Which Group Do You Want to Travel with?  
(A Study of Rating Differences among Groups on Online Travel Reviews)

Daehwan Ahn  
College of Business Administration, Seoul National University  
599 Gwanakno Daehakdong  
Gwanakgu, Seoul, Korea  
+82-2-880-2550  
pennis32@snu.ac.kr

Hyanghee Park  
College of Communication, Seoul National University  
599 Gwanakno Daehakdong  
Gwanakgu, Seoul, Korea  
+82-2-880-2550  
hyanghee@snu.ac.kr

Byungjoon Yoo  
Graduate School of Business, Seoul National University  
599 Gwanakno Daehakdong  
Gwanakgu, Seoul, Korea  
+82-2-880-2550  
byoo@snu.ac.kr

ABSTRACT

The purpose of this paper is to empirically examine that which group the travelers are travelling with can have a significant impact on travelers’ satisfaction with the hotel they visited to. The data was crawled and collected on Booking.com which is the most popular and well-known travel website by using our web crawler developed in Python. We analyzed 314 hotels rating data of two to five star hotels located in New York City conducting econometric analysis. Consequently, it is discovered that satisfaction of traveler groups decreases in the order of couple, friends, family, solo, and business. The group of couples expressed the highest satisfaction while the group of business travelers showed the lowest satisfaction. By conducting text analysis with 125,076 reviews, we found that such satisfaction differences can be caused by differences of travelers’ experiences depending on groups.

CCS Concepts

Information systems → World Wide Web → Web applications
→ Social networks

Keywords

Online Reviews; Online Travel; Smart Tourism; Hotel Rating; Group Difference

1. INTRODUCTION

A convergence of social media and traveling is changing travelers’ trip experiences dramatically. According to a study, it is found that travelers’ behavior of sharing their experiences and information not only increases the convenience, but dramatically changes their way of searching for destinations and their experiences (Lewis et al., 2010). Consequently, Social media plays an important role in carrying out marketing strategies related to travelling decision making. Moreover, by increasing the efficiency (Gretzel et al. 2006), it became the most important channel in the travel industry.

Travel websites based on social media play a significant role in sharing travelers’ experiences and travel information as a platform. Also, the travel websites are used frequently to make a reservation on airline tickets, hotels, and travel related products. Such phenomenon indicates that it shows a dramatic growth in travel websites and other related markets. In 2011, revenue occurred in online travel reservation services reached approximately 340 billion dollars and it is reported that about 39% of Americans make a reservation through online travel websites (Barclays Capital, 2012). Moreover, Booking.com which is the No.1 online travel website has monthly unique visitors reaching 40 million in 2015 (“Top 15 most popular”, 2015). There is not much research on travel websites based on social media although the importance of such research increases. The majority of the research focuses on CB studies related with the factors affecting on travelers’ satisfaction, motivations, and decision making (Cohen et al., 2014). Moreover, such past research are not conducted by using secondary behavioral data of real users but mostly based on experiments and surveys. Thus, in this study, we proceed our research using real users’ behavioral data based on social network based websites. In our research, we collected reviews and rating scores of each hotel from Booking.com which is the No.1 company in the traveling websites industry.

The purpose of this paper is to find out how different compositions of group members can have an effect on their traveling experiences. We developed a web crawler using a programming language called Python in order to collect both the data opened to the public and the detailed data that is not open on the webpage. These data was analyzed using econometric analysis and text analysis.

The remainder of this paper is organized as follows. Chapter 2 reviews the past research related to this study. Chapter 3 includes related research hypotheses. Chapter 4 explains the data collection procedure and description about our data set. Chapter 5 describes the method and model used to compare different travel groups’ satisfaction with hotels, and Chapter 6 presents analysis results and related findings. Finally, discussion on the results, limitations, and future studies are shown in the last chapter.

2. RESEARCH BACKGROUND

2.1 Satisfaction

Satisfaction is regarded as a main CB construct since the extent to which consumers are satisfied can have an effect on future business performance in the form of, for instance, revenue, competitive advantage and market image (Anderson, Fornell, & Lehmann, 1994). By examining satisfaction and its systems, marketers can...
find useful information they can apply in order to increase customers’ satisfaction, by developing strategic decisions such as segmentation, targeting, and marketing mix (Cohen et al., 2014).

Some researchers take more interest in tourist & staff relationships (Hutchinson et al., 2009; Nam et al., 2011; Tsang & Ap, 2007); there is much less research, however, on tourist-tourist interactions as a determinant of satisfaction (Getz et al., 2001; Huang & Hsu, 2010; Wu, 2007). While these studies can give an insight into understanding how interactions between travelers affect satisfaction, there is not an in-depth study yet (Cohen et al., 2014).

Some researchers examined the relationship between tourists by adopting the concepts of personal involvement (Hwang et al., 2005; Prayag & Ryan, 2012). Recently, a self-congruence perspective is adopted in various research (Bosnjak et al., 2011; Nam et al., 2011), while a few studies investigating travelers’ psychological characteristics, such as novelty seeking, as determinants of satisfaction (Petrick, 2002; Williams & Soutar, 2009). Perceiving the complication of tourist satisfaction, some researchers investigate the extent to which satisfaction varies depending on consumer types (Magnini et al., 2011; Petrick & Sirakaya, 2004).

Borges, et al. (2010) find out that, depending on who people are shopping with, the shoppers can feel the values of shopping differently. More interestingly, they discovered the difference of shopping experiences depending on three different groups which are composed of friends, family, and alone respectively. This result is closely related to our research in that travelers feel satisfied with their hotel differently depending on who they are traveling with.

2.2 Rating

There are some past research which focus on factors affecting on travelers’ reviews and scores on each hotel.

Gretzel & Yoo (2008) find out how prices and places can affect the rating and scores of hotels on the website. Sparks & Browning (2011) argue that hotel customers can be affected by the early negative reviews. Gidumal & Gonzalez Lopez-Valcarcel (2011) explains in their study that early reviews of hotels tend to be negative and if the reviews are few, the average ratings are low and it can affect potential customers searching for hotels.

Stringam & Gerdes (2010) identified the most used words when reviewers are rating the hotels they stayed through online hotel reviews. The most used words are as follows; “clean”, “staff”, “breakfast”, “bed”, “price”, “restaurant”, “pool”, “bathroom”, “airport”, “downtown”. Chaves, Gomes & Pedron (2011) also examine similar research and find out that the most frequently shown words are “room”, “staff”, “location”, “cleanliness”, “friendliness” and “helpfulness”. Sridhar & Srinivasan (2012) shows that the reason the ratings of online hotel websites are important is that online reviewers are affected by other opinion leaders’ evaluations.

2.3 Social Media and Travel

Nowadays, it became almost impossible to live without technology since consumers use it in various ways ranging from looking for information, purchasing, sharing their thoughts and experiences to entertainment activities. Buhalis & Law (2008) argue that it is more obvious in the tourism industry because people tend to sell and purchase tour products online.

Xiang & Gretzel (2010) consider social media as an important platform for providing valuable information but for travelers’ experiences. Social media already plays a significant role in all stages of the vacation cycle: before, during and after the trip (Fotis, Buhalis, & Rossides, 2011). Not surprisingly, a number of research began to examine how various social media are affecting tourism CB. In the study by Vermeulen and Seegers (2009), they focus on the impact of online hotel reviews on consumers’ decision, Paphanassiss and Knolle (2011) studied the usage of online reviews and Zehrer et al. (2011) investigate on user responses to travel blog information and recommendations.

Recently, in several studies, it is shown that the reputation resulting from eWOM behavior influences organisational performance variables such as price (Yacouel & Fleischer, 2012). Also, Serra Cantallops & Salvi (2014) examine how eWOM can influence on customers’ intention of reserving hotels on the website. In addition, Vermeulen & Seegers (2009) identify purchase decision process of hotel customers. Thus, travel websites can better understand how social media is affecting the way consumers are perceived, and how such perceptions have an effect on tourists’ decision making and behavior (Cohen et al., 2014).

Some recent research also identify how technology changes the travelers’ experience during the trip, explaining how smartphones affect their experience (Dickinson et al., 2012; Wang et al., 2012).

2.4 Group Decision Making

When traveling, the difference of group members can change decision making process and its consequences. Our study focuses on the difference of average scores on hotels depending on the different compositions of group members. In this context, the past studies on group decision making become an important theoretical background in our study.

Past CB research in tourism typically focuses on individual-level analysis as opposed to households, groups of friends or work colleagues (Cohen et al., 2014). Some research on tourism and CB (Dimanche & Havitz, 1995; Moutinho, 1993) require for a transition in scale to examine family decision-making processes. Campo-Martinez et al. (2010) regard group composition important mentioning that individual satisfaction can be different from that of a bigger travelling group. It is due to the fact that collective satisfaction of group members are more important to revisit the places than that of individuals.

Kozak (2010) finds out that different members of a household are engaged in travel decisions all together, with the specific dynamics dependent on power relations among family members. It is certainly true that families are a ‘decision making unit’, but each family member may seek information individually and, sometimes, may need to negotiate disagreements (Bronner & de Hoog, 2008, p. 967). Thus, understanding how travel decisions are made within families, and how different family members may affect decision making processes is considered as an important factor for marketing strategies (Dimanche & Havitz, 1995; Litvin et al., 2004).

In the study conducted by Hong et al. (2009), multiple decision makers (both partners) are engaged in the decision making process of revisiting the same place repeatedly. Furthermore, Wang et al. (2004) evaluate decision power of children in the process of family vacation decision-making, showing that, in the context of Taiwanese families, children have extremely little impact on choosing the family’s group package tour. In addition, there is only few studies in tourism which identify the influence of other reference groups, such as friends or work colleagues rather than just family or households, on or within travel decision-making (for an exception see Decrop & Snelders, 2004). One other notable
exception is the research of Hsu et al. (2006), who analyze the interpersonal impact of family, friends and travel agents on choosing Hong Kong as their trip destination, and find that the most important factor is the word of mouth from primary reference groups (i.e. family and friends). However, it is found that, nowadays, the online networks are the most influential factor on travel decision-making and behavioral changes (Qu & Lee, 2011). In addition, Wang et al.’s (2012) study of mobile online networks uncover how smartphone applications are already being used to increase novel inter-tourist interactions in the physical world in real time.

3. RESEARCH HYPOTHESES
People interact with and belong to different types of groups for different purposes. Traveling is a social experience. Therefore, it is possible that which group people are traveling with can have a positive or negative impact on their traveling experience. According to research conducted by Harrel et al. (1980) perceived crowding might cause different behaviors and evaluation of the shopping trip. Needless to say, it is obvious that the purpose of going on a trip is to have hedonic and entertaining feelings with their companion. Thus, travelers might be affected by the companion’s feeling or evaluation on the trip or hotel. Married couples generally (88%) report conflicting with each other during the product purchase process (Sprio, 1983). Husbands are less willing to give away the option to wife on the purchase decision, specifically when both spouses have a different option and preference in their mind (Ward, 2006). Friends or peers are another major source of influence on customers’ behavior. The presence of friends can raise the desire to buy (Mangleburg et al., 2004). Family members increase a sense of responsibility and discourage wastefulness and extravagance. In short, choices and evaluations made by shoppers can be easily affected by companions’ feelings that they are shopping with. Considering that hotel reservation is one of the shopping behaviors, satisfaction or evaluation on the hotel where they stayed in can be different in this context. Hence, we hypothesize the following:

Hypothesis 1. There is a statistically significant difference in ratings between groups.

In line with hypothesis 1, we conducted an analysis in order to see whether there are group differences in terms of categories such as Cleanliness, Comfort, Location, Facilities, Staff, Value for Money and Wi-Fi. The description of the categories is defined in Table 1 below. In a shopping situation, for example, western children tend to greatly affect their parents’ product choices. Also, Chinese adolescents affect over two out of three of their parents’ daily shopping choices (Wang et al., 2007; McNeal and Yeh, 2003). In traveling environment, parents consider their children for their experience when they choose a destination and hotel for their trip. Moreover, parents’ satisfaction and evaluation on their travelling experiences and hotel can also be influences by opinions by their children. Especially, groups with family consider certain factors (e.g. bed, kitchen, toilet and staff) more important other factors (e.g. Wi-fi, cleanliness, noise etc.) Thus, hypothesis 2 is as follows below.

Hypothesis 2. There is a statistically significant difference in ratings between groups based on categories (interaction between group and category).

In line with this shopping experience, characteristics of various groups are different in traveling environment, and it will determine travelers travelling experiences. Therefore, the hypothesis is as follows.

Also, we hypothesize the following hypothesis 3 in order to find the reason why there is a statistically difference between groups (H1) and in ratings between groups based on categories (H2). It is assumed that the travelling experiences are different depending on which groups the traveler is traveling with.

Hypothesis 3. There is a difference in travelling experiences between groups.

We conducted regression analysis to examine hypothesis 1 and 2 and text analysis to examine hypothesis 3.

4. DATA COLLECTION
We collected the data from Booking.com which is the No.1 website related to traveling web service. A programming language called Python was used to develop a web crawling program which collects data automatically. By using this program, it was possible to collect not only the data presented in the website but also unpresented detailed data in the website.

In this study, we chose New York City to narrow down our research scope because New York is the biggest global city where people with diverse ethnic backgrounds. Moreover, travelers make a trip to NYC for various purposes such as honeymoon, family trip, business trip, and shopping tour, etc. Therefore, the purposes of traveling to NYC are relatively evenly distributed. Finally, NYC was chosen as our sample data since most of the reviews are written in English.

We collected 314 hotels comprised of the minimum number of reviews which are necessary for the analysis. The data collected in this study includes diverse travelers’ detailed reviews and scores (ratings) on the hotel they stayed in. Variables used in this research are as follows:

<table>
<thead>
<tr>
<th>Table 1. Definition and description of the variables</th>
</tr>
</thead>
<tbody>
<tr>
<td>Variable</td>
</tr>
<tr>
<td>Grade</td>
</tr>
<tr>
<td>Group</td>
</tr>
<tr>
<td>Category</td>
</tr>
<tr>
<td>Reviews</td>
</tr>
<tr>
<td>Score</td>
</tr>
</tbody>
</table>

It is possible to collect categorical detailed data by group if source code is analyzed by using a developed crawling tool although...
scores (ratings) by group are originally set up unseen on the real webpage. Thus, we developed a crawling tool in order to analyze the detailed data and result, we were able to collect categorical data by group. The collected data in a form is shown in the below Table 2. In this study, a total of 314 hotel data is collected.

### Table 2. Example of collected data – Linden Hotel (3 Stars)

<table>
<thead>
<tr>
<th>Location</th>
<th>Reviews</th>
<th>Services</th>
<th>Comfort</th>
<th>Value</th>
<th>Staff</th>
<th>Wi-Fi</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Family</td>
<td>8.5</td>
<td>8.6</td>
<td>8.6</td>
<td>8.4</td>
<td>8.5</td>
<td>8.5</td>
<td>8.5</td>
</tr>
<tr>
<td>Couple</td>
<td>7.7</td>
<td>7.5</td>
<td>7.5</td>
<td>7.5</td>
<td>8.5</td>
<td>7.5</td>
<td>7.6</td>
</tr>
<tr>
<td>Friends</td>
<td>8.7</td>
<td>8.5</td>
<td>8.3</td>
<td>8.2</td>
<td>8.1</td>
<td>8.4</td>
<td>8.4</td>
</tr>
<tr>
<td>Solo</td>
<td>7.8</td>
<td>7.6</td>
<td>7.8</td>
<td>7.3</td>
<td>7.2</td>
<td>7.2</td>
<td>7.6</td>
</tr>
<tr>
<td>Business</td>
<td>9.1</td>
<td>8.7</td>
<td>8.7</td>
<td>8.4</td>
<td>8.2</td>
<td>8.2</td>
<td>8.2</td>
</tr>
<tr>
<td>Total</td>
<td>8.2</td>
<td>8.1</td>
<td>8.2</td>
<td>7.9</td>
<td>7.8</td>
<td>7.8</td>
<td>8.1</td>
</tr>
<tr>
<td>Reviews</td>
<td>105</td>
<td>550</td>
<td>223</td>
<td>119</td>
<td>80</td>
<td>997</td>
<td></td>
</tr>
</tbody>
</table>

In this paper, review data was collected to conduct text analysis. The reviews are comprised of hotel rates, title, date, username, the number of the writers’ review, detail of the trip (purpose of travel, group, room type, nights of stay, written on mobile or PC etc.). The reviews in booking.com are divided into pros & cons. A total of the reviews are 125,076 containing 532,045 sentences. The example of collected reviews are presented in Figure 1.

**Table 2. Example of collected reviews – Linden Hotel (3 Stars)**

**Figure 1. Example of collected reviews**

### 5. RESEARCH METHODOLOGY

#### 5.1 Regression Analysis

We performed a regression analysis (OLS) with additive variables such as travelers’ group characteristics, hotel review category, and hotel grades. In this case, dependent variables are average scores (ratings) of hotels, and we controlled the difference made by the number of reviews by adding the number of reviews as an independent variable. In addition, an interaction term was added in order to see if there is a category that affects average scores depending on the characteristics of each group. The econometrics model is as follows:

\[
Y_i (Score_i) = \beta_0 + \beta_1 (Number \ of \ Reviews_i) + \beta_2 (Type \ of \ Group_i \ * \ Review \ Category_i) + \delta_1 (Hotel \ Grade_i) + \delta_2 (Type \ of \ Group_i) + \delta_3 (Review \ Category_i) + \epsilon
\]

#### 5.2 Text Analysis

For text analysis, we applied two measures: word frequency and the relative centrality of words within a semantic network (Chaves et al., 2011; Oh et al., 2010; Stringham & Gerdes, 2010).

Eigenvector centrality is a measure of the influence of a node in a network. A node with a high eigenvector centrality is linked to other nodes with high eigenvector centrality. This index is especially useful when determining which node is the most important node in the network. In other words, the more eigenvector centrality score get, the more the node means important. For example, Google’s PageRank also uses the eigenvector centrality measure to calculate which web page the most important is (“Centrality,” n.d.).

The eigenvector centrality \( x_i \) of node \( i \) is given by:

\[
x_i = \frac{1}{k_i} \sum_k a_{k,i} x_k
\]

where \( A = (a_{i,j}) \) be the adjacency matrix of a graph.

This measure was used to compensate the defect which can occur when the word frequency is used only (Oh et al., 2010). For instance, if the words ‘new’ and ‘well’ have an equal frequency, the eigenvector centrality between the two words can be different depending on with which word that they co-occur. In other words, the eigenvector centrality of ‘new’ would be higher than ‘well’ if ‘new’ co-occurs with other important words (e.g. facility, room, hotel) while ‘well’ co-occurs with words of less importance (e.g. known, done).

### 6. RESULTS

#### 6.1 Regression Analysis

The results of the OLS model show that hotel ratings are affected by which group travelers belong to. Also, it is identified that the scores are high in the order of group of couple, friends, family, solo, and business. However, in these results, the interaction of group characteristics and review categories are not statistically significant. This explains that there is no category which gives especially a higher or lower score depending on different groups.

**Table 3. Regression Analysis Results**

<table>
<thead>
<tr>
<th>Category</th>
<th>OLS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Grade 3</td>
<td>0.83020 *** (0.03377)</td>
</tr>
<tr>
<td>Grade 4</td>
<td>0.96810 *** (0.03299)</td>
</tr>
<tr>
<td>Grade 5</td>
<td>1.26900 *** (0.03837)</td>
</tr>
<tr>
<td>Group Couple</td>
<td>0.33620 *** (0.06547)</td>
</tr>
<tr>
<td>Group Family</td>
<td>0.15090 * (0.06516)</td>
</tr>
<tr>
<td>Group Friend</td>
<td>0.21250 ** (0.06511)</td>
</tr>
<tr>
<td>Group Solo</td>
<td>0.13260 * (0.06513)</td>
</tr>
<tr>
<td>Category Comfort</td>
<td>-0.34200 *** (0.06511)</td>
</tr>
<tr>
<td>Category Location</td>
<td>0.54140 *** (0.06511)</td>
</tr>
<tr>
<td>Category Service</td>
<td>-0.71780 *** (0.06511)</td>
</tr>
<tr>
<td>Category Staff</td>
<td>-0.07994 (0.06511)</td>
</tr>
<tr>
<td>Category Value</td>
<td>-0.98410 *** (0.06511)</td>
</tr>
<tr>
<td>Category Wi-Fi</td>
<td>-0.66320 *** (0.06516)</td>
</tr>
</tbody>
</table>
There are several reasons why the average scores given by various groups are different. First of all, past studies defined shopping as a social experience and showed that the values of shopping can be affected by who the shoppers are with (Borges, et. al, 2010). Applying it to hotel experiences, satisfaction that travelers feel can also differ depending on who they travel with even though they stay in the same hotel.

Moreover, roles that the travelers play for their trip can change depending on who they are travelling with. For example, travelers tend to make a plan for a trip precisely by themselves by searching for information on hotels, transportation, and schedules when they travel with parents. It is because they tend to take their parents and show them around the new places. However, if the same traveler makes a trip with friends or alone, the roles he or she plays can change dramatically. He or she might shift such chores onto friends or make a schedule vaguely when they travel with friends or alone. Therefore, it should be considered that satisfaction and values that travelers feel can be affected by who they travel with.

Next, preparation processes and situations might differ when travelers prepare for their trip. For instance, if a couple travels together, they might spend more time and money searching for information on hotels and reservations compared to a travelers who makes a trip alone. In case of a business trip, the budget is normally set or limited and the travelers tend to decide where to stay considering certain characteristics of hotels. It is due to the fact that business travelers normally do not have much time to look for hotels, so they decide a hotel taking account of locations which are close to their meeting or work. To sum up, a chain of such situations might have affected the average scores of hotels.

Finally, it should be considered that the demographics of travelers can be different. The average age of couple travelers can be much younger than those of family travelers or business travelers. Because of such various reasons, we suppose that the average scores of hotels can differ depending on different groups.

In case of hotel grades (stars), the more the hotel grades increase, the higher the average scores of hotels rise. In particular, when hotel grades increase from two-star to three-star hotels, the gap of average scores increases dramatically. On the other hands, the lowest gap of average score was hotel changes from three-star to four-star hotels. This means that three-star hotels are not only much better than two star hotels, but also have power to compete with four-star hotels. In reality, in other travel websites including Booking.com, it was the three-star hotels which take the highest percentage of the number of hotels and have the largest number of hotel guests. Moreover, three-star hotels are in keen competition compared to other hotel grades and these situations might have had an effect on guests’ satisfaction.

Average scores depending on different categories decreases in the order of Location > Cleanliness >= Staff > Comfort > Wi-Fi > Service > Value. Figure 2 shows the distribution of average scores depending on different categories more easily.

<table>
<thead>
<tr>
<th>Reviews</th>
<th>-0.00043 *** (0.00003)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Interaction Terms of</td>
<td>Not Significant</td>
</tr>
<tr>
<td>Group x Category</td>
<td>7.39300 *** (0.05520)</td>
</tr>
</tbody>
</table>

Notes: The Dependent Variable is the Score of Review. Standard errors are presented in parentheses.

In this figure, the categories which get the highest and lowest average score are location and value for money respectively. We suppose that it is because the object of this study was limited to New York City. In case of New York City, since most of the hotels are located downtown, satisfaction with locations tend to be high. On the other hand, value for money that travelers feel seems low because prices are too high in New York City. Also, given that variance of the two categories are low, we can find that the average scores are given by travelers relatively consistently. The one thing to take a close look at is the score of Wi-Fi. In case of Wi-Fi, the variance of the average score is extremely high and it seems to have a number of extremely negative scores compared to other categories. We suppose that the hotels tend to consider Wi-Fi relatively less important while they care about other categories such as clean, comfort, locations, etc. Considering the fact that mobile devices become more common and travelers is likely to be connected online anytime, more attention on the Wi-Fi category is necessary in hotel management.

One interesting result is that the more hotel reviews increase the lower average scores get. There are several reasons such interesting situation takes places. First, some of the hotels might give high scores by themselves on several travel websites in order to increase early scores. In case of Booking.com, given that reviews written after 14 months are deleted regularly and there is no abnormally high score of early reviews observed, it is almost impossible that the scores are manipulated by the hotels. Secondly, the hotels which have a number of reviews mean popular places, travelers tend to have a high expectation of the hotels before having an experience of staying. According to the study of Zeithaml, Berry, & Parasuraman (1993), expectations play a significant role in determining satisfaction. Teas (1993) argue that expectations may act as the standard that consumers want when assessing the product or service. According to the expectancy theory, a travel experience that satisfies or surpasses travelers’ expectations will be remembered positively (Cohen et al., 2014). Discrepancy theory explains that individuals’ satisfaction is determined by the differences between the perceived outcomes an individual gets and the expected outcomes (Anderson et al., 2012). In this context, a number of existing reviews on the website may decrease visitors’ satisfaction because their expectations might have set too high. In reality, such situations take place quite commonly in the mobile application market and review websites. It is also observed that the score of apps or products on websites becomes lower and stable as the number of reviews increases.

Such phenomenon is results opposite to some studies which argue that early reviews on hotel tend to be negative (Sparks & Browning, 2011; Bulchand-Gidumal et al., 2011). We suppose that it is due to
the fact that Booking.com regularly deletes reviews written after 14 months. In this aspect, the policy of Booking.com can decrease errors that early negative reviews might cause and have an advantage of providing the most recent information and status of hotels to travelers.

### 6.2 Text Analysis

The most frequently used top 10 words in pros reviews are ‘Location’, ‘Room’, ‘Staff’, ‘Hotel’, ‘Great’, ‘Good’, ‘Helpful’, ‘Friendly’, ‘Clean’, ‘Nice’. The findings are similar to results from the previous research results (Chaves et al., 2011). In the case of cons reviews, the most frequently used words are ‘Room’, ‘Hotel’, ‘Breakfast’, ‘Small’, ‘Bed’, ‘Staff’, ‘One’, ‘Time’, ‘Night’, ‘Wi-Fi’. Such results have no distinctive difference between groups.

The most frequently used words shown in pros reviews have a greater proportion than those shown in cons reviews have. In other words, it shows that the factors which reviewers feel satisfied with are similar. However, there can be various reasons why they feel dissatisfied with the hotels that they stayed in.

Table 4 below shows differences of important words which appeared in pros and cons reviews between groups. For example, in the case of family group, the reviewers consider location and breakfast the most important, and they take their daughters’ opinion seriously. Moreover, the results show that they enjoy buffet and like to go shopping in stores and watch a show. In the case of solo group, they are sensitive to noises around the hotel and consider the condition of room and service important. Moreover, solo groups tend to be more cost-sensitive, and spend more time using the internet even during the trip, and take hotel security as an important factor. Also, it is shown that they tend not to endure slow work processing.

<table>
<thead>
<tr>
<th>Group</th>
<th>What do they think more important?</th>
<th>What do they complain more about?</th>
</tr>
</thead>
<tbody>
<tr>
<td>Family</td>
<td>Location (6.19%, 1.000), Breakfast (1.07%, 0.308), Daughter (0.14%, 0.053), Buffet (0.06%, 0.020), Store (0.08%, 0.031), Show (0.08%, 0.027)</td>
<td>Bed (1.39%, 0.335), Breakfast (1.60%, 0.271), Crowd (0.13%, 0.033), food (0.25%, 0.064)</td>
</tr>
<tr>
<td>Couple</td>
<td>Location (5.88%, 1.000), New (0.45%, 0.162), View (0.77%, 0.221), Bar (0.53%, 0.163), Restaurant (0.53%, 0.180), Coffee (0.29%, 0.102), Wine (0.09%, 0.031), Decoration (0.01%, 0.005), Downtown (0.06%, 0.023)</td>
<td>Small (1.61%, 0.392), Noise (0.48%, 0.118), Toilet (0.16%, 0.051), Weather (0.02%, 0.007), Wi-Fi (0.77%, 0.147)</td>
</tr>
<tr>
<td>Friends</td>
<td>Location (6.03%, 1.000), Bed (1.43%, 0.400), Shop (0.25%, 0.086), Show (0.08%, 0.026), Girl (0.03%, 0.014), Starbucks (0.04%, 0.015)</td>
<td>Bed (1.48%, 0.341), Staff (1.08%, 0.288), Taxi (0.10%, 0.022), People (0.38%, 0.135)</td>
</tr>
<tr>
<td>Solo</td>
<td>Room (5.09%, 1.000), Free (0.54%, 0.155), Bathroom (0.48%, 0.014), Wi-Fi (0.47%, 0.154), Food (0.21%, 0.077), Atmosphere (0.13%, 0.037)</td>
<td>Noise (0.52%, 0.135), Dirty (0.32%, 0.091), Service (0.68%, 0.206), Wi-Fi (0.88%, 0.174), Price (0.53%, 0.143), Secure (0.07%, 0.029), Slow (0.32%, 0.069)</td>
</tr>
<tr>
<td>Business</td>
<td>Room (5.15%, 1.000), Comfort (1.45%, 0.444), Restaurant (0.57%, 0.184), Gym (0.11%, 0.032), Value (0.36%, 0.109), Pillow (0.08%, 0.027)</td>
<td>Noise (0.51%, 0.119), Dirty (0.33%, 0.092), Service (0.84%, 0.234), Secure (0.07%, 0.028), Elevator (0.82%, 0.165), Slow (0.31%, 0.067)</td>
</tr>
</tbody>
</table>

Notes: Frequencies and Eigenvector centralities are presented in parentheses.

There are some results which are not listed in the above table. For example, travelers traveling with groups (family, couple and friends) use the word ‘recommend’ more often than those who travel alone (solo and business). In case of food, the family and couple group mention a variety of food but other groups talk about less diverse food. Additionally, traveler who travel alone complains more on the elevator than those with groups. On the other hand, travelers who travel with groups gives more criticism on the bed in the hotel than those travelling alone.

Table 5 shows that solo and business travelers give low scores compared to travelers with groups. However, they write relatively many reviews compared to those traveling with group. In addition, the solo and business travelers stay in the hotel for a short time and they tend to review less on their phone.

<table>
<thead>
<tr>
<th>Group</th>
<th>Avg. Rating</th>
<th># of Reviews</th>
<th># of Stay (Days)</th>
<th>Reviews written on Mobile (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>family</td>
<td>8.03</td>
<td>3.46</td>
<td>3.49</td>
<td>38.8</td>
</tr>
<tr>
<td>couple</td>
<td>8.19</td>
<td>3.84</td>
<td>3.40</td>
<td>39.0</td>
</tr>
<tr>
<td>friends</td>
<td>8.15</td>
<td>3.05</td>
<td>3.33</td>
<td>40.6</td>
</tr>
<tr>
<td>solo</td>
<td>7.97</td>
<td>4.34</td>
<td>2.95</td>
<td>36.8</td>
</tr>
<tr>
<td>business</td>
<td>7.61</td>
<td>4.85</td>
<td>2.84</td>
<td>34.6</td>
</tr>
</tbody>
</table>

As shown in Table 4 and 5, the results indicate that there are differences on the factors that the reviewers consider important depending on different groups. According to our results, it is speculated that the experiences that travelers have can be different.
as traveling groups are different. We conjecture that this is the main reason that makes a difference on travelers’ satisfaction depending on who they are travelling with.

7. DISCUSSION

Travel related websites have been showing fast and continuous growth and becoming as influential as monthly unique visitors of booking.com exceeded 40 million (“Top 15 most popular”, 2015). Travel related websites became essential service which can give users an opportunity to collect information and deal with reservations and discounts. Moreover, those websites made it possible for travelers to organize itinerary and complex schedules, and even to share the feelings and experiences they had during their vacation. In other words, such travel sites play a significant role in providing essential service ranging from giving useful information in the process of preparation to expressing their feelings and reviews after their trip. However, there has not been numerous research using user-generated travel data even though such accumulated huge data make the service to have a better environment to make a leap forward. In that sense, this research is meaningful since we conducted empirical analysis using travelers’ behavioral data.

The purpose of this research is to show who you are travelling with can have a significant effect on travelers’ satisfaction with their hotel. It was possible to identify that satisfaction of travelers decrease in the order of couple, friends, family, solo, business by conducting regression analysis based on 314 average data of two to five star hotels located in New York City. The group of couple showed the highest satisfaction while business travelers expressed the lowest satisfaction. After our text analysis using reviews, it showed that there is a difference in traveling experiences depending on groups. We identified that such different traveling experience could be the reason why travelers feel satisfied differently depending on groups. We expect that such results can be used to plan and develop a marketing strategy or promotion in business.

One of the limitations of this paper is that a self-selection bias problem among groups still remains. In other words, before comparing the targets, it is possible to have a difference already among them. Also, the results found in this paper are not about causation but correlation effect. It is a difficult problem to confirm causation because of its unique characteristics of travelling which makes it hard to solve an endogeneity issue. In case of traveling, there are too many factors that affect the satisfaction of traveling and those are interconnected all together. Therefore, it is hard to control all the factors when evaluating their satisfaction. For example, there are not a lot of travelers who makes a trip to the same place several times. Moreover, there is little possibility that the travelers stays in the same hotel with various groups. Additionally, although they go to the same place with the same group, their satisfaction cannot be the same.

To solve this problem, we supported our research results by examining the literature reviews and theories in an alternative way, but the endogeneity problem still remain. We expect to find better ways to solve such problem in the future empirical research.

8. REFERENCES


