How Online Reviews in a Year Predict Online Sales in the Next on Expedia.com + Agoda.com + Hotels.com? A Panel Study of Hotels

Snehasish Banerjee
The York Management School
University of York
York, England, UK
e-mail: snehasish.banerjee@york.ac.uk

Stefanie Bonfield
Department of Psychology
University of York
York, England, UK
e-mail: sb1699@york.ac.uk

Abstract—**This paper investigates how ratings, titles as well as descriptions of online reviews predict online sales. Using data from Expedia.com, Agoda.com and Hotels.com; a log-linear regression model was developed for a panel of 75 Asian hotels. The model explained 69.40% variance in the dependent variable for luxury hotels, 40.30% for budget hotels, and 38.80% for mid-scale hotels. In particular, title length was negatively related to sales for luxury and mid-scale hotels. The use of positive words in titles was positively related to sales for luxury hotels but had a negative association for budget hotels. Moreover, the use of positive (negative) words in descriptions was positively (negatively) related to sales for budget hotels.**

Keywords-e-tourism, e-WOM, online booking, online sales

I. INTRODUCTION

A. Background and Research Goal

With the ubiquity of the Internet, people not only book hotels online but also read user-generated reviews before confirming where to stay. About 96% of Internet users believe that reading reviews is important, and 83% of them hesitate to book a hotel unless they read reviews [1]. Clearly, reviews make or break the fate of hotels. While a plethora of positive reviews extoling a hotel inspires confidence to proceed with a booking, an abundance of negative entries highlighting criticisms can have a damning effect.

On the research front, scholars have widely studied how the numerical ratings of reviews are able to predict sales [2,3]. However, an online review submitted on most contemporary websites consists of three parts: a rating, a title, and a description. Hence, scholars seem to have been examining the role played by the first part while ignoring the second and the third. Taking into account these two textual components of reviews is also conceptually important. After all, text has a greater influence on purchase decisions than ratings [4]. In fact, people can use ratings to decide which reviews to read, and further lean on titles as well as descriptions to determine whether to make bookings [5].

Hence, the goal of this paper is to investigate how all the three parts of reviews—ratings, titles and descriptions—can cumulatively predict online sales of hotels on three booking-cum-review websites: Expedia.com, Agoda.com and Hotels.com. To offer a granular treatment, the paper separately examines three types of hotel category, namely, luxury, budget and mid-scale. This was necessary because hotel category dictates, at least in part, how reviews affect people's expectation and perception [6]. In consequence, the relation between reviews and sales is unlikely to be consistent across luxury, budget and mid-scale hotels [7].

The significance of this paper is two-fold. First, it extends the literature on the link from reviews to sales by including the textual components of titles and descriptions. This contributes to a deeper understanding of the role that different components of reviews play in predicting sales. Second, the findings have potential business implications. They will shed light on how reviews predict sales differently for luxury, budget and mid-scale hotels. Accordingly, hoteliers and administrators of review websites could work together to guide users on how to write titles and descriptions as a possible step toward improved sales performance.

B. Related Literature

In one of the earliest related works, an increment in ratings was found to lead to an improvement in relative sales rank of books on Amazon.com and Barnesandnoble.com [2]. The trend was later identified in other contexts ranging from video games [8] to hotels [3]. However, the finding is far from unanimous. For example, ratings were non-significant in predicting movies' box office revenues [9]. To this end, some works argue that reviews are able to dictate sales not due to their ratings but by virtue of their sheer volume [10]. When a product or a service garners a substantial quantity of reviews, the online buzz surrounding it has a significant bearing on its sales [7]. This suggests that controlling for the volume of reviews is imperative to meaningfully tease out how reviews per se are able to predict sales.

Interestingly, the earliest work [2] pointed that the length of reviews was a significant predictor of sales. This suggests that users pay attention, at least to some extent, to the text in reviews rather than solely relying on numerical figures conveyed through ratings. Yet, no attempt has hitherto been made to granularly shed light on the role played by the textual components of titles and descriptions in reviews.

The rationale to study titles and descriptions separately is hinged on relevance theory. It posits that individuals assess the relevance of information by optimising cognitive efforts through cost-benefit analyses [11]. Information is deemed relevant if it provides maximal value using minimal effort. Specifically, when information is presented as titles and descriptions, the former serves as a relevance optimiser for the latter [12]. Titles of reviews have a bearing on the extent to which the descriptions are perceived as being relevant. Put differently, titles could be utilised as a way to identify descriptions that are worth reading. This in turn suggests that the ways in which title-centric message properties relate to sales will differ from description-centric ones.

With respect to both titles and descriptions, this paper takes into consideration three message properties, namely, length, positivity and negativity. Length was important to study because short titles are known to enhance reading ease [12]. Additionally, length of descriptions has been shown to predict sales [2]. Furthermore, positivity and negativity were studied. This was because texts allow for a verbal representation of numerical ratings via positive and negative sentiment words. In fact, [13] found that verbal valence had a greater impact on purchase intentions than numerical ratings. In a related vein, [14] argued that the valence of a review is better reflected in the textual content rather than the numerical ratings.

II. METHODS

The data for this paper included 118,580 reviews for a panel of 75 Asian hotels posted from January 2011 to December 2014 across three booking-cum-review websites: Expedia.com, Agoda.com and Hotels.com. The data were aggregated at a yearly level per hotel to examine how reviews in a given year (t) could predict online sales in the next year (t+1). The number of reviews posted was used as a proxy for online sales [3]. A yearly interval was appropriate to account for seasonal variations. A log-linear regression model was developed to carry out the investigation. The details of data collection and analysis are as follows.

A. *Sample and Data*

Asia was chosen as the context of investigation. Since the early 2010s, it has been deemed as one of the fastest growing regions in the world in terms of not only domestic but also international tourist arrivals [15,16]. If a large number of these tourists relied on reviews to book hotels online, how the entries were related to online sales during this period makes for an interesting revelation. In particular, five popular tourist destinations were identified: Bangkok, Hong Kong, Kuala Lumpur, Singapore, and Tokyo. Such a setting represents a departure from previous works that studied China [3], the UK [7], and the US [17].

To carry out a meaningful investigation, obtaining data from authenticated booking-cum-review websites—known to facilitate booking followed by verified review submission from bona fide travellers—was imperative. For this purpose, some of the top travel websites were short-listed. From the list, three websites that solicit reviews in the form of ratings, titles and descriptions were identified for data collection. These include Expedia.com, Agoda.com, and Hotels.com.

Next, it was necessary to create a panel of hotels. For this purpose, hotels in a given tourist destination as featured on Expedia.com, Agoda.com, and Hotels.com were filtered based on hotel category. Properties labelled consistently by the three websites as five-stars, one-star, and three-stars (not user ratings, but website-assigned ratings) were treated as luxury, budget and mid-scale hotels respectively. In each category, the top five hotels in terms of the volume of reviews were selected. Thus, from each tourist destination, 15 hotels were selected (5 luxury + 5 budget + 5 mid-scale).

In sum, the panel had 75 hotels (5 tourist destinations x 15 hotels). There were 25 luxury hotels, 25 budget hotels, and 25 mid-scale hotels; which were confirmed in March 2015. A web scraper collected reviews for these hotels posted from January 2011 to December 2014. Some 118,580 reviews were retrieved from the three websites. The yearly volume of reviews per hotel ranged from 58 to 1,597. The dataset is summarised in Table I.

TABLE I. DATASET SUMMARY

Descriptors	**Frequency**
Reviews	118,580
Hotels	75 (25 luxury + 25 budget + 25 mid-scale)
Reviews by website	27,084 reviews from Expedia.com + 76,344 reviews from Agoda.com + 15,152 reviews from Hotels.com
Reviews by year	10,638 reviews in 2011 + 23,218 reviews in 2012 + 38,775 reviews in 2013 + 45,949 reviews in 2014
Reviews by hotel category	50,817 reviews for luxury hotels + 29,501 reviews for budget hotels + 38,262 reviews for mid-scale hotels
Reviews by tourist destination	24,895 reviews for hotels in Bangkok + 21,460 reviews for hotels in Hong Kong + 24,336 reviews for hotels in Kuala Lumpur + 26,602 reviews for hotels in Singapore + 21,207 reviews for hotels in Tokyo

B. *Measures and Analysis*

This paper seeks to answer the question, "How online reviews in a year predict online sales in the next?" Given the data from 2011 to 2014, a three-year panel study was conducted. The premise was to find out how reviews for a given hotel in 2011, 2012 and 2013 could predict online sales in 2012, 2013 and 2014 respectively.

A one-year interval was chosen because it was long enough to allow reviews to manifest their effect on online sales. It also fits well with the common business policy to conduct annual performance evaluation [17,18]. In any case, an interval longer than a year would have been inapt because users seldom read reviews posted more than a year ago to make a booking decision. Furthermore, selecting an interval shorter than a year would have biased the analysis due to seasonal fluctuations dictated by regional public holidays for domestic tourists (e.g., Thai New Year in Bangkok, Hong Kong Special Administrative Region Establishment Day in Hong Kong, Islamic New Year in Kuala Lumpur, Chinese New Year in Singapore, Shōwa Day in Tokyo) and more global festivals for international tourists (e.g., Christmas).

The data were averaged at a yearly level for each of the 75 hotels. In particular, the data corresponding to 2011, 2012 and 2013 had to be averaged based on the independent variables, namely, ratings, title length, title positivity, title negativity, description length, description positivity, and description negativity. There were 16,309 non-English reviews in this pool. The titles and the descriptions of these entries were converted to English using Google Translate.

Additionally, the data corresponding to 2012, 2013 and 2014 required averaging based on the dependent variable of online sales, which was calculated as the number of reviews posted. This is a suitable measure because only bona fide travellers who booked rooms from the authenticated websites of Expedia.com, Agoda.com and Hotels.com could write exactly one review on these platforms [3].

Given a hotel and a year, ratings were calculated based on the average of all the numerical ratings. Title length was calculated as the average number of words per title. Title positivity (negativity) was computed as the proportion of positive (negative) words per title on average. Similarly, description length was calculated as the average number of words per description. Description positivity (negativity) was computed as the proportion of positive (negative) words per description on average. The measurements of length, positivity and negativity were obtained using the Linguistic Inquiry and Word Count tool [19].

Guided by prior works such as [3], a log-linear regression model for the online sales of hotel i in year t+1 was developed based on the reviews received in year t as follows: $\log(\text{online sales}_{i,t+1}) = \text{Constant} + (\psi \times \text{control variables}) + (\beta_1 \times \text{ratings}_{i,t}) + (\beta_2 \times \text{title length}_{i,t}) + (\beta_3 \times \text{title positivity}_{i,t}) + (\beta_4 \times \text{title negativity}_{i,t}) + (\beta_5 \times \text{description length}_{i,t}) + (\beta_6 \times \text{description positivity}_{i,t}) + (\beta_7 \times \text{description negativity}_{i,t})$

Three control variables were added hierarchically in the initial block of the log-linear regression model. First, the effect of the five tourist destinations was controlled by creating dummy variables. Bangkok was treated as the reference for comparison. Second, given that prior works have found online buzz to predict sales [10], the volume of online reviews for hotel i in year t was added as a control variable in modelling the property's online sales in year t+1. It was logarithm transformed to avoid multicollinearity. Third, the effect of rating variance for hotel i in year t was controlled. After all, prior works have often hypothesised that a higher variance in ratings for a hotel results in fewer online bookings albeit without empirical support [3].

The independent variables were added in the next block of the regression model. The motive of the analysis was to identify β1 through β7 separately for luxury, budget and mid-scale hotels by accounting for the control variables. A three-year analysis meant that there were 75 data points per hotel category (25 luxury hotels x 3 years = 75; 25 budget hotels x 3 years = 75; 25 mid-scale hotels x 3 years = 75).

III. Empirical Setup and Results

The initial log-linear regression model suffered from multicollinearity. Variance inflation factors (VIF) suggested that description length and description positivity were the correlated variables (VIF > 5, r = 0.91, $p < 0.001$). There was no basis to aggregate the two variables. After all, despite the statistical correlation, they are conceptually disparate. When even variable transformation did not solve the problem, one of the two had to be dropped.

Specifically, description length was removed. This was because even in the unstable model with multicollinearity, it was nowhere close to statistical significance (p = 0.61 for luxury hotels, p = 0.67 for budget hotels, p = 0.57 for mid-scale hotels). On the other hand, description positivity was statistically significant particularly for budget hotels (p = 0.04). Therefore, removing description length from the model was deemed to be an informatively lossless option while getting rid of description positivity could be lossy.

After dropping description length, the revised model attained statistical stability. The VIF values then ranged from 1.14 to 4.95, confirming no multicollinearity. Table II shows the results (the initial block is omitted for brevity).

With respect to the control variables, the dummy variable corresponding to Kuala Lumpur was negatively related to sales for budget hotels (β = -0.29, $p < 0.05$). This shows that budget hotels in the Malaysian capital had lower sales compared with those in Bangkok (reference for comparison). In addition, higher volume of reviews in a given year was associated with more bookings in the next year for luxury hotels (β = 0.40, $p < 0.001$), budget hotels (β = 0.40, $p < 0.05$), as well as mid-scale hotels (β = 0.64, $p < 0.001$). This confirms the buzz effect of reviews on sales as highlighted in prior research [7,10].

For luxury hotels, the model explained 69.40% variance in the dependent variable. Title length was negatively related to online sales (β = -0.44, $p < 0.001$). In addition, title positivity was positively associated with the dependent variable (β = 0.24, $p < 0.05$).

For budget hotels, the model explained 40.30% variance in the dependent variable. Title positivity was negatively related to online sales (β = -0.51, $p < 0.05$). While description positivity showed a positive relation (β = 0.80, $p < 0.001$), description negativity had a negative association (β = -0.37, $p < 0.05$).

For mid-scale hotels, the model explained 38.80% variance in the dependent variable. Title length was negatively related to online sales (β = -0.30, $p < 0.05$).

TABLE II. Regression Coefficients

	Luxury	**Budget**	**Mid-Scale**
Tourist destination			
Bangkok	Baseline	Baseline	Baseline
Hong Kong	-0.01	0.17	0.11
Kuala Lumpur	0.04	-0.29*	0.06
Singapore	0.01	-0.19	0.10
Tokyo	-0.15	0.06	0.00
Log(Volume)	0.40***	0.40*	0.64***
Rating Variance	0.20	0.12	-0.26
Ratings	0.22	0.06	-0.06
Title Length	-0.44***	0.21	-0.30*
Title Positivity	0.24*	-0.51*	-0.27
Title Negativity	-0.12	-0.07	0.17
Description Positivity	-0.04	0.80***	0.22
Description Negativity	0.17	-0.37*	0.06
Control Variables R^2	**46.00%**	**16.80%**	**24.50%**
Full Model R^2	**69.40%**	**40.30%**	**38.80%**

Note: *** $p < 0.0001$, * $p < 0.05$

IV. DISCUSSION

The key thrust of this paper lies in the finding that titles and descriptions of reviews were able to explain sales. The log-linear regression model developed by including titles and descriptions had decent explanatory power as shown in Table II (R^2_{luxury} = 69.40%, R^2_{budget} = 40.30%, $R^2_{mid\text{-}scale}$ = 38.80%). This in turn extends the literature that has mostly focused on the link from ratings to sales [2,3] by demonstrating that even titles and descriptions of reviews have a crucial role to play in this context [4,12].

Additionally, the paper suggests that a positive relationship between ratings and sales cannot be taken for granted. After all, ratings were not statistically related to sales for luxury ($\beta = 0.22$, $p > 0.05$), budget ($\beta = 0.06$, $p > 0.05$) as well as mid-scale hotels ($\beta = -0.06$, $p > 0.05$)—contradicting works such as [8] but concurring with the likes of [9]. Overall, scholars do not seem to have fully cracked the code on the link from ratings to sales.

Besides, the length and the positivity of titles had a significant bearing on sales. Specifically, length was negatively related to sales for luxury ($\beta = -0.44$, $p < 0.001$) and mid-scale hotels ($\beta = -0.30$, $p < 0.05$). In other words, an increase in the length of titles corresponded with a fall in sales, and vice-versa. Prior research suggests that titles are succinct entries that highlight the gist of descriptions [12]. If they are overly long-winded or lengthy, they fail to play their role of relevance optimiser for descriptions effectively [11]. Therefore, lengthy titles perhaps deterred users' willingness to pay attention to descriptions. In turn, this could have translated to a lower propensity to make bookings on the three websites.

The use of positive words in titles was positively related to sales for luxury hotels ($\beta = 0.24$, $p < 0.05$) but had a negative association for budget hotels ($\beta = -0.51$, $p < 0.05$). A positive relation between title positivity and sales—empirically demonstrated for the first time in this paper—is not hard to fathom.

Interestingly however, an increase in the use of positive words in titles corresponded with a fall in sales for budget hotels. This counter-intuitive finding can be explained in light of the literature on review spam [20]. For hotels that offer no-frills accommodation, positive words in titles were perhaps perceived as being too good to be true. Questioning the genuineness of the reviews, users might have decided against chancing their luck in such budget hotels.

Furthermore, the positivity and the negativity of descriptions were found to have a significant bearing on sales for budget hotels. Specifically, the former showed a positive relationship ($\beta = 0.80$, $p < 0.001$) while the latter exhibited a negative association ($\beta = -0.37$, $p < 0.05$). An increase in the use of positive words in descriptions corresponded with a rise in sales, and vice-versa. Conversely, an increase in the use of negative words in descriptions corresponded with a fall in sales, and vice-versa.

In this vein, [7] found sentiment of reviews—measured through ratings—to have a greater impact on sales for high-end hotels vis-à-vis economy-type properties. In contrast, this paper found sentiment of review descriptions to be significantly related to sales for budget hotels only. For these hotels, title positivity was also found to shape sales. Overall, the findings of the paper justify the a priori assumption on which it was conceived: Titles and descriptions of reviews will predict online sales albeit differently for luxury, budget and mid-scale hotels.

V. CONTRIBUTIONS AND LIMITATIONS

This paper extends the literature [2,3,8,17] by showing that not only ratings but also the textual components of titles and descriptions in reviews predict online sales. Besides, based on the relevance theory [11], the paper conjectured that the ways in which title-centric message properties relate to sales will differ from description-centric ones. This emerged as being generally true. For example, while title positivity was positively related to sales for luxury hotels, description positivity remained a non-significant predictor. Again, while title positivity was negatively related to sales for budget hotels, description positivity exhibited a positive association.

Of particular note was that at least one title-centric message property was significant in predicting sales for luxury (length and positivity), budget (positivity), as well as mid-scale (length) hotels. This was however not true for description-centric message properties, which were significant only for budget hotels (positivity and negativity). Perhaps, titles of reviews are perceived as being more relevant than descriptions, especially for high-end hotels. Further empirical works are needed to illuminate and confirm the possibility.

Apart from extending the literature by shedding light on how titles and descriptions of reviews predict sales, the paper corroborates the buzz effect of reviews on sales highlighted in previous works [7,10]. However, by not being able to reconcile the equivocal prior findings on the link from ratings to sales [8,9], it invites scholars to theorise reasons and mechanisms for the divergent patterns.

On the practical front, this paper helps hoteliers understand how reviews can have a bearing on hotels' sales through booking-cum-review websites such as Expedia.com, Agoda.com and Hotels.com. This understanding is important because it will assist hoteliers understand the extent to which a given website is helping them convert potential travellers to actual customers. The understanding in turn could dictate their business strategies with specific websites.

The findings further show that there is no one size fits all strategy in managing the link from reviews to sales. As shown in Table II, the strongest predictors of sales for luxury, budget and mid-scale hotels were title length, description positivity, and volume of reviews respectively. Luxury hotels could liaise with review website administrators to enhance users' likelihood to write short titles. Budget hotels could improve their sales performance by enticing users to include positive words in descriptions. Finally, mid-scale hotels would do well to solicit reviews in large volumes so as to harness the online buzz effect.

The paper also has implications for review website designers of Expedia.com, Agoda.com and Hotels.com. It found that title length had either a negative or a non-significant relationship with sales. An increase in the length of titles either corresponded with a fall in bookings, or had no effect. Clearly, long titles are detrimental for the fate of businesses. However, short titles can have a positive association with hotels' revenues. Therefore, review website designers are recommended to impose length restrictions so that users are compelled to submit short titles. If the websites can generate bookings for hotels by doing so, it would be a win-win situation for all.

These contributions and implications notwithstanding, this paper is not without limitations. For one, it used the proxy variable of number of reviews to measure online sales. This was informed by the literature [3]. On authenticated booking-cum-review websites such as Expedia.com, Agoda.com and Hotels.com; only travellers who have made a booking are able to post a review. Therefore, volume of reviews should reflect online sales on these websites. Nonetheless, it is expected that not all who book through the trio of Expedia.com, Agoda.com and Hotels.com would write a review. Hence, the findings need to be viewed in light of this underlying self-selection bias among reviewers.

Moreover, the translation validity of the 16,309 non-English reviews in the dataset could not be ascertained. These entries were in several languages ranging from Indonesian (e.g., "…fasilitas kurang…") and Korean (e.g., "…객실은 오래 된 듯 했지만 청소…") to Vietnamese (e.g., "…ổ cắm sạc không…") and Norwegian (e.g., "…Veldig bra…"). They were translated to English using Google Translate in order to be processed by the Linguistic Inquiry and Word Count tool [19]. It would have been better to ascertain the translation validity through approaches such as back-translation. This however could not be carried out due to lack of access to individuals who were effectively bilingual.

Additionally, this paper particularly focused on hotels in five popular Asian tourist destinations, namely, Bangkok, Hong Kong, Kuala Lumpur, Singapore and Tokyo. It is also constrained by the window of data collection. Hence, caution is advocated in generalising the findings.

Going forward, this paper serves as a call for scholars to revisit the link from review ratings to sales. The findings have been largely inconclusive thus far. Interested scholars might want to carry out a comprehensive meta-analysis of the current literature to clarify the inconsistencies in the scholarly understanding. Besides, future research could look to replicate the current panel study with a longer timeframe. The context of investigation could also be extended to sharing economy platforms such as Airbnb, the popularity of which has of late been on the rise.

ACKNOWLEDGMENT

The first author thanks Dr Alton Chua for his supervision of the PhD dissertation, which eventually inspired this work.

REFERENCES

[1] C. Weismayer, I. Pezenka, and C.H.K. Gan, "Aspect-based sentiment detection: Comparing human versus automated classifications of TripAdvisor reviews," In B. Stangl, and J. Pesonen (eds.), Information and communication technologies in tourism, pp. 365-380, Cham, Switzerland: Springer, 2018.

[2] J.A. Chevalier, and D. Mayzlin, "The effect of word of mouth on sales: Online book reviews," Journal of Marketing Research, vol. 43, pp. 345-354, 2006.

[3] Q. Ye, R. Law, B. Gu, and W. Chen, "The influence of user-generated content on traveler behavior: An empirical investigation on the effects of e-word-of-mouth to hotel online bookings," Computers in Human Behavior, vol. 27, pp. 634-639, 2011.

[4] S. Ludwig, K. De Ruyter, M. Friedman, E.C. Brüggen, M. Wetzels, and G. Pfann, "More than words: The influence of affective content and linguistic style matches in online reviews on conversion rates," Journal of Marketing, vol. 77, pp. 87-103, 2013.

[5] N. Hu, N.S. Koh, and S.K. Reddy, "Ratings lead you to the product, reviews help you clinch it? The mediating role of online review sentiments on product sales," Decision Support Systems, vol. 57, pp. 42-53, 2014.

[6] K. Xie, L. Kwok, and W. Wang, "Monetizing managerial responses on TripAdvisor: Performance implications across hotel classes," Cornell Hospitality Quarterly, vol. 58, pp. 240-252, 2017.

[7] I. Blal, and M.C. Sturman, "The differential effects of the quality and quantity of online reviews on hotel room sales," Cornell Hospitality Quarterly, vol. 55, pp. 365-375, 2014.

[8] F. Zhu, and X. Zhang, "Impact of online consumer reviews on sales: The moderating role of product and consumer characteristics," Journal of Marketing, vol. 74, pp. 133-148, 2010.

[9] W. Duan, B. Gu, and A.B. Whinston, "Do online reviews matter?—An empirical investigation of panel data," Decision Support Systems, vol. 45, pp. 1007-1016, 2008.

[10] Y. Liu, "Word of mouth for movies: Its dynamics and impact on box office revenue," Journal of Marketing, vol. 70, pp. 74-89, 2006.

[11] D. Sperber, and D. Wilson, Relevance: Communication and cognition. Oxford, UK: Blackwell, 1986.

[12] S.D. Ascaniis, and U. Gretzel, "What's in a travel review title?," In M. Fuchs, F. Ricci, and L. Cantoni (eds.), Information and communication technologies in tourism, pp. 494-505, Vienna, Austria: Springer, 2012.

[13] A.S. Tsang, and G. Prendergast, "Is a "star" worth a thousand words? The interplay between product-review texts and rating valences," European Journal of Marketing, vol. 43, pp. 1269-1280, 2009.

[14] A. Chong, B. Li, E. Ngai, E. Ch'ng, and F. Lee, "Predicting online product sales via online reviews, sentiments, and promotion strategies: A big data architecture and neural network approach," International Journal of Operations & Production Management, vol. 36, pp. 358-383, 2016.

[15] UNWTO, "UNWTO tourism highlights," 2014, available at: https://www.e-unwto.org/doi/pdf/10.18111/9789284416226

[16] UNWTO, "UNWTO tourism highlights," 2015, available at: https://www.e-unwto.org/doi/pdf/10.18111/9789284416899

[17] W.G. Kim, H. Lim, and R.A. Brymer, "The effectiveness of managing social media on hotel performance," International Journal of Hospitality Management, vol. 44, pp. 165-171, 2015.

[18] B. Hanson, A.S. Mattila, J.W. O'Neill, and Y. Kim, "Hotel rebranding and rescaling: Effects on financial performance," Cornell Hospitality Quarterly, vol. 50, pp. 360-370, 2009.

[19] J.W. Pennebaker, R.J. Booth, and M.E. Francis, Linguistic Inquiry and Word Count Software, Austin, TX: LIWC.net, 2007.

[20] S. Banerjee, and A. Chua, "Authentic versus fictitious online reviews: A textual analysis across luxury, budget, and mid-range hotels," Journal of Information Science, vol. 43, pp. 122-134, 2017.