

An analytical examination of accommodation sales and the importance of electronic ‘word-of-mouth’ appraisals via internet travel sites

Yoshiki Nakamura (Department of Business Administration, Aoyama Gakuin University, nakamura@busi.aoyama.ac.jp)

Nozomi Oomiya (Faculty of Advanced Engineering, Nippon Institute of Technology, n.oomiya@nit.ac.jp)

Abstract

Electronic word-of-mouth (eWOM) is now a standard in daily life for Internet communication. This study focuses on the eWOM on travel sites, specifically, on the eWOM of reviews and assessments of three hotels through the Internet. We set two research questions. One is “Are sales influenced by reviews and assessments?” The other is: “Is there any time difference between ‘sale’ and ‘reviews and assessments?’” In order to answer the research questions, text-mining method was used to extract nouns from the reviews and to quantify them. Using the output, correlation analysis, principal component analysis, and multiple regression analysis was calculated and indicated the trends and relationships between reviews and assessments. This study discovers two things: one is recognizing the linkage between reviews and assessments. The other is that hotels can more easily find problem trends in their hotels and predict people’s stays. This information provides hotels with preparation time. Possibilities for future research include: (1) making the objects of text-mining not only nouns, but also adjectives, adverbs, and other parts of speech; (2) expanding the subjects for other hotels and other sites of travel; and (3) using other industrial engineering methods for analysis.

Keywords

text-mining, travel site analysis, word-of-mouth analysis, principal component analysis, multiple regression analysis

1. Introduction

In the 1980s, companies were able to sell many products without as much effort as has been required in recent years. The reason is that there was a shortage in home appliances and people needed product itself. Therefore, people only watched how to use products during those days [Kotha, 1996]. “Word-of-mouth” (WOM) is one of the useful techniques to learn about new products. WOM is defined as the communication between consumers regarding a product, service, or a company wherein the sources are considered to be independent of any commercial influence [Litvin et al., 2008; Cantalops and Salvi, 2014]. In the marketing perspective, a “mass marketing” was mainstream during that era and was accomplished mainly through newspaper advertisements, outdoor signs and billboards, and radio and television commercials [Davis, 1987; Pine, 1993; Kotha, 1995; Talalay, 2013]. After the late of 90’s, however, Internet technology was born and marketing concepts drastically changed. Specifically, WOM was transformed into electronic WOM (eWOM) [Murugesan, 2007] [Holsapple et al., 2018].” eWOM is communication regarding products through Internet sites [Abubakar and Ilkan, 2016; Hennig-Thurau et al., 2004]. Specifically, the consumer cannot only receive information from web sites about products, but they can also read other users’ comments and evaluations. Customers provide information for other users who might be interested in products and services as well. On the other hand, the business side takes advantage of eWOM due to its vast volumes of useful knowledge about products and services. As a result, eWOM is for not only for customers but also for employees,

competitors, and enterprise partners [Holsapple et al., 2018]. eWOM services include blogs; search engines such as Google and Yahoo; wikis that can be edited by users like Wikipedia; social network services like Twitter, Instagram, Facebook, and LinkedIn; and sales and review sites like Amazon, Yelp, and TripAdvisor.

In terms of related research, WOM research historically began in the 1950s [Whyte, 1954; Engel et al., 1969; Arndt, 1967]. Specifically, researchers attempted to measure the effect of WOM from neighborhood communication. For example, Katz and Lazarsfeld [1955] proved that WOM is seven times more effective through newspapers and magazines when furniture or food was being purchased. WOM is also twice as effective as commercials on the radio. In the 1960s, the WOM research field was shifted to television (TV) ratings and new products [Brown and Reingen, 1987; Reingen and Jerome, 1986]. There were also studies that conducted mathematical analysis to measure how helpful WOM was [Duhan et al., 2001].

This research focuses on travel web sites, specifically, travel sites’ research surveys [Filiari and McLeay, 2014; Edwards et al., 2017; Wu et al., 2016]. Filiari and McLeay [2014] focused on the travel sites’ reviews and analyzed the relation between reviews and product ranking, information accuracy, information value-added, information relevance and information timeliness. Wu et al. [2016] proposed investigating the potential operation of niche market identifications in the tourism industry and developed the clustering approach to search niche markets. Berezina et al. [2015] examined the underpinnings of satisfied and unsatisfied hotel customers and correlated their reviews and hotel values. He et al. [2017] found that star rating reviews are correlated with online customer textual reviews. A case study also revealed that both extremely satisfied and

extremely dissatisfied hotel customers shared a common interest in five categories: food, location, rooms, service, and staff. Vu et al. [2019] implicitly revealed tourist activity preferences through complex travel patterns by analyzing Twitter tweets. There have been also numerous studies on targeting eWOM [Feldman et al., 1998; Seol et al., 2011]. Sen and Lerman [2007], for example, found the existence of this negativity effect in eWOM and examined the influence of the reader’s attributions.

As mentioned above, any previous researches regarding how much WOM influences sales have not been advanced. Specifically, discussions are mainly from the customer perspective. Therefore, it is necessary to understand the relationship between WOM and sales and suggest the problems that exist and the improvements that should be made by analyzing the comments of reviews. There is advice for hotels regarding the things they should consider in terms of their problems.

In this research, three Japanese hotels were selected from one of the travel sites. The targets of eWOM in this study nominate the site’s “review” and “assessment items.” The review consists of the customer’s comments and impressions. This study tries to quantify the review and its numerical output is used to determine the relationship between reviews, assessment items, and sales. After analyzing the relationship, we conduct an analysis from the hotel perspective and make problem-solving suggestions from the reviews. Through this analysis, we can propose the best way for hotel staff to use eWOM.

2. The research questions and objects

2.1 Research questions

Figure 1 illustrates the research approach. Our research assumes a “sale” being represented by the sum of price plans for each customer and each month. There are two research questions:

- Q1: Are sales influenced by reviews and assessments?
- Q2: Is there any time difference between “sales” and “reviews and assessments?”

Q1 tries to understand the influence of reviews and assessments. In the past, WOM was one of the ways customers made purchasing or hotel decisions. It is, however, difficult to

measure the effect of WOM. Recently, it has become easier to capture the information and to analyze the linkage between eWOMs. Q2 tries to investigate the time trend between sales and the reviews and assessments. This question assumes the following premise: “Before booking hotels, customers check and refer to past reviews and assessments of the hotel at which he or she wants to stay. There are some time lag characteristics.” If these questions are clearly addressed, it will serve as a guideline for hotels regarding the specific evaluation items and reviews that should be managed and how to increase sales. Therefore, this research indicates the number of months it takes for hotels’ “sales” to appear on the Internet after their “reviews and assessments.”

The contribution of this research consists of two aspects. One is recognizing the linkage between reviews and assessments. The other is making it easier for hotels to find problem trends with their hotels and to predict people’s stays. This information provides preparation time for hotels.

2.2 The research object and data

This research focuses on one of the Japanese travel sites, which we call “website A.” Website A serves as not only a place for booking a hotel, but also for creating a tour plan, booking flights, and renting cars. The site also posts information about sightseeing and events near the hotel. After customers stay at the booked hotel, he or she can input various data: one type is basic demographic information such as an anonymous name, dates they stayed, gender, age, plan, and price. The other type of data is a review of their impressions or complaints, as well as assessment items such as “overall,” “room,” “bath,” “breakfast,” “dinner,” “service,” and “cleanliness.” Their items are assessed using a five-point Likert scale. Some hotels reply to guests’ review comments.

This research focuses on three hotels: the first is called “X,” a ryokan in the Miyagi prefecture. The second is called “Y,” a hotel in Tokyo. The third is called “Z,” a ryokan in the Fukui prefecture. The reasons why these accommodations select are all luxury type and very famous in Japan. The data were accumulated between October 2014 and March 2016. The full sample size consists of 669 hotels and there are each review and assessment data. These reviews pick only Japanese.

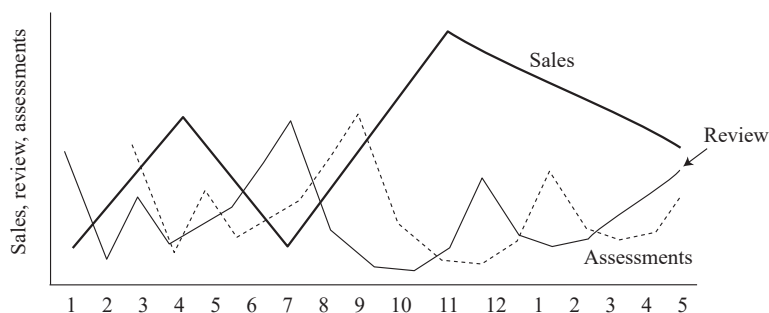


Figure 1: Research approach

3. Results of the data evaluation

3.1 Basic calculations

Table 1 demonstrates the average price and assessment items for each hotel. "Overall" is the average of three hotels. The price average is 15,702 yen per night. Ryokan Z has the highest price at 18,767 yen per night. All of the hotels are proved to be the luxury type.

The average of the overall assessment items is over 4.0 points. Hotel Y's assessments note that there are no outputs for "breakfast" and "dinner" because hotel Y does not serve meals. Every hotel gets higher evaluations. However, "bath" in hotel Y receives a lower evaluation because the hotel is located in the center of Tokyo and its rooms are usually small. Nevertheless, "service" and "cleanliness" for hotel Y are higher than ryokan X and ryokan Z. On the other hand, both ryokan X and ryokan Z are located in regional hot springs. Their strengths, therefore, are "bath" and meals.

3.2 Review of calculations

There is the review part in the assessment field. The customer enters a review of their impressions and opinions. This is very useful information for not only other customers but also for hotel management. Therefore, text-mining analysis is conducted on the reviews and nouns are extracted from the reviews

using KHCoder software [2020]. The number of commonly extracted words among the three hotels is 140 words. The rate of appearance is calculated after the extraction is completed. The formula for calculating the rate of a word's appearance is as follows:

$$\text{Rate of word appearance} = \frac{\text{Count of the word}}{\text{Total number of the words in reviews}} \quad (1)$$

Table 2 presents data regarding the rate of a word's appearance and its ranking for each hotel. The first-place word is "buffet" for ryokan X, "hotel" for hotel Y, and "room" for ryokan Z. Each word expresses its own strength points and relevance for sales. Ryokan X and Z's customers usually indicate family hotel use. Therefore, there are many related words that appear such as "kids" and "dinner."

Next, correlation analysis is conducted between a word's appearance and the hotel's assessment items, where "*" indicates a 5 % significance difference and "***" indicates a 1 % significance difference in Tables 3 through 5. The words select as a part and they gain the significance difference. In Table 3, ryokan X's words include "reception," "impression," and "opportunity." They are words of appreciation. This is the same as the other hotels, where there is a positive correlation demon-

Table 1: Averages of price, overall assessment, and assessment items

	Price	Overall	Room	Bath	Breakfast	Dinner	Service	Cleanliness
Overall	15702.928	4.333	4.180	4.081	4.433	4.476	4.328	4.242
X	14870.663	4.440	4.216	4.281	4.439	4.670	4.258	4.176
Y	13471.000	4.260	4.180	3.776	N/A	N/A	4.500	4.390
Z	18767.121	4.298	4.145	4.185	4.427	4.283	4.226	4.161

Source: Created by the authors based on the Website A.

Table 2: The word appearance ranking

	X (Word appearance: Ranking)	Y (Word appearance: Ranking)	Z (Word appearance: Ranking)
Buffet	0.430 : 1	0.020 : 18	0.109 : 9
Bath	0.284 : 2	0.061 : 10	0.198 : 5
Dinner	0.284 : 3	0.020 : 19	0.203 : 4
Rooms	0.145 : 7	0.296 : 2	0.270 : 1
Hotel	0.121 : 9	0.593 : 1	0.089 : 11
Pool	0.103 : 11	0.151 : 4	0.008 : 35
Family	0.106 : 10	0.019 : 20	0.093 : 10
Reception	0.038 : 19	0.098 : 7	0.028 : 26
Opportunity	0.049 : 15	0.025 : 16	0.033 : 22
Couple	0.039 : 18	0.020 : 17	0.047 : 17
Smile	0.028 : 28	0.004 : 27	0.044 : 18
Menu	0.032 : 25	0.004 : 28	0.016 : 29
Impression	0.020 : 33	0.008 : 25	0.031 : 23
Bath	0.023 : 30	0.030 : 15	0.003 : 36

Note: Left row indicates the word appearance ranking; right row shows the ranking of word's appearance.

Table 3: Correlation analysis between a word’s appearance and ryokan X’s assessment items

Word	Overall	Room	Bath	Breakfast	Dinner	Service	Cleanliness
Reception	-0.593*	-0.453	-0.865**	-0.196	-0.431	-0.194	-0.134
Opportunity	0.408	0.254	0.713**	-0.103	0.374	-0.080	-0.049
Impression	-0.719**	-0.351	-0.667*	-0.273	-0.744**	-0.439	-0.328
Bath	-0.807**	-0.500	-0.675*	-0.486	-0.771**	-0.606*	-0.514

Notes: * is significant at the 5 % level. ** is significant at the 1 % level.

Table 4: Correlation analysis between a word’s appearance and hotel Y’s assessment items

Word	Overall	Room	Bath	Service	Cleanliness
Rooms	0.236*	0.178	0.159	0.183	0.191
Hotel	-0.071	-0.007	-0.234*	-0.115	-0.183
Pool	-0.073	-0.023	-0.204*	-0.090	-0.139
Smile	-0.142	-0.240*	-0.089	-0.302**	-0.172
Room	-0.183	-0.147	0.043	-0.401**	-0.147
Bath	-0.255*	-0.058	-0.327**	-0.029	-0.205*

Notes: * is significant at the 5 % level. ** is significant at the 1 % level.

Table 5: Correlation analysis between a word’s appearance and ryokan Z’s assessment items

Word	Overall	Room	Bath	Breakfast	Dinner	Service	Cleanliness
Dinner	-0.082	-0.043	-0.024	-0.220*	-0.311**	-0.153	0.014
Family	-0.082	-0.172	-0.119	-0.033	0.035	0.019	-0.245**
Smile	0.190*	0.126	0.110	0.160	0.136	0.196*	0.186*
Menu	-0.191*	-0.107	-0.124	-0.357**	-0.298**	-0.091	-0.092

Notes: * is significant at the 5 % level. ** is significant at the 1 % level.

strating a good impression, such as “appreciate,” “thanks,” and “smile.” A negative correlation displays a negative impression and comments. For example, words such as “poor reception” and “towels have poor absorption” appear. This output seems to be the same as for other hotels. Through these analyses, the hotel staff understands what kind of the review words are linked to the assessment.

3.3 Principal component analysis and analysis of the output

Principal component analysis was used to attempt to search for extracted words. The output attempts to capture the overall trends in reviews. The process is as follows. (1) The correlation value is calculated between the word’s appearance before conducting principal component analysis. (2) The words are put together according to high correlation and similar usage. (3) An eigenvalue is calculated and a number of synthetics is decided. (4) Each synthetic item’s name is checked before and after the review. (5) The principal component scores are calculated and analyzed using statistical methods.

Next process is 18 words remains after selecting by the correlation value (see Table 7’s first column). Table 6 presents the eigenvalues and six components fulfill more than 40 %. Therefore, six synthetic variables were set for this research. Table 7 presents each component according to the principal

component load. Principal component 1 is defined as “facility” because there are many comments about rooms and gardens in the reviews. In the same manner, principal component 2 is named “service,” principal component 3 is named “atmosphere,” principal component 4 is named “satisfaction,” principal component 5 is named “family,” and principal component 6 is “thanks.”

Next, multiple regression analysis was conducted for answering research question 2. For this regression, the rate of price increase was set using the following formula:

$$\Delta_t = (\Delta_t \text{'s price} - \Delta_{t-1} \text{'s price}) \div \Delta_{t-1} \text{'s price} \quad (2)$$

t : month ($t = 1, 2, 3$)

The reason why this research considers the rate of increase is review must be influence to the increase and decrease of prices. Their influences are changed depended on the month elapsed. As a result, this research sets three items. Δ_1 is the rate of increase from the previous month. In the same manner, Δ_2 is the rate of increase from two months earlier, and Δ_3 is the rate of increase from three months earlier.

In Tables 8, 10, and 12, the results of multivariate analysis of objective variables, such as price, and explanatory variables, such as the assessment items in each hotel, are presented.

Table 6: Eigenvalues

Principal component	Eigenvalue	Contribution rate	Cumulative rate
1	1.560	8.665 %	8.665 %
2	1.386	7.701 %	16.366 %
3	1.269	7.052 %	23.418 %
4	1.205	6.693 %	30.111 %
5	1.130	6.278 %	36.389 %
6	1.096	6.090 %	42.480 %

Table 7: Principal component load

Word	Facility	Service	Atmosphere	Satisfaction	Family	Thanks
Garden	0.563	-0.105	0.254	0.080	0.334	-0.128
Rooms	0.462	-0.143	0.169	0.181	-0.024	-0.324
Feeling	0.388	-0.085	-0.081	0.115	0.120	-0.273
Hotel	0.343	-0.182	0.161	-0.430	-0.134	0.236
Staff	0.313	0.036	-0.536	0.182	0.242	0.155
Bath	-0.226	0.486	-0.016	-0.134	-0.082	-0.401
Buffet	-0.448	0.323	0.007	0.035	0.225	-0.270
Inside	-0.054	0.236	0.001	-0.431	0.342	0.228
The Best	-0.042	0.203	0.151	0.309	0.525	-0.235
Couple	0.013	0.048	0.396	0.378	-0.070	0.352
Fun	-0.205	0.171	0.348	0.422	-0.119	0.284
Atmosphere	0.087	0.198	0.312	-0.307	0.358	0.223
Word Of Mouth	0.046	0.125	0.131	0.351	0.102	0.229
Smile	0.149	0.156	-0.602	0.226	0.108	0.208
Parents	-0.192	-0.040	-0.194	-0.022	0.314	0.341
Children	-0.413	-0.575	-0.034	0.115	0.191	-0.102
Swimming Pool	-0.352	-0.644	-0.006	0.023	0.180	-0.014
Opportunity	0.009	0.132	-0.200	0.173	-0.375	0.045

Table 8: Multiple regression analysis between price and assessment items for ryokan X

	Room	Bath	Breakfast	Dinner	Service	Cleanliness	CT	R2	DW
Δ_1	-0.328 (-1.270)	0.700 (2.889)	-0.859 (-2.354)	-0.118 (-0.458)	0.559 (2.485)	0.083 (0.394)	0.038 (0.043)	0.919	1.678
Δ_2	-0.330 (-1.030)	0.702 (2.097)	-0.861 (-1.873)	-0.118 (-0.391)	0.558 (2.011)	0.086 (0.259)	0.042 (0.039)	0.904	1.574
Δ_3	-0.307 (-0.922)	0.660 (1.884)	-0.495 (-0.786)	-0.088 (-0.277)	0.558 (1.938)	-0.077 (-0.197)	-0.955 (-0.607)	0.918	1.776

Table 9: Multiple regression analysis between price and principal component score for ryokan X

	Facility	Service	Atmosphere	Satisfaction	Family	Thanks	CT	R2	DW
Δ_1	-0.050 (-0.176)	-0.115 (-0.551)	0.651 (1.685)	0.306 (0.635)	0.117 (0.339)	-0.329 (-0.774)	-0.015 (-0.176)	0.764	2.664
Δ_2	-0.035 (-0.110)	-0.104 (-0.451)	0.555 (1.186)	0.395 (0.704)	0.065 (0.166)	-0.390 (-0.804)	-0.030 (-0.304)	0.743	2.564
Δ_3	0.149 (0.427)	-0.129 (-0.574)	0.579 (1.276)	0.518 (0.933)	0.200 (0.499)	-0.504 (-1.048)	0.040 (0.345)	0.814	1.882

Table 10: Multiple regression analysis between price and assessment items for hotel Y

	Room	Bath	Service	Cleanliness	CT	R2	DW
Δ_1	-0.881 (-1.894)	0.261 (1.074)	0.703 (1.604)	-0.007 (-0.020)	-0.422 (-0.262)	0.686	2.363
Δ_2	-0.914 (-2.268)	0.250 (1.187)	0.603 (1.574)	0.198 (0.643)	-0.728 (-0.519)	0.803	2.136
Δ_3	-0.724 (-2.057)	0.316 (1.765)	0.406 (1.204)	0.484 (1.602)	-2.100 (-1.503)	0.886	1.961

Table 11: Multiple regression analysis between price and principal component score for hotel Y

	Facility	Service	Atmosphere	Satisfaction	Family	Thanks	CT	R2	DW
Δ_1	-0.275 (-1.600)	0.229 (1.152)	0.291 (1.767)	0.363 (1.346)	0.149 (0.730)	0.094 (0.247)	0.535 (1.780)	0.824	2.957
Δ_2	-0.304 (-1.569)	0.269 (1.185)	0.474 (1.341)	0.627 (1.185)	0.045 (0.161)	0.330 (0.577)	0.633 (1.728)	0.832	2.647
Δ_3	-1.342 (-3.343)	0.869 (3.372)	0.922 (3.528)	0.459 (1.489)	-0.191 (-1.048)	0.533 (1.593)	1.605 (3.847)	0.963	1.800

Table 12: Multiple regression analysis between price and assessment items for ryokan Z

	Room	Bath	Breakfast	Dinner	Service	Cleanliness	CT	R2	DW
Δ_1	0.711 (1.895)	-0.288 (-0.761)	-0.258 (-1.282)	-0.382 (-1.324)	0.265 (0.824)	0.041 (0.130)	-0.281 (-0.254)	0.880	2.814
Δ_2	-0.664 (-1.256)	0.011 (0.021)	-0.019 (-0.070)	0.228 (0.598)	-0.095 (-0.217)	0.786 (1.821)	-1.089 (-0.737)	0.822	3.105
Δ_3	-0.497 (-0.994)	-0.737 (-1.533)	0.083 (0.353)	1.052 (2.862)	-0.647 (-1.755)	0.600 (1.392)	0.581 (0.370)	0.926	2.008

In Tables 9, 11, and 13, the results of multivariate analysis between objective variables, such as price, and explanatory variables, such as the principal component score for each hotel, are presented. The upper row presents the partial regression coefficients and the lower row with brackets represents the t-values. The right two columns present the coefficients of determination (R2) and Derby-Watson ratios. R2 is a statistical measure, and its item clears how close the data are to the fitted regression line. DW is a test for autocorrelation in the residuals.

In the partial regression coefficients presented in Table 8, the increasing rate demonstrates the effect of each evaluation item. Specifically, both ‘bath’ and ‘service’ are positive for all three periods. This suggests that ryokan X needs to pay attention to ‘bath’ and ‘service.’ In Table 9, ‘atmosphere,’ ‘satisfaction,’

and ‘family’ are always positive. Particularly, the R2 of Δ_3 is higher. This recommends that ryokan X needs to pay attention to these three things before three months.

Similarly, hotel Y’s R2 of Δ_3 is higher in Table 10. It is also important that hotel Y manages ‘bath’ and ‘service.’ Since hotel Y is a luxury hotel in Tokyo, the impact of customer service is higher and is essential to the price. In Table 11, ‘service,’ ‘atmosphere,’ ‘satisfaction,’ and ‘thanks’ are all positively affected during the three previous periods.

Ryokan Z in Table 12 always indicates positive values for ‘cleanliness.’ The t-value of Δ_3 ’s ‘dinner’ is higher. A unique feature is that both Δ_1 and Δ_2 are positive partial regression coefficients in Table 13; ‘Facility,’ ‘service,’ ‘atmosphere,’ ‘satisfaction,’ and ‘thanks.’ This suggests that if a review has negative comments, ryokan Z needs to manage it as soon as

Table 13: Multiple regression analysis between price and principal component score for ryokan Z

	Facility	Service	Atmosphere	Satisfaction	Family	Thanks	CT	R2	DW
Δ_1	0.353 (1.381)	0.989 (1.602)	0.910 (1.066)	1.713 (1.324)	-0.997 (-0.921)	0.304 (0.491)	-0.236 (-1.396)	0.811	1.840
Δ_2	0.334 (0.887)	0.674 (0.674)	1.106 (0.785)	1.304 (0.587)	-1.245 (-0.684)	0.730 (0.671)	-0.166 (-0.661)	0.619	2.539
Δ_3	-0.705 (-15.310)	-2.590 (-18.208)	-5.430 (-18.498)	-9.609 (-19.434)	7.061 (19.718)	-4.963 (-18.980)	0.800 (19.127)	0.998	1.743

possible.

We examined them to solve our research questions by these outputs. Question 1 is 'Are sales influenced by reviews and assessments?' We discuss the rate of prices and find the linkage among the sales, review and assessments. The answer is 'yes.' Question 2 is 'Is there any time difference between "sales" and "reviews and assessments?" The answer is 'yes,' and we find some notice from reviews.

4. Conclusion

eWOM analysis is important not only to the customers who are referred but also for the company to manage and solve problems. Therefore, this study focuses on a Japanese travel site and tries to analyze the relationships between sales and reviews and assessment items. Six hundred and sixty-nine assessments were extracted and each review attempt was numerically labeled according to principal component analysis. These reviews created six components. Multiple regression analysis was conducted using two approaches: one was between price and the assessment items, the other was between price and the principal component scores. Through this analysis, this research indicates that the influence of the time difference depends on the specific hotel. We can also discuss which parts are most important for the hotel to consider. This discussion provides important aspects for capturing information about trends and problems.

Future works should include: (1) making the object of text mining not only nouns, but also adjectives, adverbs, and other parts of speech, (2) expanding the subjects to other hotels and other sites of travel, and (3) extending the hotel's and tourist's demographic information to include factors such as gender, city, and birthplace.

References

- Abubakar, A. M. and Ilkan, M. (2016). Impact of online WOM on destination trust and intention to travel: A medical tourism perspective. *Journal of Destination Marketing and Management*, Vol. 5, No.3, 192-201.
- Arndt, J. (1967). Role of product-related conversations in the diffusion of a new product. *Journal of Marketing Research*, Vol. 4, No. 3, 291-295.
- Berezina, K., Bilgihan, A., Cobanoglu, C., and Okumus, F. (2015). Understanding satisfied and dissatisfied hotel customers: Text mining of online hotel reviews. *Journal of Hospitality Marketing and Management*, Vol. 25, No. 1, 1-24.
- Brown, J. J. and Reingen, P. H. (1987) Social ties and word-of-mouth referral behavior. *Journal of Consumer Research*, Vol. 14, No. 3, 350-362.
- Cantalalops, A. S. and Salvi, F. (2014). New consumer behavior: A review of research on eWOM and hotels. *International Journal of Hospitality Management*, Vol. 36, 41-51.
- Davis, S. (1987). *Future perfect*. Boston: Addison Wesley.
- Duhan, D. F., Johnson, S. D., Wilcox, J. B., and Harrell, G. D. (1997). Influences on consumer use of word-of-mouth recommendation sources. *Journal of the Academy of Marketing Science*, Vol. 25, No. 4, 283-295.
- Edwards, D., Cheng, M., Wong, I., Zhang, J., and Wu, Q. (2017). Ambassadors of knowledge sharing: Co-produced travel information through tourist-local social media exchange. *International Journal of Contemporary Hospitality Management*, Vol. 29, No. 2, 690-708.
- Engel, J. E., Blackwell, R. D., and Kegerreis, R. J. (1969). How information is used to adopt an innovation. *Journal of Advertising Research*, Vol. 9, No. 4, 3-8.
- Feldman, R., Fresko, M., Hirsh, H., Aumann, Y., Liphstat, O., and Reimer, U. (1998). Knowledge management: A text mining approach. *Paper presented at the 2nd International Conference on Practical Aspects of Knowledge Management*.
- Filieri, R. and McLeay, F. (2014). E-WOM and accommodation: An analysis of the factors that influence travelers' adoption of information from online reviews. *Journal of Travel Research*, Vol. 53, No. 1. 44-57.
- He, W., Tian, W., Tao, R., Zhang, W., Yan, G., and Akula, V. (2017) Application of social media analytics: A case of analyzing online hotel reviews. *Information Review*, Vol. 41, No. 7. 921-935.
- Hennig-Thurau, T. Gwinner, K. P., Walsh, G., and Gremler, D. D. (2004). Electronic word-of-mouth via consumer-opinion platforms: What motivates consumers to articulate themselves on the internet? *Journal of Interactive Marketing*, Vol. 18, No. 1, 38-52.
- Holsapple, C. W., Hsiao, S., and Pakath. R. (2018). Business social media analytics: Characterization and conceptual framework. *Decision Support Systems*, Vol. 110, 32-45.
- Katz, E. and Lazarsfeld, P. F. (1955). *Personal influence: The part played by people in the flow of mass communications*. Glencoe, III., Free Press.
- KHCOrder. <http://khcoder.net/>. Retrieved in March 2020.
- Kotha, S. (1995). Mass customization: Implementing the emerging paradigm for competitive advantage. *Strategic Management Journal*, Vol. 16, No. S1. 21-42.
- Kotha, S. (1996). From mass production to mass customization: The case of the national industrial bicycle company of Japan author links open overlay panel. *European Management Journal*, Vol. 14, No. 5, 442-450.
- Litvin, S., Goldsmith, R., and Pan, B. (2008). Electronic word-of-mouth in hospitality and tourism management. *Tourism Management*, Vol. 29, No. 3, 458-468.
- Murugesan, S. (2007) Understanding web 2.0. *IT Professional*, Vol. 9, No. 4, 34-41.
- Pine II, B. J. (1993). *Mass customization: The new frontier in business competition*. Boston Harvard Business School Press.
- Talalay, L. (2013). The past as commodity: Archaeological images in modern advertising. *Public Archaeology*, Vol. 3, No. 4, 205-216.

- Reingen, P. H. and Jerome, B. K. (1986). Analysis of referral networks in marketing: Methods and illustration. *Journal of Marketing Research*, Vol. 23, No. 4, 370-378.
- Sen, S. and Lerman, D.. (2007). Why are you telling me this?: An examination into negative, consumer reviews on the web. *Journal of Interactive Marketing*, Vol. 21, No. 4, 76-94.
- Seol, H., Lee, S., and Kim, C. (2011). Identifying new business areas using patent information: A DEA and text mining approach. *Expert Systems with Applications*, Vol. 38, No. 4, 2933-2941.
- Vu, H., Li, G., and Law, R. (2019). Discovering implicit activity preferences in travel itineraries by topic modelling. *Tourism Management*, Vol. 75, 435-446.
- Whyte, W. H. (1954). The web of word of mouth. *Fortune*, Vol. 50, 140-143.
- Wu, C., Ho, G. T., Lam, H. Y., Ip, W. H., Choy, K. L., and Tse, M. (2016). An online niche-market tour identification system for the travel and tourism industry. *Internet Research*, Vol. 26, 167-185.

(Received April 7, 2020; accepted April 27, 2020)