The Impact of Online Reviews and Ratings on Customer Hotel Preferences, Case Study: TripAdvisor

Isabella Natasha*, Robert La Are*
*Swiss German University, Tangerang, Indonesia
*isabellanatasha25@gmail.com
*robert.are@sgu.ac.id

Abstract: Online hotel reviews is an important source for customers to find out the products and service quality. Online hotel reviews consist of reviews and ratings. The best online reviews website in TripAdvisor. In this era, online reviews are really important and could affect customer preferences in choosing a hotel. The aim of this research is to analyze in what extent do online reviews and ratings affect customer preferences. This study use survey by distributing a total of 100 questionnaires in order to gain the data. Then, the data will be proceed through SPSS to test the validity, reliability, classical assumption, and multiple regression analysis. The results indicated that both of the variables (online reviews and ratings) positively impact customer preferences in choosing a hotel. However, online reviews has a bigger impact towards customer preferences rather than online ratings. It is recommended for TripAdvisor to weed out the fake reviews and guarantee the trustworthy reviews. For the hotel managers, it is recommended to response both, positive and negative reviews that made by customers in TripAdvisor, because it could change someone’s perceptions into the good feeling of the hotel.

Keywords: Online reviews, Online ratings, EWOM, TripAdvisor, Customer preferences

1. Introduction

Nowadays, the internet is the predominant of travel arrangements and has become a new channel for the distribution of tourism information (Buhalis et al., 2008). Internet plays an important role in business today in connecting people around the globe. According to Gerwin (2017), In the tourism and hospitality industry, there are a lot of traveller websites that provides comprehensive information, such as TripAdvisor, Booking.com, Hipmunk, and Expedia. However, according to Banerjee & Chua (2016), TripAdvisor is the world’s most largest and popular travel website community. There are more than five million users who visit the website 30 million times per month on average. TripAdvisor has 435 million reviewers and opinions in it (Xiang and Gretzel, 2010). In addition, PhoCusWright concluded that 92.3% of customers open TripAdvisor as their main travel website (TripAdvisor, 2013). TripAdvisor also provides travel guide informations, such as hotels and vacation rental in the surrounding area, flight tickets, top places to go, top restaurants to eat, and things to do.

2. Literature Review

2.1. Internet

Today, the internet is the most powerful tool in the world, a collection of various services and resources. Internet gives a huge impact to everyone lifes especially in this digital era. In the tourism industry, internet is the major of travel arrangements and has become a new channel for the distribution of tourism information (Buhalis et al., 2008). Internet plays an important role in business and connecting people around the globe.
2.2. **Social Media**

Social media are the various platforms, generally internet based, that enable communication among users using a unique frame of reference. Social media has been cited as a methodology for reducing the knowledge translation gap, creating communities of practice, and reducing traditional hierarchical divisions. (Roland, 2018).

2.3. **Online Review Website (TripAdvisor)**

Today, online review website significantly influences customer behavior and attitude towards customer decision (Xie et al., 2015).

2.4. **Online Reviews**

According to Omer et al. (2014), online reviews can be described as a communication platform that provides information to consumers about the function of the products and services through internet-based technologies. Reviews are text-based and often comprise of large information repositories beyond the analytical capabilities of traditional econometric and statistical methods (Guo et al., 2017).

2.5. **Ratings**

Ratings are numerical and easily understood but have a huge impact on hotel sales. As part of the review process, reviewers are able to allocate star ratings to hotels (one to five stars) in five specific aspects including hotel location, cleanliness, hotel room experience, service quality, value of money, as well as an overall rating of satisfaction (Guo et al., 2017).

3. **Methodology**

3.1. **Population and Sample**

The target population for this study have several criteria. Firstly, the population is hotel customer and should have known about online hotel reviews. Secondly, the population are lives in Indonesia. Using sample to study and find the result that could be generalized as the population point of attention (Sekaran & Bougie, 2013). The one who uses TripAdvisor as their media to read the online reviews and choose a hotel through online review preferences is the criteria to be a sample of this study.

3.2. **Sample Size**

This study will use Z test for a sampling size. Z test is use to measure the sample and determine how many samples will be used in this research, thus the population is unknown. Here is the formula:

\[
\text{Whereas:}
\]

\[ n = \text{Sample size} \]
\[ Z = \text{Z value} \]
\[ \sigma = \text{Standard Deviation} \]
\[ e = \text{Margin of error} \]

The confidence level used for determine the sample is 90% (1,645) with 10% margin of errors.

The calculation shows the minimum number of sample is 68. Therefore, the required respondent is 68. However, to maximize the result, there will be 30 respondents added for the pre-test. A total of 100 respondents will be taken as a sample for this research. In addition, according to Febriana and Munawaroh (2015), if the sample size are below than 100, it is too small, 100-200 is medium.

3.3. **Data Collection Method**

100 questionnaires will be distributed to the representative sample according to the target market within the population. The questionnaire in this study was made using open and closed questions and using likert scale type by using the five-point scale in order for the respondent to choose how strongly he or she agree or disagree with the question given. Likert scale is a tool to measure which item or the grain of the
questions contains a tiered choice, while in this study is given a scale of 1 - 5. 1 is the lowest scale that means strongly disagree, in contrast 5 is the highest scale means strongly agree.

There are a total of 3 variables; Rating ($X_1$), Review ($X_2$), and Customer Preferences ($Y$). All the variables using ordinal scales. In this research, there are 5 rating scales from strongly disagree (1) until strongly agree (5). $X_1$ has 5 questions, $X_2$ has 5 questions, and lastly $Y$ has 5 questions. In conclusion, there are a total of 15 questions in main variables part.

a. Data Collection Method

- Hypothesis #1:
  $H_0$: Reviews have no impact on customer preference in choosing a hotel.
  $H_1$: Reviews have an impact on customer preference in choosing a hotel.

- Hypothesis #2:
  $H_0$: Ratings have no impact on customer preference in choosing a hotel.
  $H_1$: Ratings have an impact on customer preference in choosing a hotel.

4. Result & Discussion

TripAdvisor users in Indonesia have been participated in the study, which the number of respondents meets the minimum required sample for this study. Based on the collected data that has been spread to 100 respondents, most of the TripAdvisor users are female with the number of 67 respondents (67%). TripAdvisor users are mostly in the age of 20-30 years old (53%) and 31-40 years old (44%). 73 respondents are live in Jakarta, which is the highest number of domicile profiles, moreover the highest number of occupational profile is employee with the number of 64 respondents. Lastly, 57% of the respondents admitted that they always use TripAdvisor as their travel guide.

4.1. Multiple Regression Test

The regression using more than one independent variable is called multilinear regression (Deryakulu & Büyüköztürk, 2002). In this study, through variate regression analysis, the relations between rating and review towards customer preferences are analysed, the equation representing the linear relations between the variables are formulated (Uyanik & Guler, 2013).
Based on the table 1, R score of $X_1$ and $X_2$ – $Y$ is 0.573. It represents the value of R. R can also be considered as simple correlation coefficient (Malhotra and Birks, 2007). It shows the correlation between independent variables and dependent variable.

Range score of R is 0 - 1, the higher the correlation coefficient, the stronger and greater relationship they have. If R has 0 score, that means there is no linear association between the variables (Hair et al., 2010).

According to R value interpretation, it can be concluded that R value of $X_1$ and $X_2$ – $Y$ (0.573) is moderate association.

On the other hand, the “Adjusted R Square” column represents a modified version of R Square that has been adjusted for the number of predictors in the model.

As it is shown on the table above, the Adjusted R Square of $X_1$ and $X_2$ – $Y$ is 0.314, it means online reviews and ratings explain 31.4% of the variability of customer preferences and the rest of the percentage is explained by other independent variable(s).

According to the table above, the adjusted R square of $X_1$ shows reviews positively impact customer preferences by 53.8% and adjusted R square of $X_2$ shows ratings positively impact customer preferences by 26.9%. Therefore, it can be seen that reviews has a greater impact than ratings in affecting customer preferences in choosing a hotel.

### 4.2. Hypothesis Test

Hypothesis test is the main goal of this research and it’s used to determined whether a hypothesis can be accepted or not. In this research, there are two ways to do a hypothesis test;

- **Partial**: Using t-Test
- **Simultaneous**: Using F Test

In this research, there are a total of two hypothesis that will be analyzed. The first one is the relation between online reviews ($X_1$) and customer preferences ($Y$). Secondly, the relation between online ratings ($X_2$) and customer preferences ($Y$).
4.2.1. **T-Test**

T-test is conducted to find out whether each independent variable partially predict or not predict the dependent variable. The $H_{1.1}$ and $H_{2.1}$ will be accepted if $t_{score} > t_{table}$ and Sig. < 0.05. On the other hand, the $H_{1.0}$ and $H_{2.0}$ will be accepted if $t_{score} < t_{table}$ and Sig. > 0.05.

$t_{table}$ score can be found by using the degree of freedom (n-k-1) with confidence level of 95% or 5%. Whereas:

- **N**: Total sample
- **K**: Total variables Degree of freedom = n – k – 1
- $100 – 2 – 1 = 97$

With degree of freedom 97, and 0.05 (2 tails), it can be found that $t_{table}$ score is 1.9847.

### Table 3: Hypothesis Test with T-Test Coefficients

<table>
<thead>
<tr>
<th>Model</th>
<th>Unstandardized Coefficients</th>
<th>Standardized Coefficients</th>
<th>95.0% Confidence Interval for B</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>B</td>
<td>Std. Error</td>
<td>Beta</td>
</tr>
<tr>
<td>1</td>
<td>7.029</td>
<td>1.989</td>
<td>3.534</td>
</tr>
<tr>
<td>RASUM</td>
<td>.296</td>
<td>.098</td>
<td>.310</td>
</tr>
</tbody>
</table>

Based on the table above, it can be shown that $t_{score}$ for variable reviews ($X_1$) – customer preferences ($Y$) is 3.267 (> 1.9847) with sig. 0.002 (< 0.05). $t_{score}$ for variable ratings ($X_2$) – customer preferences ($Y$) is 3.035 (> 1.9847) with sig. 0.003 (< 0.05). Therefore, the hypothesis test of $H_{1.1}$ and $H_{2.1}$ using t-Test are both accepted.

4.2.2. **F-Test**

F-test is conducted to find out whether each independent variables simultaneously predict or not predict the dependent variable. The hypothesis will be accepted if $F_{score} > F_{table}$ and Sig. < 0.05.

$F_{table}$ score can be found by using the formulation as follows: (k; n-k) with confidence level of 95% or 5%. Whereas:

- **N**: Total sample
- **K**: Total variables
- **K; n - k**
- $2; 100 – 2 = 98$

With the number of 98, and 0.05 (2 tails), it can be found that $F_{table}$ score is 3.089

Source: Data Processing Result
Table 4: Hypothesis Test with F-Test

ANOVA

<table>
<thead>
<tr>
<th>Model</th>
<th>Sum of Squares</th>
<th>Df</th>
<th>Mean Square</th>
<th>F</th>
<th>Sig.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Regression</td>
<td>204.010</td>
<td>2</td>
<td>102.005</td>
<td>23.661</td>
<td>.000*</td>
</tr>
<tr>
<td>Residual</td>
<td>418.180</td>
<td>97</td>
<td>4.311</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>622.190</td>
<td>99</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Predictors: (Constant), RASUM, RESUM
Dependent Variable: CPSUM

Source: Data Processing Result

Based on the table above, it can be shown that F score is 23.661 (> 3.089) with sig. 0.000 (< 0.05). Therefore, the hypothesis test of H_{1.1} and H_{2.1} using F-Test are both accepted.

4.3. Hypothesis Test Summary

According to the hypothesis test that has been conducted, there is a hypothesis test summary to conclude it all.

Table 5: Hypothesis Test Summary

<table>
<thead>
<tr>
<th>Hypothesis</th>
<th>t-Test (Partial)</th>
<th>F Test (Simultaneous)</th>
<th>Hypothesis Conclusions</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>t score &gt; t table (1.98 (47)</td>
<td>Sig. &lt; 0.05</td>
<td>F score &gt; F table (3.089)</td>
</tr>
<tr>
<td>H_{1.1}: RE-CP</td>
<td>3.267</td>
<td>0.002</td>
<td>23.661</td>
</tr>
<tr>
<td>H_{2.1}: RA-CP</td>
<td>3.035</td>
<td>0.003</td>
<td></td>
</tr>
</tbody>
</table>

Source: Data Processing

Based on the table above, it can be concluded that H_{1} in hypothesis #1: “Reviews have an impact on customer preference in choosing a hotel” is accepted and H_{0} is rejected. Moreover, H_{1} in hypothesis #2: “Ratings have an impact on customer preference in choosing a hotel” is accepted and H_{0} is rejected.

4.4. Discussion about The Impact of Online Reviews (X_{1}) On Customer Preferences (Y)

According to Buhalis et al. (2008), online reviews give a big impact and influence customer decision making when selecting a product. In addition, online reviews are also believed as an important source to find out about product quality and service excellence (Dickinger, 2015). In this research, the result is positive, which means online hotel reviews have an impact towards customer preferences in choosing a hotel.

According to the table of correlation of X_{1} towards Y, the adjusted R square is 0.538. Therefore, it can be concluded that online reviews have an impact on customer preferences by 53.8%.

Based on the questionnaire that has been spread, almost all of the respondents are reading online hotel reviews before they book a hotel, they read online hotel reviews to find out the hotel’s service and product quality. Most of the respondents read online hotel reviews on lesser-known hotel or budget hotel than well-known hotels, that proves online hotel reviews have a bigger impact towards lesser-known hotels than well-known hotels. Positive reviews make respondents feel confident to book on that hotel. Moreover, the larger volume of online hotel reviews reflects that many people are interested in that hotel.
However, according to the respondents (33% said strongly agree and 34% said agree), they have a bad impression when they read some of the negative reviews. As Papathanassis and Knolle (2011) had observed, negative reviews spread faster than positive reviews.

4.5. Discussion about The Impact of Online Ratings ($X_2$) On Customer Preferences ($Y$)

Nowadays, online ratings have become an important source of product information about the quality of services (Gavilan et al., 2018). The numerical scale of rating makes customer easier to know about the product quality and its influence customer when choosing a product (Chen & Xie, 2008). In this research, the result is positive, which means online ratings positively impact customer preferences in choosing a hotel.

According to the Correlation of $X_2$ towards $Y$ table, the adjusted R square is 0.269. Therefore, it can be concluded that online ratings have an impact on customer preferences by 26.9%, it has half smaller effect rather than online reviews (53.8%).

Based on the questionnaire that has been spread, most of the respondents agreed that higher ratings indicate the hotel has a good reputation. In contrast, low ratings reflect that the hotel has a bad service and quality. Therefore, the level (low-high / 1-5) of ratings measure the value of the hotel. The respondents tend to choose a hotel with higher rating rather than lower rating, however high ratings are trustworthy if supported by a high number of reviews.

5. Conclusion and Recommendation

5.1. Conclusion

The objective of this research is to find out in what extend online hotel reviews and ratings impact customer preferences in choosing a hotel. To answer the research questions, multiple linear analysis test on SPSS will be used to check whether the hypothesis is accepted or not.

To conclude all the results, there are some important points from this research:

1. According to the questionnaire that has been spread to 100 respondents, most of the TripAdvisor users are female with the number of 67 respondents (67%). TripAdvisor users are mostly in the age of 20-30 years old (53%) and 31-40 years old (44%). 73 respondents are live in Jakarta, which is the highest number of domicile profiles.
2. moreover the highest number of occupational profile is employee with the number of 64 respondents. Lastly, 57% of the respondents admitted that they always use TripAdvisor as their travel guide.
3. Online reviews and ratings influence customer preferences in choosing a hotel by 31.4%.
4. Online reviews positively impact customer preferences in choosing a hotel by 53.8%. Therefore, the first hypothesis ($H_{1.1}$) is accepted.
5. Online ratings positively impact customer preferences in choosing a hotel by 26.9%. Therefore, the second hypothesis ($H_{2.1}$) is also accepted.
6. According to the result of the correlation, online hotel reviews has a bigger impact than online ratings in affecting customer in choosing a hotel. The correlation of $X_1$ towards $Y$ is twice bigger than $X_2$ towards $Y$.

5.2. Recommendation

The results of this research have been able to provide some recommendations for managerial implication and further research.

5.2.1. Recommendation for Managerial Implication

Firstly, for TripAdvisor, there are a lot of people using TripAdvisor. However, everyone can make a review whether it’s fake or not. It is recommended to filter and weed out the fake reviews, thus only genuine reviews are shown. Take Zomato as an example, they have trustworthy reviews stamp on the application. On top of that, it will make the customers trust the reviews and assume that TripAdvisor is a trustworthy source.
Secondly, for hotel companies, customers are more reliant on online reviews because they assume customer reviews are more reliable than companies (Tong Bao, 2014). As it is also shown on the result of correlations, online reviews and ratings positively impact customer preferences in choosing a hotel. However, the customer will have a bad impression whenever negative reviews are occur, it is recommended for the hotel manager in response all the reviews, both positive and negative reviews, because it will change someone’s perception if the manager response nicely and will boost up the image of the hotel.

5.2.2. Recommendation for Further Research

There are some recommendations for future research who wants to observe the similar objects.

This study focuses on online hotel reviews and online ratings, it is recommended to add more indicator or the value of the variables. Therefore, it can be seen which aspects has the most significant impact towards customer hotel preferences.

Furthermore, this study focuses only on the hotel industry, it is also recommended to add or use another component of the object research, thus it can be used for a better purpose in another industry, such as F&B, online shopping, and so on.

Last suggestion is to add the number of respondents and variety of questionnaire, in a purpose of getting more precise and accurate data in solving a problem. The future research might also use other approaches to gather the data and use another method for data processing.

References:


