staff (Beta=0.14), followed by room (Beta=-0.11) and services (Beta =0.07). As staff and services are related to front desk, consumers seem to be satisfied with services provided by front desk staff. However, the negative effect of room on ratings and its co-occurring words imply that consumers are probably not satisfied with cleanliness or bed conditions in the rooms. Other attributes in the top 10 are also room related. Room space, view, quietness, and modernity are positively associated with product ratings, while reservation staff, wall placement, and Internet/WiFi are negatively associated with product ratings. The 10 most important attributes clearly articulate that staff services and room quality,
the core lodging elements, are important from the perspective of consumers. Therefore, hotel managers need to pay attention to these attributes as they could effectively increase or badly affect ratings. Notably, many of these attributes, such as professional staff and spaciousness of a room, are not offered as filtering options in hotel reservation sites or descriptions on hotel websites. That is, attributes discussed by consumers can be different from those used by hotel management.

The next 10 most important attributes are a mixture of room-related attributes and supplementary services. Regarding room-related attributes, clean and suite are positive drivers on ratings, but bathroom and noise are negative drivers. Among supplementary services, lobby has a negative effect on ratings, while treatment, free (breakfast or parking), walk (to near places), and decoration (in lobbies and rooms) have positive effects on ratings. It is interesting that supplementary service related attributes are ranked after core service related attributes. Though it has a negative effect as expected, price is ranked in twentieth place. A possible reason for this low importance could be that after a customer has stayed at a hotel, price is not as important as when the customer searched for hotels. Another reason could be that because customers can determine a price that fits their budgets, the impact of price can be less than the impacts of options that customers cannot determine (e.g., front desk services).

Among the last 10 important attributes, many are related to the environment outside hotels. Shop, river, water, and park are closely related to the environment of Chicago, a metropolitan city next to Lake Michigan. The frequencies of water and park have a negative effect on review ratings, possibly because these consumers were not satisfied with hotel conditions related to water and park (e.g., lack of access to the lake or parks). In addition, some supplementary service attributes (i.e., corner, smile, and room amenity) have a positive effect on ratings, while interior related attributes (i.e., luxurious, door, and renovation) have mixed effects.

To address the second research question about the trend of attribute importance, we compare the standardized coefficients (Beta) of attributes over the years. The yearly estimation results of top 10 attributes selected from the accumulated data analysis are presented in Table 3, and the rest 20 attributes are omitted but available on request. The top 10 attributes include hotel staff, hotel rooms (associated with cleanliness and beds), hotel services, room/bathroom spaces, hotel views, reservation services, room/bathroom walls, Internet/WiFi, quietness, and modernity. The coefficients of determination ($R^2$) are around 10% across the models. Notably, the effect
signs of hotel attributes are constant over time. For example, the effect of staff on hotel ratings has always been positive, while the effect of room has always been negative in 2011 through 2016. We test whether the standardized coefficients of each attribute change over time (Beta difference test) by examining overlapped confidence intervals of the highest and lowest Betas, a test approach suggested by Cumming (2009).

< Table 3 >

In Figure 3, we depict the importance rank changes with the significance flags. The major findings are summarized as follows. First, staff is the most important and uniquely salient attribute across the years in that the average of Beta is 0.14, and the change in Beta values over time is not significant (POL\(^1\)=0.87, p>0.1). This positive effect of staff is twice as important as the positive effects of the other attributes. This suggests that hotels should invest more in human capital than in anything else.

Second, services, space, view, quietness, and modernity form an attribute importance group with positive effects. The importance averages of these attributes are around 0.05 with narrow fluctuations. The change in Beta values over time for space is significant at the 0.05 level (POL=0.56, p<0.05) and the changes in Beta values over time for view and quietness are significant at the 0.1 level (POL=0.81 and 0.80, respectively, p<0.1), implying that the importance of space, view, and quietness statistically changed over the years. However, the changes in Beta values over time for service and modernity are not significant (POL=1.10 and 1.40, respectively, p>0.1).

Third, four attributes (i.e., room, reservation, wall, and Internet/WiFi) form another attribute importance group with negative effects. The importance averages of these attributes except for room are around -0.06, while room has almost twice the negative effect on ratings compared to the other negative attributes. Notably, the importance levels of room, wall, and Internet/WiFi fluctuated during the observation period with a pattern of being slightly less important. The importance of reservation became the least important in 2013, but it has become stronger since then. The changes in Beta values over time for room, reservation, and Internet/WiFi are significant at the 0.05 level (POL=0.19, -0.57, 0.14, respectively, p<0.05) while the change in Beta values over time for wall is not significant (POL=1.14, p>0.1). The

\(^1\) POL stands for proportion overlap, which is a ratio of the overlapped interval length to the average single arm length of the confidence intervals for two different Betas (Cumming 2009).
consistent negative effects indicate that hotel managers should take quick action because improving these attributes will lead to higher customer satisfaction.

< Figure 3 >

6. Conclusions

6.1. Summary

The purpose of this study is to examine important hotel attributes from the perspective of consumers over time. By applying a text mining technique to consumer reviews of hotels in Chicago over six years, we address two important research questions: what are important attributes in determining review ratings and has the importance of these attributes changed over time. We examine the relationships between numeric ratings and hotel attributes and select the 30 most important attributes based on standardized coefficients. Interestingly, the importance orders of the attributes are closely related to core services and room quality, supplementary services, and attributes outside hotels.

The comparison of the effects of attributes in yearly data reveals that staff, who are the primary service conductors, are consistently the most important in the hotel industry. Among the attributes with positive effects, space, view, and quietness, which are related to rooms, have shown significant importance changes in affecting review ratings over the years, while staff, services related to the front desk, and modernity have not. The attributes with negative effects show fluctuations over the years. Reservation has become more important in determining review ratings after the decline in 2013, while room conditions related to cleanliness and beds and Internet/WiFi have become less important after the increase in 2014 and 2012, respectively. The important attributes from the perspective of consumers and the changes in their importance may require the attention of hotel managers to improve information provision and service quality, which we discuss in detail in the next subsection.

6.2. Managerial implications

Our findings provide managerial implications as to which attributes hotels need to maintain and improve and how hotels or third party reservation websites should provide attribute information. Our outcomes are based on the effects of attributes on ratings, which represent consumer satisfaction, and we suggest that improving these attributes can enhance consumer
satisfaction. Although attributes found important in consumer reviews are included in filtering options and hotel descriptions on websites, not all of the important attributes are included. From the perspective of consumers, attributes related to core services, service conductors, and room quality are important and filtering options and descriptions on relevant websites should include these attributes. Hotel managers can also put more emphasis on the attributes frequently mentioned by consumers when they provide descriptive information regarding their hotels and conduct customer satisfaction surveys. Attributes with negative effects indicate that consumers are not satisfied with these attributes. Therefore, hotel managers need to pay attention to those negative attributes (e.g., rooms may have problems with cleanliness or Internet/WiFi).

Hotel managers can also utilize the trends of attribute importance over the years. In our data, the signs of positive and negative effects do not change over time. Regarding attribute importance that influences ratings, some attributes, such as staff and reservation, are consistently important or become more important after a certain year, suggesting that hotel managers need to pay more attention to these attributes. In contrast, other attributes, such as Internet/WiFi and room, become less important gradually or after a certain year, implying that hotel managers have improved these attributes and that managers may pay less attention to them. That is, the changes of attribute importance suggest that hotel managers should regularly investigate attributes that become more (less) important to avoid underinvestment (overinvestment) in these attributes.

6.3. Limitations and future research

We present several limitations that need future research attention. First, it is necessary to examine attribute sentiment rather than attribute frequency. A limitation of attribute frequency is that frequency does not indicate the valence of attributes, but sentiment can deal with this issue. A more advanced text analysis should be applied to determine the tones of attributes. Second, it is necessary to examine hotel attributes by hotel class. Hotels falling into the same hotel star class may have similar important attributes. By considering hotels within specific classes, hotel managers could improve customer satisfaction with attributes that are more appropriate to their hotel class. Third, our study identifies the important hotel attributes based on online reviews. However, not all hotel customers write online reviews. Those consumers who do not write online reviews may have different opinions and evaluation criteria that need to be considered in future research.
References


Table 1. Descriptive Statistics of Ratings over Years

<table>
<thead>
<tr>
<th>Year</th>
<th>2011</th>
<th>2012</th>
<th>2013</th>
<th>2014</th>
<th>2015</th>
<th>2016</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number</td>
<td>13,678</td>
<td>21,259</td>
<td>26,944</td>
<td>33,958</td>
<td>38,781</td>
<td>40,648</td>
<td>175,268</td>
</tr>
<tr>
<td>S.D.</td>
<td>0.996</td>
<td>0.954</td>
<td>0.959</td>
<td>0.997</td>
<td>1.014</td>
<td>1.009</td>
<td>0.993</td>
</tr>
</tbody>
</table>

The numbers are based on 149 hotels.
### Table 2. Estimation Results of the Accumulated Data regarding Top 30 Attributes

<table>
<thead>
<tr>
<th>Rank</th>
<th>Variable</th>
<th>Co-occurring Words</th>
<th>Coef.</th>
<th>S.E.</th>
<th>Beta</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Staff</td>
<td>Front desk</td>
<td>0.28</td>
<td>0.01</td>
<td>0.14</td>
</tr>
<tr>
<td>2</td>
<td>Room</td>
<td>Clean, Bed</td>
<td>-0.14</td>
<td>0.01</td>
<td>-0.11</td>
</tr>
<tr>
<td>3</td>
<td>Service</td>
<td>Front desk</td>
<td>0.17</td>
<td>0.02</td>
<td>0.07</td>
</tr>
<tr>
<td>4</td>
<td>Space</td>
<td>Bathroom, Room</td>
<td>0.31</td>
<td>0.02</td>
<td>0.07</td>
</tr>
<tr>
<td>5</td>
<td>View</td>
<td>Floor, Room</td>
<td>0.19</td>
<td>0.02</td>
<td>0.06</td>
</tr>
<tr>
<td>6</td>
<td>Reservation</td>
<td>Front desk</td>
<td>-0.22</td>
<td>0.02</td>
<td>-0.07</td>
</tr>
<tr>
<td>7</td>
<td>Wall</td>
<td>Room, Bathroom</td>
<td>-0.64</td>
<td>0.06</td>
<td>-0.06</td>
</tr>
<tr>
<td>8</td>
<td>Internet/WiFi</td>
<td>Breakfast, Room</td>
<td>-0.32</td>
<td>0.02</td>
<td>-0.06</td>
</tr>
<tr>
<td>9</td>
<td>Quietness</td>
<td>Room, Bed</td>
<td>0.23</td>
<td>0.02</td>
<td>0.04</td>
</tr>
<tr>
<td>10</td>
<td>Modernity</td>
<td>Room, Bathroom</td>
<td>0.22</td>
<td>0.02</td>
<td>0.04</td>
</tr>
<tr>
<td>11</td>
<td>Clean</td>
<td>Bed</td>
<td>0.10</td>
<td>0.01</td>
<td>0.04</td>
</tr>
<tr>
<td>12</td>
<td>Suite</td>
<td>View, Reservation</td>
<td>0.14</td>
<td>0.02</td>
<td>0.04</td>
</tr>
<tr>
<td>13</td>
<td>Lobby</td>
<td>Bar, Area</td>
<td>-0.12</td>
<td>0.02</td>
<td>-0.03</td>
</tr>
<tr>
<td>14</td>
<td>Treatment</td>
<td>Staff, Service</td>
<td>0.31</td>
<td>0.02</td>
<td>0.03</td>
</tr>
<tr>
<td>15</td>
<td>Free</td>
<td>Breakfast, Parking</td>
<td>0.16</td>
<td>0.02</td>
<td>0.04</td>
</tr>
<tr>
<td>16</td>
<td>Walk</td>
<td>Distance, Shop</td>
<td>0.12</td>
<td>0.01</td>
<td>0.05</td>
</tr>
<tr>
<td>17</td>
<td>Bathroom</td>
<td>Clean</td>
<td>-0.14</td>
<td>0.03</td>
<td>-0.04</td>
</tr>
<tr>
<td>18</td>
<td>Noise</td>
<td>Floor, Room</td>
<td>-0.28</td>
<td>0.03</td>
<td>-0.04</td>
</tr>
<tr>
<td>19</td>
<td>Decoration</td>
<td>Lobby, Room</td>
<td>0.15</td>
<td>0.02</td>
<td>0.03</td>
</tr>
<tr>
<td>20</td>
<td>Price</td>
<td>Location</td>
<td>-0.11</td>
<td>0.02</td>
<td>-0.03</td>
</tr>
<tr>
<td>21</td>
<td>Shop</td>
<td>Restaurant, Location</td>
<td>0.16</td>
<td>0.02</td>
<td>0.04</td>
</tr>
<tr>
<td>22</td>
<td>Corner</td>
<td>View, Floor</td>
<td>0.18</td>
<td>0.03</td>
<td>0.03</td>
</tr>
<tr>
<td>23</td>
<td>Smile</td>
<td>Front desk, Staff</td>
<td>0.33</td>
<td>0.03</td>
<td>0.02</td>
</tr>
<tr>
<td>24</td>
<td>River</td>
<td>View, Walk</td>
<td>0.12</td>
<td>0.02</td>
<td>0.03</td>
</tr>
<tr>
<td>25</td>
<td>Luxurious</td>
<td>Bathroom, Service</td>
<td>0.23</td>
<td>0.04</td>
<td>0.03</td>
</tr>
<tr>
<td>26</td>
<td>Door</td>
<td>Front desk</td>
<td>-0.16</td>
<td>0.02</td>
<td>-0.03</td>
</tr>
<tr>
<td>27</td>
<td>Water</td>
<td>Place, Block</td>
<td>-0.18</td>
<td>0.05</td>
<td>-0.03</td>
</tr>
<tr>
<td>28</td>
<td>Park</td>
<td>Walk, Distance</td>
<td>-0.07</td>
<td>0.02</td>
<td>-0.02</td>
</tr>
<tr>
<td>29</td>
<td>Amenity</td>
<td>Services, Room</td>
<td>0.15</td>
<td>0.02</td>
<td>0.02</td>
</tr>
<tr>
<td>30</td>
<td>Renovation</td>
<td>Lobby, Room</td>
<td>-0.16</td>
<td>0.03</td>
<td>-0.02</td>
</tr>
</tbody>
</table>

Coefficients of Determination ($R^2$) 0.092

1. Coefficients are significant at the 0.05 level, and standard errors are robust cluster standard errors.
2. Data are accumulated for six years from 2011 to 2016.
Table 3. Yearly Estimation Results (Presenting Top 10 Attributers Only)

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Staff</td>
<td>0.28</td>
<td>0.02</td>
<td>0.135</td>
<td>0.26</td>
<td>0.02</td>
<td>0.137</td>
<td>0.25</td>
<td>0.02</td>
<td>0.132</td>
<td>0.29</td>
<td>0.01</td>
<td>0.148</td>
<td>0.27</td>
<td>0.01</td>
<td>0.141</td>
<td>0.27</td>
<td>0.01</td>
<td>0.143</td>
<td>0.87</td>
</tr>
<tr>
<td>2</td>
<td>Room</td>
<td>-0.15</td>
<td>0.02</td>
<td>-0.115</td>
<td>-0.14</td>
<td>0.01</td>
<td>-0.112</td>
<td>-0.14</td>
<td>0.01</td>
<td>-0.112</td>
<td>-0.17</td>
<td>0.01</td>
<td>-0.131</td>
<td>-0.13</td>
<td>0.01</td>
<td>-0.101</td>
<td>-1.3</td>
<td>0.01</td>
<td>-0.104</td>
<td>0.19**</td>
</tr>
<tr>
<td>3</td>
<td>Service</td>
<td>0.17</td>
<td>0.03</td>
<td>0.064</td>
<td>0.16</td>
<td>0.02</td>
<td>0.067</td>
<td>0.18</td>
<td>0.02</td>
<td>0.073</td>
<td>0.19</td>
<td>0.02</td>
<td>0.078</td>
<td>0.18</td>
<td>0.02</td>
<td>0.071</td>
<td>0.16</td>
<td>0.02</td>
<td>0.064</td>
<td>1.10</td>
</tr>
<tr>
<td>4</td>
<td>Space</td>
<td>0.30</td>
<td>0.04</td>
<td>0.064</td>
<td>0.25</td>
<td>0.03</td>
<td>0.055</td>
<td>0.27</td>
<td>0.02</td>
<td>0.063</td>
<td>0.32</td>
<td>0.03</td>
<td>0.070</td>
<td>0.29</td>
<td>0.02</td>
<td>0.066</td>
<td>0.30</td>
<td>0.02</td>
<td>0.068</td>
<td>0.56**</td>
</tr>
<tr>
<td>5</td>
<td>View</td>
<td>0.16</td>
<td>0.04</td>
<td>0.053</td>
<td>0.17</td>
<td>0.03</td>
<td>0.059</td>
<td>0.18</td>
<td>0.03</td>
<td>0.062</td>
<td>0.21</td>
<td>0.02</td>
<td>0.073</td>
<td>0.18</td>
<td>0.02</td>
<td>0.064</td>
<td>0.20</td>
<td>0.02</td>
<td>0.073</td>
<td>0.81*</td>
</tr>
<tr>
<td>6</td>
<td>Reservation</td>
<td>-0.20</td>
<td>0.04</td>
<td>-0.069</td>
<td>-0.17</td>
<td>0.03</td>
<td>-0.061</td>
<td>-0.12</td>
<td>0.03</td>
<td>-0.042</td>
<td>-0.24</td>
<td>0.02</td>
<td>-0.079</td>
<td>-0.25</td>
<td>0.03</td>
<td>-0.078</td>
<td>-0.29</td>
<td>0.03</td>
<td>-0.085</td>
<td>-0.57**</td>
</tr>
<tr>
<td>7</td>
<td>Wall</td>
<td>-0.71</td>
<td>0.16</td>
<td>-0.060</td>
<td>-0.63</td>
<td>0.14</td>
<td>-0.058</td>
<td>-0.75</td>
<td>0.11</td>
<td>-0.066</td>
<td>-0.66</td>
<td>0.09</td>
<td>-0.057</td>
<td>-0.57</td>
<td>0.06</td>
<td>-0.053</td>
<td>-0.64</td>
<td>0.10</td>
<td>-0.058</td>
<td>1.14</td>
</tr>
<tr>
<td>8</td>
<td>Internet/WiFi</td>
<td>-0.31</td>
<td>0.06</td>
<td>-0.061</td>
<td>-0.44</td>
<td>0.05</td>
<td>-0.082</td>
<td>-0.32</td>
<td>0.04</td>
<td>-0.062</td>
<td>-0.27</td>
<td>0.03</td>
<td>-0.052</td>
<td>-0.34</td>
<td>0.05</td>
<td>-0.063</td>
<td>-0.33</td>
<td>0.03</td>
<td>-0.055</td>
<td>0.14**</td>
</tr>
<tr>
<td>9</td>
<td>Quietness</td>
<td>0.28</td>
<td>0.04</td>
<td>0.048</td>
<td>0.20</td>
<td>0.03</td>
<td>0.036</td>
<td>0.23</td>
<td>0.03</td>
<td>0.045</td>
<td>0.25</td>
<td>0.03</td>
<td>0.045</td>
<td>0.23</td>
<td>0.03</td>
<td>0.044</td>
<td>0.22</td>
<td>0.02</td>
<td>0.041</td>
<td>0.80*</td>
</tr>
<tr>
<td>10</td>
<td>Modernity</td>
<td>0.24</td>
<td>0.04</td>
<td>0.045</td>
<td>0.20</td>
<td>0.04</td>
<td>0.039</td>
<td>0.20</td>
<td>0.04</td>
<td>0.040</td>
<td>0.22</td>
<td>0.03</td>
<td>0.045</td>
<td>0.22</td>
<td>0.02</td>
<td>0.046</td>
<td>0.23</td>
<td>0.03</td>
<td>0.048</td>
<td>1.40</td>
</tr>
</tbody>
</table>

Coefficients of Determination (R²): 0.092, 0.083, 0.081, 0.105, 0.093, 0.103

1. Coefficients are significant at the 0.05 level, and standard errors are robust cluster standard errors.
2. Beta difference test is based on the test approach by Cumming (2009) and uses the lowest and highest Betas for each attribute. POL stands for proportion overlap, which is a ratio of the overlapped interval length to the average single arm length of the confidence intervals for two different Betas.
3. * (**) means that the highest and lowest Betas are significantly different at the 0.1 (0.05) level.
Figure 1. Attribute Extraction Process

- Hotel Data Collection
- Word Splitting
- Word Frequency Counting
- Attribute Selection with 500+ Word Frequency
- Top 30 Significant Attribute Selection
- Linear Regression:
  DV: review ratings; IV: hotel attributes
Figure 2. Correspondence Analysis Results Using Years and Attributes

A. Year Plot

B. Attribute Plot
Figure 3. Importance Changes of Top 10 Important Variables from 2011 to 2016

1. The x-axis represents the years, and the y-axis represents the importance of the attributes on product ratings.
2. * (**) means that the highest and lowest Betas are significantly different at the 0.1 (0.05) level.