SENTIMENT ANALYSIS ON PERCEPTION OF TOURISM IN DKI JAKARTA PROVINCE: CASE STUDY OF OLD CITY AND KEPULAUAN SERIBU

^{*1}Winata, Supriyadi Ramdan, ²Pangestuty, Farah Wulandari

^{*1}Bank Indonesia, Indonesia

²Faculty of Economics and Banking, Universitas Brawijaya, Malang – Indonesia

Note: * Indicates corresponding author

ARTICLE DETAILS	ABSTRACT
Article History Published Online: publisher use only	The purpose of this study is to determine the perceptions of tourists in two tourist destinations in DKI Jakarta which are included in the National Tourism Strategic Area (KSPN), namely Kota Tua and Kepulauan Seribu by conducting a sentiment analysis. The analysis was carried out on traveler reviews in the form of electronic word of mouth (eWOM) submitted on the tripadvisor site. The results of the study found that in general the sentiment of tourists towards these two destinations was positive, and each year there was an increasing trend. However, based on the exploration of sentiments that have negative connotations, it is known that the problems of odor, garbage and pollution are of great concern to tourists. Other aspects of concern to tourists are the need for better governance by local governments, and the need for standard facilities and services.
Keywords Sentiment Analysis, Tourism, Jakarta, Sentiment Lexicon, Match Ratio Lexicon	
*Corresponding Author Email: supriyadi_rw@bi.go.id	

Introduction

World Tourism Organization (UNWTO) notes that tourism is the world's largest export commodity. In developing countries, tourism is generally the main source of economic growth, foreign exchange, investment and job creation. In addition, the World Travel and Tourism Council (WTTC) predicts tourism contributes more than 10% of world GDP, absorbs around 8% of the world's workforce, and 12% of total world exports.

Despite having great potential, tourism development in DKI Jakarta has not been carried out holistically and integratedly, resulting in low competitiveness of Jakarta tourism. In terms of global tourism competition, Jakarta is still far behind in terms of the level of visits by foreign tourists. According to the *Master card Destination Cities Index* in 2017, Bangkok was ranked first out of 132 major destination cities in the world. The number of foreign tourists visiting Bangkok in 2016 reached 19.41 million people, with total expenditures of USD 14.1 billion. Meanwhile, the number of foreign tourists visiting Jakarta was only around 1.6 million with total expenditures of around USD 1.2 million.

However, DKI Jakarta is one of the five fastest growing destination cities, reaching 18.2% CAGR¹, along with Osaka (24.0%), Chengdu (22.7%), Colombo (20.3%), and Abu Dhabi (18.9%).

Nationally, Jakarta is the second highest area in terms of tourist visit rates after Bali. However, the majority of tourists who come to Jakarta are for business, so the length of tourist visits is lower. If the average length of tourist visit in Bali is 3.49 days for star hotels and 3.11 days for non-star hotels, then the length of visit in Jakarta is 2.08 days and 1.36 days, respectively. Thus there is a potential to increase the length of tourist visits to various destinations in Jakarta.

With all the potential that exists, the contribution of the tourism sector should be further improved, especially in Jakarta, there are destinations that are included in the development of the National Tourism Strategic Area (KSPN), namely Kota Tua and the Kepulauan Seribu. In addition, amid the declining share of the manufacturing industry due to weakening carrying capacity, tourism can become a new source of growth for the Jakarta economy. Therefore, it is very important for

¹ Compound annual growth rate

Jakarta to understand the factors that influence the level of tourist visits, and to find out the perspectives of tourists regarding tourist destinations in DKI Jakarta.

Tourism is an information-intensive industry. The development of the *internet of things* has encouraged tourists to share their experiences on various social media, submit opinions or reviews, post photos, and give ratings on travel blogs and on various sites such as tripadvisor.com. *Electronic word of mouth* (eWOM) like this is then also used by other potential tourists, in choosing certain destinations. Tourists often take advantage of opinions from social media like this, because it is not easy to evaluate tourism products and services if they do not experience them firsthand (Litvin, Goldsmith, & Pan, 2008).

Due to the rich content of information contained in eWOM, analysis of eWOM is one of the most efficient and powerful methods for understanding consumer feelings towards services, service providers, and certain products (Cheung, Luo, Sia, & Chen, 2009). By knowing the opinions of tourists, policy makers can formulate and implement appropriate policies, in accordance with consumer expectations, to increase the level of tourist visits.

Analysis using eWOM has not been widely carried out in Indonesia. Thus, this research is expected to contribute to encouraging the use of eWOM as a resource in policy formulation. This study aims to conduct a *sentiment analysis* on tourists' perceptions of two tourist objects in DKI Jakarta which are included in the development of the National Tourism Strategic Area (KSPN), namely Kota Tua and Kepulauan Seribu. This perception is obtained from traveler reviews provided through the tripadvisor site.

Literature Review

Consumers usually tend to imitate a certain experience, and they will tell each other. This communication process through *word of mouth* (WOM) allows consumers to share information and opinions. Such information and opinions can lead someone to approach or stay away from certain products, brands or services (Hawkins, Best and Coney 2004 in Litvin, Goldsmith, and Pan, 2008).

The study of WOM has developed since the 1960s and its definition continues to grow until now (Carl, 2006). At first, WOM was defined as face-to-face communication about a product or company, between people who are not commercial entities (Arndt, 1967). Meanwhile, Westbrook (1987) defines WOM as all informal communication activities aimed at other consumers to describe the ownership, use, or

characteristics of certain goods and services, or the seller. Westbrook does not specifically define informal communication clearly, but his writings indicate that it is communication built from interpersonal relationships, which is different from communication through mass media channels that aim to provide *product knowledge* from producers to consumers (Litvin, Goldsmith, & Pan, 2008). Then Buttle (1998) argues that WOM can be conveyed via electronic channels (eWOM).

With the growing use of social media by the community, now we can easily get eWOM on many things, including tourism. People are increasingly accustomed to sharing photos, or writing reviews or opinions about tourist objects that have been visited. This then attracted the interest of many researchers to take advantage of the eWOM, so that more structured information, knowledge and data could be obtained.

The method that is widely used in analyzing eWOM is *text mining*. As a technology that is still developing, *text mining* is very useful for extracting important information from a number of documents quickly (Liu et al, 2011 and He et al, 2013). The main purpose of *text mining* is to find useful models, trends, patterns, or an order from unstructured text data (Romero et al, 2008; He et al, 2015; and He et al 2015).

One application of *text mining* in particular that is now increasing in popularity is *sentiment analysis*. This method aims to extract positive and negative opinions from a particular text (Pang and Lee, 2004). An example of the use of *sentiment analysis* in tourism is carried out by Valdivia, Luzon and Herrera (2017), who conducted a *sentiment analysis* on tripadvisor reviews of three important monuments in Spain, namely the Alhambra, Mezquita Cordoba and the Sagrada Familia. The study found that tourists rated the three monuments positively, with more than 90% giving them a 4 and 5 rating (on a scale of 1-5). However, the researchers observed that there was a tendency that the respondents gave positive ratings but negative reviews, or vice versa. Therefore, the researcher affirms that the analysis is carried out in depth on the reviews, not only limited to the final score given. Reflecting on the research by Valdivia, Luzon and Herrera (2017), this study will also focus on exploring sentiment in reviews provided by tourists.

Data and Methodology

This study will use tourist reviews sourced from the tripadvisor.com site, both domestic and foreign tourists, for two tourist destinations, namely the Kepulauan Seribu and Kota Tua.

SENTIMENT ANALYSIS ON PERCEPTION OF TOURISM IN DKI JAKARTA PROVINCE: CASE STUDY OF OLD CITY AND KEPULAUAN SERIBU

Reviews for the two destinations have been available since 2010, but due to the relatively low frequency of reviews, this study will use reviews for the period 2012 to June 2018. During that period, there were 626 reviews for the Kepulauan Seribu, while for Kota Tua totaled 790.

Figure 1. Distribution of Kep. Seribu Reviews

Figure 2. Tourist Categories of Kep. Seribu

Until 2016, the number of tourists providing reviews about the Kepulauan Seribu continued to increase, reaching 166. However, in 2017 the number fell to 92, while up to the first semester of 2018, the number of tourists who gave new reviews was recorded as many as 35 people (Figure 1). From its composition, the majority of tourists are still dominated by domestic tourists with a total of 50%, while foreign tourists are 29%, then the remaining 21% are not identified because they do not provide information (Figure 2).

Figure 3. Region of Origin of Kep. Seribu

Figure 4. Country of Origin of Foreign Tourists Kep. Seribu

Of the total number of foreign tourists, the majority came from Asia and Europe, with a share of 37% and 34%, respectively. The American region accounts for 16%, Australia 12%, and the remaining 1% comes from the African region (figure 3). Meanwhile, if viewed by country, the five countries with the highest number of tourists were the United States (13%), Australia (11%), India (9%), then Singapore and the UK each contributed 6% (figure 4).

Figure 5. Distribution of Kota Tua Reviews

Figure 6. Category of Kota Tua Tourists

Meanwhile, for Kota Tua, the number of tourists who provide reviews continues to decline until 2014. The number of reviews has then increased again and in 2016 there were 197 reviews. As is the case in the Kepulauan Seribu, in 2017 this number fell to 154, while up to the first semester of 2018, the number of tourists who gave new reviews was 89 people (Figure 5). From its composition, the majority of tourists are dominated by foreign tourists with a total of 50%, while domestic tourists are 29%, then the remaining 21% are not identified because they do not provide information (Figure 6). This condition is different from the Kepulauan Seribu, which is dominated by domestic tourists.

Of the total number of foreign tourists, the majority came from Europe and Asia, with a share of 35% and 32%, respectively. Tourists from America have a share of 20%, then Australia 11%, and the remaining 2% come from Africa (Figure 3). Meanwhile, if viewed by country, the four countries with the highest number of tourists are the United States (14%), then Malaysia, the UK, and Australia each contributing 10% (figure 4).

Figure 7. Continent of Origin of Kota Tua Tourists

Figure 8. Country of Origin of Kota Tua Tourists

Reviews that have been chosen will then be processed using *sentiment analysis*. Please note that all reviews that will be analyzed are in English translation, because the method used requires this. In simple terms, the data processing process in this study can be illustrated in Figure 9. Text data in the form of reviews from tourists that have been translated into English, sourced from the tripadvisor.com website, will first be processed into several stages, including correction of abbreviated words, elimination. Special characters and numbers, convert all words to lowercase, remove excess spaces, and remove unnecessary words. Some words that the author omits in the review are general words but do not eliminate the meaning of the review, such as Indonesia, Jakarta, Rupiah, and several other words.

Figure 9. Data Analysis Process

Figure 8. Country of Origin of Kota Tua Tourists
Source: Silge and Robinson, 2017

After the data is tidied up, the next step can be carried out by the process of extracting *sentiment* by comparing it against the sentiment dictionary, or what is known as the *sentiment lexicon*. According to Silge and Robinson (2017), in general, there are three *lexicon sentiments* that are widely used, namely AFINN (from Finn Arup Nielsen), Bing (from Bing Liu et al), and NRC (from Saif Mohammad and Peter Turney).

The three *lexicons* are based on a *unigram*, or single word, and the word is associated with a *sentiment* or certain emotion. *NRC lexicon* categorize the word *binary* ("yes" / "no") into the category of positive, negative, anger, anticipation, disgust, fear, joy, sadness, surprise, and trust. *Bing lexicon* categorizes words *binary* into positive and negative categories. Meanwhile, the *AFINN lexicon* attributes words

SENTIMENT ANALYSIS ON PERCEPTION OF TOURISM IN DKI JAKARTA PROVINCE: CASE STUDY OF OLD CITY AND KEPULAUAN SERIBU

with scores ranging from -5 and 5, with negative scores indicating negative sentiment and positive scores indicating positive sentiment. Not all words in the English language are in the *lexicon* because many words have neutral properties. This method also doesn't catch words preceded by *qualifiers* like "no good" or "not true", because this method is based on unigram.

Results and Discussion

This section will present the results of text data processing based on the *sentiment analysis* method. Section 4.1 will present the results for Kepulauan Seribu, while section 4.2 will present the results for Kota Tua.

Sentiment Analysis Results for the Kepulauan Seribu

Based on the mapping with the *word cloud* (figure 10), the words that appear most frequently in tourist reviews is boat/ship. This is not surprising because the only mode of transportation that can be used to reach Kepulauan Seribu is by sea, so the word "boats" is one of the keywords used by tourists. Meanwhile, other words that are included in the 10 words with the highest frequency in reviews can be seen in Figure 11. "Beach" and "snorkeling" are also often mentioned by tourists which indicate activities that are often carried out by tourists. "Nice" and "beautiful" also have a high frequency of occurrence, which reflects the sense of admiration by tourists for Kepulauan Seribu.

Figure 10. Word Cloud Reviews of Kepulauan Seribu

Figure 11. Words with the Highest Frequency in Kepulauan Seribu Reviews

As explained in the previous section, there are three *lexicons* that are commonly used in conducting *sentiment analysis*, namely *AFINN*, *bing*, and *nrc*. To determine which *lexicon* is the best to use, select the *lexicon* that has the highest match ratio. In table 1 it can be seen that the *lexicon nrc* has the highest match ratio value of 0.21, while the *AFINN* lexicon has the lowest *match ratio* of 0.10.

Table 1. Match Ratio of Lexicon for Kepulauan Seribu

lexicon	lex_match_words	words_in_body	match_ratio
AFINN	416	4084	0.1018609
bing	596	4084	0.1459354
nrc	876	4084	0.2144956

After the most appropriate *lexicon* is chosen, next we can look at the frequency of each sentiment. Based on Figure 12, it can be seen in general that positive sentiment dominates the reviews of Kepulauan Seribu tourists, followed by joy, trust, then anticipation, while the frequency of negative sentiment is relatively low. We can also see the division of positive and negative sentiment based on *bing lexicon* in Figure 13. This figure also shows that positive sentiment dominates, while the frequency of negative sentiment does not reach half.

Figure 12. Frequency of Each Sentiment According to the NCR Lexicon

Figure 13. Frequency of Each Sentiment According to the Bing Lexicon

Figure 14. Polarity of Sentiment Direction of Kepulauan Seribu

After knowing the sentiment frequency in general, we can also see the polarity of sentiment throughout the observation period (Figure 14). Sentiment polarity has tended to decline in recent years, however, this is more due to the lower frequency of reviews. Based on the ratio of positive sentiment to total sentiment, we can observe that there is a positive trend. This shows that during the observation period, the positive rating of tourists towards Kepulauan Seribu tends to increase. However, the increase in the ratio of positive sentiment tends to form a relatively gentle slope. This means that there are no significant improvements that can make tourists give higher appreciation.

Figure 15 illustrates how the sentiment categories relate to time. It can be seen that every year, positive sentiment always dominates the reviews, followed by "joy" and "trust". This is in line with the polarity of the sentiment direction which shows that there is a positive trend regarding the sentiment of tourist reviews. We can also observe that the number of negative sentiments tends to increase in 2016 and 2017.

Figure 15. Relationship between Sentiment and Year in Kepulauan Seribu

SENTIMENT ANALYSIS ON PERCEPTION OF TOURISM IN DKI JAKARTA PROVINCE: CASE STUDY OF OLD CITY AND KEPULAUAN SERIBU

Table 2 presents the top five words that describe each sentiment in the Kepulauan Seribu review. In this case, one word can be associated with more than one sentiment. For example, "food" can be associated with "joy" and "positive". This means that tourists judge the taste of food in Kepulauan Seribu in general according to their tastes, so that it creates a sense of pleasure and becomes a positive value in the eyes of tourists.

Because the majority of tourists tend to have sentiments that have a good connotation, a more detailed discussion will be made of those with a bad connotation, namely "anger", "disgust", "fear", and "negative". In the sentiment of "anger", the five main words that appear are "money", "limited", "hot", "disappointed", and "bad".

Table 2. Words that Describe Every Sentiment on Reviews of Kepulauan Seribu

Sentiment	Five Main Words				
Anger	Money	Limited	Hot	Disappointed	Bad
Anticipation	White	Time	Sunset	Friendly	Enjoy
Disgust	Pollution	Garbage	Finally	Dirty	Bad
Fear	Swim	Escape	Destination	Bad	Avoid
Joy	Food	Enjoy	Clean	Beautiful	Beach
Negative	Garbage	Escape	Dirty	Cheap	Bad
Positive	Visit	Sea	Food	Clean	Beautiful
Sadness	Quiet	Lovely	Destination	Blue	Bad
Surprise	Trip	Sun	Money	Expect	Destination
Trust	White	Recommend	Food	Enjoy	Clean

"Money" appears in angry sentiments because tourists feel the money spent is not

worth the benefits. People are seen as just trying to get money, without providing good service. In some cases, tourists are also charged non-standard rates or rates that change when doing activities on the Kepulauan Seribu.

The word "limited" appeared in angry sentiments because tourists considered the existing facilities in Kepulauan Seribu, especially for proper accommodation, were still considered limited. Meanwhile, the word "hot" represents two aspects, namely hot weather which makes tourists uncomfortable, and the unavailability of hot water facilities in the accommodation, both for bathing and for making drinks. The unavailability of hot water for bathing has been complained of especially by foreign tourists, because of the habits they practice in their home countries. However, there are also those who do not mind this, because the hot weather conditions make them simply take a shower in cold water.

The word "disappointed" appears mainly to describe two aspects, namely disappointment with cleanliness and unkempt natural conditions, and disappointment with the services provided to tourists. This service is mainly related to tour guides, because there are foreign tourists who think that the tour guide provides all the facilities promised. The word "bad" arises because many tourists see the poor condition of the island due to irregularity and lots of garbage, the bad condition of the coral reefs because there has been a lot of damage, and poor public transportation to Kepulauan Seribu.

On the sentiment "disgust", five words that appear most frequently are "pollution", "garbage", "finally", "dirty", and "bad". Of the five words, three words are closely related, namely "pollution", "garbage", and "dirty". This shows that environmental cleanliness is still a problem that needs serious handling, because it will reduce the interest of tourists to visit again, or to carry out promotions through WOM and eWOM. If conditions like this are allowed, it will certainly cause bad impacts, both in the short term and in the long term.

In the sentiment "fear", five words that are often referred to are "swim", "escape", "destination", "bad" and "avoid". The word "swim" appears in the sentiment of fear mainly due to tourists' fear of strong currents / waves, trash, and sea urchins when carrying out these activities. The word "escape" appears because Kepulauan Seribu has been considered as a place to escape from the hubbub of Jakarta, but there are fears from tourists about the condition of Kepulauan Seribu in general.

On negative sentiments, the five words that often appear are "garbage", "escape",

"dirty", "cheap", and "bad". The word "cheap" often appears in negative sentiment because low prices are accompanied by low quality, thus becoming a negative assessment for tourists. The word "cheap" is generally mentioned in the context of the transportation costs used to and from Kepulauan Seribu.

Figure 16. *Bigram Network* for Negation Words in Kepulauan Seribu Reviews

As explained in the previous section, the *lexicon* in *sentiment analysis* is based on unigram, or single word. Therefore it is also necessary to look at how the two words together, or *bigram*, form sentiment. Figure 16 maps in full how these words appear in the Kepulauan Seribu review. Based on the identification results, the bigram is only preceded by "not", "no", and "never". It can be seen that the words beginning with "not" have more numbers than the other two.

Figure 17. Words preceded by “Not” in the Kepulauan Seribu Review

Picture 16 shows the whole word, while in Figure 17 shows the word that begins with “not” which has a big contribution to sentiment,

both in positive and negative directions. It can be seen that for the negative direction, the word "*not recommend*" or "not recommended" gave the greatest contribution. The word "not recommended" is used by tourists in several contexts, but the majority are not advised to stay on a particular island, either because of the dirty environmental conditions, or because the accommodation is considered unrepresentative.

Meanwhile, in a positive direction, “*not worry*” or “*don't worry*” gave the largest contribution to sentiment. The word “*don't worry*” is used by tourists in the context of describing the friendly conditions of the people of Kepulauan Seribu, so there is no need to worry about visiting there.

Sentiment Analysis Results for Kota Tua

Based on the mapping with the *word cloud* (Figure 18), the word that appears most frequently in Kota Tua tourist reviews is museum. This is not surprising, as there are six museums in the City Area that attract tourists. The museum is divided into two groups, namely museums belonging to the Provincial Government of DKI Jakarta (Jakarta History Museum, Fine Arts and Ceramics Museum, Wayang Museum, and Maritime Museum) and non-provincial government museums (Bank Indonesia Museum and Mandiri Museum). Thus, in order to attract a higher level of tourist visits, better museum management is a very important factor.

Figure 18. Word Cloud of Kota Tua Reviews

Figure 19. Words with the Highest Frequency

SENTIMENT ANALYSIS ON PERCEPTION OF TOURISM IN DKI JAKARTA PROVINCE: CASE STUDY OF OLD CITY AND KEPULAUAN SERIBU

While other words included in the 10 words with the highest frequency in the review can be seen in Figure 19. The words "building" and "square" (field / park) is often referred in reviews because it describes its destination. The Kota Tua area is still synonymous with Fatahillah Park and various old buildings, although there are many other attractions.

Of the three *lexicons* used, it can be seen in table 3 that the *nrc lexicon* has the highest *match ratio* value of 0.21, while the AFINN lexicon has the lowest *match ratio* of 0.10. These results are relatively the same as the findings for the Kepulauan Seribu review. Thus, Kota Tua will also use the *nrc lexicon*.

Table 3. Match Ratio Lexicon for Kota Tua

lexicon	lex_match_words	words_in_body	match_ratio
AFINN	435	4210	0.1033254
bing	637	4210	0.1513064
nrc	903	4210	0.2144893

Based on figure 20, it can be seen that in general positive sentiment dominates Kota Tua tourist reviews, followed by "trust", "joy", then "anticipation". Meanwhile, the frequency of negative sentiments is relatively low. We can also see the division of positive and negative sentiments based on the *bing lexicon* in Figure 21. The figure also shows that positive sentiment dominates, although the portion of negative sentiment is larger than the result in the Kepulauan Seribu review.

Figure 20. Frequency of Each Sentiment According to NCR Lexicon

Figure 21. Frequency of Each Sentiment According to Bing Lexicon

Figure 22. Polarity of Sentiment Direction of Kota Tua

Figure 22 shows the polarity of the sentiment direction of Kota Tua. As in the Kepulauan Seribu review, the polarity has tended to decline in recent years, but this is more due to the decline in the frequency of reviews. In trend terms, the polarity tends to move up. Based on the ratio of positive sentiment to total sentiment, we can observe that in 2015 there was a decrease in the ratio, but this ratio continued to increase until 2017, and in 2018 it declined again. However, in general there is a positive trend, which means that during the observation period, the positive rating of tourists towards Kota Tua has increased relatively.

Figure 23 illustrates how the sentiment categories relate to time. It can be seen that every year, positive sentiment always dominates the reviews, followed by “joy” and “trust”. This is in line with the polarity of the sentiment direction which shows that there is a positive trend regarding the sentiment of tourist reviews. We can also observe that with the increasing number of reviews in 2016 and 2017, sentiments with bad connotations, such as “negative”, “disgust”, and “anger” also increased.

Figure 23. Relationship between Sentiment and Years in Kota Tua

Table 4 presents the top five words that describe each sentiment in the Kota Tua review. As in the Kepulauan Seribu review, the majority of tourists tend to have sentiments that have a good connotation, so a more detailed discussion will be made of those with bad connotations, namely “anger”, “disgust”, “fear”, and “negative”.

The main five words that appear in an angry sentiment are “smell”, “money”, “hot”, “bad”, and “abandoned”. The word “smell” appears in many angry sentiments because tourists are very disturbed by the smells arising from highly polluted rivers or canals.

The word “money” is often mentioned by tourists in several contexts, namely (1) Tourists think that they should spend more money on arranging Kota Tua to make it more attractive, (2) Some tourists feel annoyed by the existence of individuals asking for money when tourists want to take photos certain objects, and (3) some cafes / restaurants considered too expensive by tourists. The word “hot” often appears because many tourists cannot stand the hot weather. Then, the words “bad” and “abandoned” are widely used by tourists to describe the condition of buildings that have been left damaged and abandoned.

Table 4. Words that Describe Every Sentiment on Reviews in Kota Tua

Sentiment	Five Main Words				
Anger	Smell	Money	Hot	Bad	Abandoned
Anticipation	Time	Public	Fun	Friendly	Enjoy
Disgust	Smell	Garbage	Finally	Dirty	Bad
Fear	Watch	Government	Bad	Avoid	Abandoned
Joy	Love	Fun	Food	Enjoy	Beautiful
Negative	Government	Forget	Dirty	Cheap	Bad
Positive	Worth	Visit	Food	Café	Building
Sadness	Music	Late	Bad	Art	Abandoned
Surprise	Unique	Trip	Shopping	Money	Art
Trust	Guide	Food	Enjoy	Bank	Architecture

SENTIMENT ANALYSIS ON PERCEPTION OF TOURISM IN DKI JAKARTA PROVINCE: CASE STUDY OF OLD CITY AND KEPULAUAN SERIBU

Regarding sentiment *disgust*, the words that often appear are “*smell*”, “*garbage*”, “*Finally*”, “*dirty*” and “*bad*”. There are three closely related words, namely “*smell*”, “*garbage*”, and “*dirty*”, which refer to the condition of a clean environment. From this it can be seen that tourists have a very high attention to the cleanliness of the environment, because a comfortable environment will attract them to come and stop or stay longer.

On the sentiment of *fear*, the words that often appear are “*watch*”, “*government*”, “*bad*”, “*avoid*”, “*abandoned*”. The word “*watch*” appears in the sentiment of “*fear*” because many tourists feel afraid when they see the heavy traffic in Kota Tua, and cause them to be afraid when they want to cross the road to a certain destination.

The word “*government*” is often mentioned because tourists think that local governments should be able to do more of their role in structuring Kota Tua Area to make it friendlier to tourists. On the word “*avoid*”, many tourists write in their reviews in order to avoid the weekend to visit because it is very crowded, and avoid the heat of the weather.

On negative sentiment, the words that often appear are “*government*”, “*forget*”, “*dirty*”, “*cheap*”, and “*bad*”. The word “*cheap*” arose in negative sentiment because tourists considered the cost to enter the museum in Kota Tua to be very cheap, but indeed the collection is very insufficient. This then becomes a negative assessment in the eyes of tourists. Tourists consider that a museum should have more collections to display.

Figure 24 maps in full how the Bigram begins with “*not*”, “*no*”, and “*never*”. It can be seen that the words beginning with “*not*” have more numbers than the other two. For the prefix “*never*”, there are only two linked words, namely “*never forget*” and “*never stop*”.

Figure 24. Bigram Network for negative words in Kota Tua Reviews

Figure 25 shows words that are preceded by a “*not*” which have a major contribution to sentiment, both for positive and negative directions. It can be seen that for the negative direction, the word “*not worth*” or “*not worth it*” gave the biggest contribution. The word “*not worth*” is used by tourists because tourists feel that the effort put into seeing the buildings in Kota Tua is not worth it because many are in poor condition. For the positive direction, the word “*not miss*” gave the largest contribution compared to others. This word is used by tourists in reviews as a form of information so as not to miss certain attractions in Kota Tua.

Figure 25. Words preceded by “Not” in the Kota Tua

Conclusion

Based on the results of the *sentiment analysis* on reviews submitted by tourists for the Kepulauan Seribu and Kota Tua, it can be concluded that the following are:

1. In general, tourists have good sentiments towards Kepulauan Seribu and Kota Tua, and every year the polarity of sentiment moves in an increasing trend.
2. All tourists, both in Kepulauan Seribu and Kota Tua, consider that the cleanliness of these two destinations still needs to be improved, due to the large amount of garbage, the appearance of an odor, and high pollution. This is important to do so that tourists can enjoy the two destinations more, thereby increasing the *length of stay*.
3. Apart from cleanliness issues, tourists also feel that Kepulauan Seribu and Kota Tua can be more tidied up, so that they can give a good impression to tourists, and they want to promote these two destinations. The regional government is also expected to be able to play a more role in carrying out this arrangement.
4. Based on the reviews given by tourists, the problem of standard facilities and service standards is also very important, so that

tourists don't feel disadvantaged because they are given poor facilities or services.

References

- Arndt, J. (1967). Role of Product-Related Conversations in the Diffusion of a New Product. *Journal of Marketing Research*.
- Buttle, FA (1998). Word of Mouth: Understanding and Managing Referral Marketing. *Journal of Strategic Marketing*.
- Carl, WJ (2006). What's All the Buzz About?: Everyday Communication and the Relational Basis of Word-of-Mouth and Buzz Marketing Practices. *Management Communication Quarterly*.
- Chen, M., Mao, S., & Liu, Y. (2014). Big Data: A Survey. *Mobile Networks and Applications*.
- Cheung, MY, Luo, C., Sia, CL, & Chen, H. (2009). Credibility of Electronic Word-of-Mouth: Informational and Normative Determinants of On-Line Consumer Recommendations. *International Journal of Electronic Commerce*.
- DeMauro, A., Greco, M., & Grimaldi, M. (2015). What is Big Data? A Consensual Definition and A Review of Key Research Topics. *AIP Conference Proceedings*.
- Hashem, IA, Anuar, NB, & Mokhtar, S. (2014). The Rise of Big Data on Cloud Computing: Review and Issues. *Information Systems*.
- Litvin, S., Goldsmith, RE, & Pan, B. (2008). Electronic Word-of-Mouth in Hospitality and Tourism Management. *Tourism Management*.
- McAfee, A., & Brynjolfsson, E. (2012, October). Big Data: The Management Revolution. *Harvard Business Review*.
- Puschmann, C., & Burgess, J. (2014). Metaphors of Big Data. *International Journal of Communication*.
- Song, H., & Liu, H. (2017). Predicting Tourist Demand Using Big Data. In Z. Xiang, & D. Fesenmaier, *Analytics in Smart Tourism Design, Tourism on the Verge*. Switzerland: Springer International.
- Valdivia, A., Luzon, MV, & Herrera, F. (2017). Sentiment Analysis in Tripadvisor. *IEEE Intelligent Systems*.
- Westbrook, RA (1987). Product/Consumption-Based Affective Responses and Postpurchase Processes. *Journal of Marketing Research*.