

# Exploring consumer behavior through user-generated content on TripAdvisor. The case of Holguin destination

## Explorando el comportamiento del consumidor a través del contenido generado por el usuario en TripAdvisor. El caso del destino Holguín

Elizabeth del Carmen Pérez Ricardo<sup>1\*</sup>, Ernesto Batista Sánchez<sup>2</sup> and Justa Ramona Medina Labrada<sup>3</sup>

<sup>1</sup> M.Sc. Tourism Management, Department of Tourism, University of Holguin, Holguin, Cuba

<sup>2</sup> M.A. International Tourism, Kemmy Business School, University of Limerick, Limerick, Ireland

<sup>3</sup> M.Sc. Tourism Management, Department of Tourism, University of Holguin, Holguin, Cuba

\* Corresponding author: [eliza941025@gmail.com](mailto:eliza941025@gmail.com)

Received: 2022-11-15

Accepted for publication: 2022-11-30

Published: 2022-12-31

### ABSTRACT

*The present study explores consumer behavior of tourists through user-generated content by presenting a method that utilizes online reviews and text processing techniques in analyzing behavior patterns. The current study examined 7,856 travel reviews from TripAdvisor written by international tourists visiting Holguin (Cuba) destination. This method includes the overall process of converting raw data into useful information, which comprises selection, pre-processing, data mining and interpretation. Data are collected from online review platforms rather than from traditional databases, and the first step is data crawling. Then, unstructured review content is transformed into suitable formats for analysis. The final step is exploratory data analysis, which includes both data mining and interpretation. Findings show that the cognitive dimension of the experience predominates in consumer's behavior. The study identifies topics that could be used by destination management organizations to promote this destination and highlights the advantages of applying a data science approach.*

**Keywords:** Consumer behaviour; User-generated content; Text mining; Sentiment analysis; Tourist destination

### RESUMEN

El presente estudio analiza el comportamiento del consumidor a través del contenido generado por el usuario, al presentar un método que utiliza reseñas en línea y técnicas de procesamiento de texto para analizar patrones de comportamiento. Fueron examinados 7.856 reseñas de viajes de TripAdvisor, escritas por turistas internacionales que visitan el destino Holguín (Cuba). Los datos se recopilan de plataformas de opinión online en lugar de bases de datos tradicionales y este contenido no estructurado se transforma en formatos adecuados para el análisis exploratorio, que incluye tanto el procesamiento como la interpretación de datos. Los hallazgos muestran que la dimensión cognitiva de la experiencia predomina en el comportamiento del consumidor. El estudio identifica temas que las organizaciones de gestión de destinos podrían utilizar para promover este destino y destaca las ventajas de aplicar un enfoque de ciencia de datos.

**Palabras clave:** Comportamiento del consumidor; Contenido generado por el usuario; Minería de texto, Análisis de sentimiento; Destino turístico

## INTRODUCTION

At present, competitive dynamics in tourism industry demands an in-depth analysis about the creation of better relationships with different marketing environments. Customers have been essential for tourism companies. However, the technology and globalization impact on tourists are not completely understood. The information technology advances and the introduction of new communication methods have caused changes in consumer behavior, who is constantly seduced by different products and companies in competitive environments (Öter, 2018).

In this sense, consumer behavior is generally focused on the activities that people do when they search, select, buy, consume and evaluate products, services, ideas or experiences that they hope will satisfy their needs or desires. Most consumer behavior models include three stages: pre-purchase, consumption, and post-purchase. In the first stage, potential tourists recognize the need or feel some motivation to travel on their vacations, so they look for information and evaluate alternatives to plan they trip.

During the second stage, the tourists develop some events and activities they choice in the destination. The last stage includes the evaluation of tourist experience after trip, which is reflected in feelings of satisfaction or disagreement with the products or services consumed. The analysis of these sensations after purchase and tourist product consumption is important, because they can affect the future consumer behavior related to (re) choice or not the destination and in the decision-making of potentials clients with whom the consumer has shared his travel experiences (W. Wang & Wang, 2018).

Cohen, Prayag, and Moital (2014) refer that tourism researchers have developed individual concepts, specific influences and particular research contexts, without considering some of these assessments in the general environment of travel or tourist behavior. Research concepts, influences and contexts can be studied for a specific stage (before the visit, during the visit and after the visit). Consumer evaluations are important for future decision-making and are essential for tourism managers. Understanding consumer behavior is the fundamental propose of marketing management (Kotler & Keller, 2006).

Several studies carried out in recent years address different contexts such as: marketing and hotels (Fetscherin,

2019; Tan, Lee, Hew, Ooi, & Wong, 2018) food and beverage process (Shin, Im, Jung, & Severt, 2018; Tsai & Wang, 2017); tourism and travel (Guo, Barnes, & Jia, 2017; Huang & Lu, 2017) and tourist destinations (Díaz-Meneses, 2019; Jiménez-Barreto, Rubio, & Campo, 2020; Pestana, Parreira, & Moutinho, 2020). These studies about consumer behavior in tourist destinations focus mainly on issues such as brand authenticity, satisfaction, emotion and motivation as key factors for choosing a destination and market segmentation strategies. In addition, these papers refer perspectives such as the impact of the experience quality in service and creation of destination image, the influence of the millennial consumer brand perception value in the destination loyalty, the effect of the cultural intelligence of the tourist in his evaluation of the destination and the new tendencies in the search for information and its influence on the loyalty towards the destinations. However, the study of post-consumer behavior in destinations, based on the integration of its components, is a topic that is rarely developed.

On the other hand, companies, which traditionally learned about customers' needs, preferences, consumption habits, and other details by applying market research such as surveys, are now learning about customers in new ways (Couture, Arcand, Senecal, & Ouellet, 2015). These transformations have produced a change in the company's marketing strategies and business administration (A. Serra Cantallops & Salvi, 2014), especially on the tourism industry, where it is being used, increasingly, for marketing studies and decision-making.

User Generated Content (UGC) is a new source of data generated by potential or actual consumers via the internet. It can be found on the web in different ways such as text (discussion forums, blogs, etc.), photos, videos, music, audios and other ways that gives users the opportunity to create content, interact with each other and share. This information exchange contains knowledge, user experiences, as well as opinions about a product, a service or an experience and it is considered one of the largest sources of big data in this industry (Oum & Han, 2011).

Also, tourists share their travel experiences through important online platforms in the industry such as TripAdvisor (Guo et al., 2017; Vu, Li, Law, & Zhang, 2019). Online travel opinions usually contains opinions of hotels, res-

taurants and attractions (Fazzolari & Petrocchi, 2018) that are fundamental elements that customers consult before planning their trips and booking tourist destinations (Xiang, Schwartz, Gerdes, & Uysal, 2014). These opinions are published by consumers who have bought and used a product or service, including the consumer's own experiences during consumption, as well as evaluations of the product used. They are considered a specific example of e-WOM (electronic word of mouth) and it is important for customers and brands because the influence on the purchase intention of other consumers. This practice is an essential source of information for post-consumer behavior studies today.

Pantano, Priporas, and Stylos (2017) state that "while many studies have investigated the effect of online opinions on tourist's decisions, none have directly investigated the extent to which open data analyzes can favor the tourist response to a certain destination". In addition, (Antonio, Correia, & Perdigão Ribeiro, 2020) refer that a major challenge in tourist destination management is how to track the behaviour of tourists. Destination managers need to know the details of specific locations visited by tourists, what attracts tourists at each location, personal reflections on tourists' experiences and future travel behavioural intentions. In general, most current approaches are unable to address these issues in a decision-centric, integrated and comprehensive manner. Most of the existing methods for analysing social media data are focused on finding answers to specific questions that are predefined in their studies and not on developing a general understanding of tourists' movement, interests and experiences.

The proposed study goes beyond previous studies in the use of a methodology to develop and evaluate a new method of analysis based on big data with significant content focused on tourism. The method integrates techniques to identify patterns of consumer behavior in a tourist destination, taking into account the main components of the destination. This document specifies the details of the proposal that is of value for the strategic planning of destination management and as a tool to support operational decision-making. The objectives are to present a general framework to analyze consumer behavior in a tourist destination through the processing of user-generated content and to demonstrate the effectiveness of the proposed method through a case study of Holguin tourism destination.

In this article, consumer post-consumption behavior is used to refer to tourists' evaluations of a destination's tourist attractions, accommodations and gastronomic activities. Subjective opinions refer to the sentiments (positive, negative) that tourists expressed in the review comments about their experience at the visited destinations. Holguin destination was selected for the case study because it is a popular destination in Cuba. It is the third most important tourist destination in the country and also has one of the largest tourism projections in the Caribbean.

The analysis of the case study results, which are based on approximately 8,000 reviews posted by international tourists about attractions, hotels and restaurants, can provide insights into the general behaviors and perceptions of tourists. The proposed method can help researchers and tourism practitioners gain comprehensive insights into the consumer behavior by utilizing online travel reviews as alternative data sources to traditional surveys and questionnaires approaches.

The rest of the article is organized as follows. The second section provides a review of the literature on the consumer behavior and an overview of consumer behavior studies in the context of tourist destination. Limitations of the existing studies are also highlighted. The third section introduces the method used for extracting and processing online travel reviews. The fourth section presents the results of a case study on international tourists visiting Holguin and an analysis thereof, followed by a discussion of the practical implications of the research outcomes. The article includes the details of study limitations and identification of avenues for further research.

## **METHODOLOGY**

This section presents the method of exploring consumer behavior based on user-generated content. The use of a mixed methods approach from Data Science, especially Text Mining, Data Mining, and Machine learning, also makes it possible for less biased and more consistent analysis, uncovering unknown patterns and trends. While the algorithms employ the same criteria in all analyzed texts, a human analyst can hardly maintain the same standards and objectives over time (Antonio et al., 2020).

This method includes the overall process of converting raw data into useful information, which comprises selection,

pre-processing, data mining and interpretation. In this case, data are collected from online review platforms rather than from traditional databases, and the first step is data crawling. Then, unstructured review content and metadata are transformed into suitable formats for analysis. The final step is exploratory data analysis, which includes both data mining and interpretation. A series of analysis are conducted to provide comprehensive insights into consumer behavior. Figure 1 presents destination review analysis framework; the details are presented in the subsequent sections.

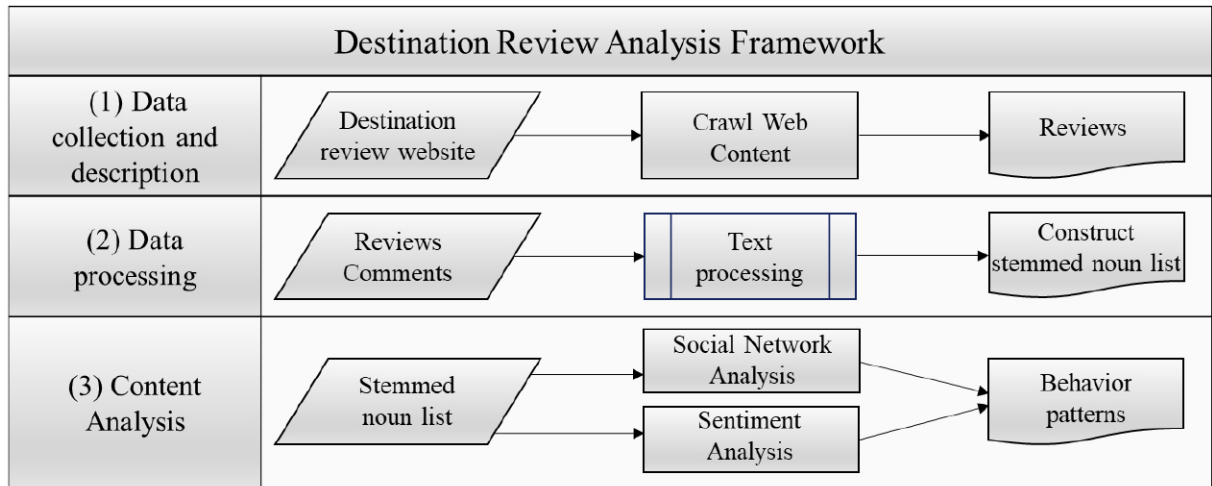
**Data collection and description**

The first step in the framework is to crawl the data from

the platform. Web crawler software is developed to navigate through the review websites and extract their content automatically. In this article was used TripAdvisor (www.tripadvisor.com) as a data source. TripAdvisor is a popular travel review site, which features a comprehensive database of attractions, restaurants and hotels, with millions of reviews by tourists from all over the world. Among the available data on TripAdvisor, review text is the most important for this research. Review text is the actual comment provided by reviewers after visiting destinations. It contains rich information about tourist’s behavior.

**Data processing**

**Figure 1. Destination review analysis framework**



The processing of the data will allow to normalize the entries obtained from the web crawling and facilitate subsequent analysis; however, improper use can lead to the loss of important information. Among the most frequent techniques used in text mining for this type of processing are: tokenization, filtering and stemming.

Text processing is conducted as follows. Each review, is first loaded into a text tokenization algorithm, in which the stream of text is broken into words, phrases, symbols, or other meaningful elements called “tokens.” A filter is applied to the tokens to normalize all letters to lower case and remove symbols and numbers. The remaining tokens are inputted into the stemming module to reduce inflected words to their stem, base, or root form.

The stemmed words then go through part of speech tagging (POST), where each word is tagged with its corresponding type, such as noun, verb, or adjective. In this framework, it was adopted the English lexicon available in GATE (<http://gate.ac.uk/>) for POST. GATE is a widely used package for text processing with a large vocabulary database (Vu et al., 2019). Then, words of noun and adjective type in a stemmed list for further processing were kept and others were discarded. The advantage of this text-processing technique is that the researcher can inspect this list and select the nouns and adjectives describing attractions, restaurants and hotels for further analysis. This technique facilitates the exploration of all possible destinations-related aspects that tourists are interested in or concerned about.

**Content analysis**

Analysis of the social network. In this step, the social network is built from the result of the data processing. It represents the connections between the agents that make up the systems to be studied and they have a great weight in their influence. They are useful in research to study emergent behaviors in previously created systems. In this investigation, the VosViewer-1.6.10 software was used to visualize the network.

Sentiment analysis: Analyzing tourist sentiments expressed in review comments is important for tourism managers to gain insightful understanding about the experience and subjective opinions of tourists toward the attractions, restaurant and hotels services. Platforms, such as TripAdvisor, usually provide a rating function, which reflects the overall sentiment of reviewers. However, detailed sentiment information about various aspects of destinations is unavailable. Tourists often comment on various aspects about their experience in their reviews. Relying on the overall rating of a review or predefined features is insufficient for an insightful understanding about tourist experiences.

Therefore, this research proposes to analyze the sentiment of tourists from online travel reviews at sentence level for detailed insights. First it is necessary to separate the review into sentences using sentence splitters. Review comments are broken into sentences based on a list of abbreviations or hand-coded rules to identify the end of a sentence. Examples of sentence-end indicators are full stop (.), exclamation mark (!), and question mark (?) (Vu et al., 2019). The sentiment is finally estimated in the sentences. For this analysis it was used MiningCloud Add-in for Excell v-3.5.

A sentence has three possible sentiment labels (positive, negative, and neutral). Table 1 lists examples of review sentences and their sentiment labels. S1 and S4 express positive sentiment of reviewers toward restaurant features, such as atmosphere, service, and food items. S2 express negative sentiment toward hotels and S3 is labeled as neutral because they mainly include facts rather than express any subjective opinion.

**Table 1. Example of review sentences and their sentiment labels.**

ID	Comment	Sentiment Label
S1	"It is a fantastic place and my favourite restaurant in Holguin"	Positive
S2	"I visited the resort during Christmas I was disappointed with room safe, air condition unit and patio door broken"	Negative
S3	"The coach taxi leave me in central drop point"	Neutral
S4	"It is a beautiful restaurant outside with music and a great lasagne"	Positive

**RESULTS**

**Data collection and description**

The data were collected from TripAdvisor with the data extraction and web crawler software described in the methodology section. In carrying out this step a specific crawling technique was used, which is applied in studies that involve web mining, because only files on a specific topic were selected, thus reducing network traffic. The data were selected using two ways: web crawling and an API access (Application Programming Interface). This data includes reviews generated by consumers on TripAdvisor. The data collection was focused on attractions, hotels and restaurants located in Holguin tourism destination, one of the most popular tourist destinations in Cuba. The software navigated through the listed attractions, hotels and restaurants in Holguin tourism destination to extract the reviews and associated information about the reviewers. The data extraction software was employed in December 2020, navigating through the web sites and collecting approximately 8,000 review comments.

The type of content that was selected for the study were opinions, of which there are a large number in the destination. Reviews written in English were chosen considering the availability of specialized tools for the analysis of natural language processing techniques. Furthermore, 70% of the total reviews were written in this language. The time frame selected for the data extraction was January 2015 until January 2020.

Once the content segment to be studied had been selected, the sample was determined. In this case, the population is taken for the analysis, as explained in García-García, Reding-Bernal, and López-Alvarenga (2013), since it fulfills the assumption of 100% representativeness and it is possible to carry out its extraction and analysis by using specialized software. It covers all the reviews, written in English, generated by customers about the most representative components of Holguin tourism destination, in the period January 2015 - January 2020.

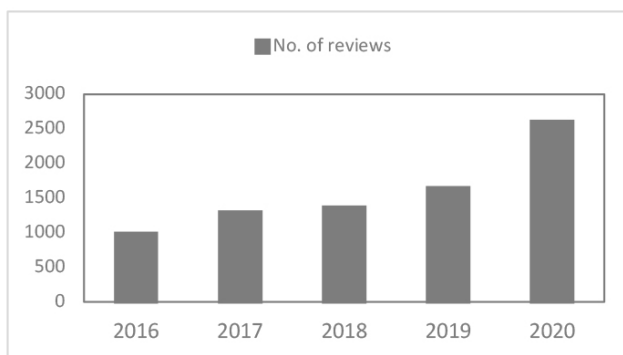
TripAdvisor is an open platform, in which any registered user can post travel reviews. Reviews posted by local residents was removed, given that international tourists are the focus of this analysis. 7,856 reviews were left, posted by international tourists for 15 attractions, 15 restaurants and five hotels, as shown in Table 2.

**Table 1. Example of review sentences and their sentiment labels.**

Components	No. of components	No. of reviews	No. of reviews/ No. of components
Hotels	5	6364	1272.85
Restaurants	15	447	29.82
Attractions	15	1043	69.53

Figure 2 shows the distribution of reviews by year. Most of the reviews were posted since 2018, and the number of reviews has been increasing in recent years. The fast growing number of reviews is probably due to the availability of review websites and the change in tourist behavior in information search and sharing about travel (Vu et al., 2019). Given that the data set is relatively new, the whole data set was considered in this case study.

**Figure 2. Destination review analysis framework**



### Data processing

**Tokenization:** In this step, a pre-processing of the extracted text is carried out, which is developed in three moments, tokenization, filtering and stemming, using the NLTK natural language processing library, also developed in Python, which some of the languages supported by default are: English, Spanish, German or Portuguese. Then punctuation marks, special symbols and characters that do not constitute text were eliminated in the comment, so that tokenization allowed each opinion to be fragmented into a set of phrases and keywords, finally obtaining a list of tokens.

**Filtering:** After the tokenization process, the data was filtered based on the grammatical elements criteria, which made it possible to reduce the definitive list of tokens to nouns and adjectives, as well as to eliminate most of the stop words. Those tokens that were not significant for the future classification of the text (articles, prepositions, numerals or symbols) were removed from the previously generated list.

**Stemming:** Finally, the stemming process is developed. With the application of this technique, it was possible to further reduce the defined tokens, since the terms derived from the same word were associated in a single token, thus avoiding linguistic variations.

From the data processing, a set of relevant information was obtained to carry out different analysis. The definitive token list covers a total of 701,568 keywords, which are distributed by reviews of each component of the destination, 87,347 correspond to attractions, 603,984 to hotels and 10,237 to restaurants. In some cases, it was not possible to eliminate all stop words due to the existence of spelling errors in the reviews.

### Content analysis

#### Analysis of the social network

In this step, the results of the analysis of comments are interpreted. Hence, the social network was used in order to understand the most co-occurring words, their density in relation to the number of comments of each component

and the interactions between them and thus define the main nodes.

From this information it is possible to identify that according to the perception of consumers regarding the cognitive element of the image, beach, hotel and food stand out as attractive. However, when analyzing the social network and the comments in depth, it is perceived that 67% of the comments concerning the beaches do not refer to them properly, but are allusive to the hotels surroundings, generating a wrong image of the attractiveness for those who, when reading the comments, try to get a valuation of the place.

Regarding the affective aspect, the attractions of the destination are considered pleasant places where the quality is good and the staff's assessment varies according to the site. In the case of hotels, among the main terms associated with this component that appear in the 6364 reviews, the following stand out: hotel, room, beach, food and bar. Most of the reviews address elements of the catering service and room cleaning.

Furthermore, beaches and hotels are identified as a relevant factor in the destination image. This information shows that beaches and hotels are one of the main attractions of Holguin tourism destination, given that it is what tourists highlight the most in their reviews. In the case of restaurants, its main cognitive dimensions are related to consumer's perception of food, meal, service, staff and place. Even more, its affective dimension is positively related to the staff and the place.

Overall, six nodes are identified at the destination, as shown in figure 3, which represent the global semantic network. The first cluster, "hotels and restaurants" is the most representative of the destination and its main cognitive elements are: hotel, restaurant, food, drink, staff

and time, which are essential factors that tourists consider during the valuation of the destination. The internal relationship established in this community is the largest in the semantic network. On the other hand, in clusters two, three and four, different variables are interrelated, the most important according to their density are experience and holiday, which are linked to hotel, restaurant, food, among others. In the case of nodes five and six, the least relevant are considered in terms of density and interrelation between variables, in them there are isolated dimensions related to attractions and emotions. It can also be concluded that the cognitive dimension of the experience prevails over the affective dimension according to its presence in the comments studied. Likewise, hotels and restaurants predominate as the most important components.

### Sentiment analysis

For the sentiment analysis, two practices were combined, automatic processing, performed through the Mining Cloud Add-in for Excel v-3.5 software, and human processing to maximize the efficiency and precision of several results. The reviews were classified through the Mining Cloud according to three levels: positive, negative and neutral, and the subjectivity / objectivity of the review and its irony were also considered as another indicator.

As shown in figure 4, positive opinions predominate in the destination (73%), being the number of negative (17%) and neutral (10%) reviews the least significant. However, regarding the Agreement / Disagreement indicator, 44% of the positive comments with internal contradiction in the content stood out, due to the fact that the tourists expressed negative feelings or disagreement with some product or service of the destination in some sentence of the review.





In addition, the proposed approach provides more insights into the consumer behavior and opinions of tourists than prior works using online reviews (Miah et al., 2017) because textual reviews are analyzed and detailed information about the subjective opinions of tourists can be explored by sentiment analysis, which is effective in assessing tourist satisfaction and identifying shortcomings for future improvement.

Fundamental and important insights into tourist's behavior in terms of preferred attractions, hotels and restaurants, and their subjective opinions can be obtained efficiently. The proposed method can be applied to big data set for large-scale studies owing to the automation of data collection and utilization of software tools and APIs for data preprocessing, which allows for convenient subsequent statistical analysis. The effectiveness of the proposed method is demonstrated in a case study on Holguin tourism by using a large-scale data set.

The case study demonstrates the capability of large-sale online reviews in capturing comprehensive information about tourists' preference. Rather than using existing approaches to data collection (e.g., survey and questionnaire), this study employs online review data supported by text processing techniques. The analysis is carried out on a large-scale data set to provide a comprehensive understanding of tourists' preferences in terms hotels, attractions and restaurants. The result of this study has practical for destination management organizations, given the importance of the analysis of post consumption tourism behavior.

Findings from this study can help tourism enterprises to identify the elements affecting online tourism reviews and find correlations between them. This information extracted from the opinions is important to improve the commercial management of the destination and identify competitive advantages of the organizations (Fang, Ye, Kucukusta, & Law, 2016; Miguéis & Nóvoa, 2017). With this approach, companies can better and more accurately understand user needs and improve their quality of service.

Furthermore, text mining is a knowledge discovery process with which managers can design branded advertisements and develop themes that meet consumer demand, thus increasing their willingness to buy and providing a better experience for tourists.

The use of big data tools by tourist managers is considered a key element for destination management because the systematization and analysis of this content allows a reduction in the time, effort and money spent on image formation by of tourism managers. This image of the destination that is formed by means of the content generated by the user has a greater reach and a high speed of propagation, given the visibility and the possibility of sharing information and data should be used as a beginning point for specific policies of tourism managers, who, knowing how the image is perceived, can invest their resources in emphasizing certain segments, improving others and investing in points of the destination with unexploited tourist potentials.

In terms of technical limitations, the text-processing framework is mainly designed for reviews in English. Specific text-processing techniques and language lexicon should be further investigated to process reviews from other languages. Future studies can investigate other tools to better support sentiment analysis of review comments in other languages. It is worth mentioning that review websites, such as TripAdvisor, do provide API for direct access to their data. However, the data collected via TripAdvisor API are subject to its own terms and conditions, which may have restrictions on data analysis and result publication. Researchers and business managers are suggested to get themselves familiar with the terms and conditions before carrying out their research using TripAdvisor API, and be familiar with potential restrictions for using the alternative web crawling approach presented in this study.

The proposed method is a general analysis framework applicable to data from review websites other than TripAdvisor. Future research can consider combining review data from multiple travel platforms to obtain representative

results and analysis. Text processing tools and sentiment analysis API for languages other than English can be integrated into our framework for destination reviews written by tourists from non-English-speaking countries.

Future study can investigate other sophisticated text mining techniques to extract patterns according to specific needs from the processed data resulted from our framework. Apart from text, photos can provide insights into tourists' own experience and specific interests. Future research can focus on developing techniques for analyzing photos together with the review comments for in-depth understanding. In this work, the main focus was on tourists who traveled to a tourism destination.

**REFERENCES**

**Antonio, N., Correia, M. B., & Perdigão Ribeiro, F. (2020).** Exploring User-Generated Content for Improving Destination Knowledge: The Case of TwoWorld Heritage Cities. *Sustainability*, 12, 19. doi:10.3390/su12229654

**Cohen, S. A., Prayag, G., & Moital, M. (2014).** Consumer behaviour in tourism: Concepts, influences and opportunities. *Current Issues in Tourism*, 17(10), 872-909. doi: 10.1080/13683500.2013.850064

**Couture, A., Arcand, M., Senecal, S., & Ouellet, J. F. (2015).** The Influence of Tourism Innovativeness on Online Consumer Behavior. *Journal of Travel Research*, 64(1), 66-79. doi: 10.1177/0047287513513159

**Díaz-Meneses, G. (2019).** A multiphase trip, diversified digital and varied background approach to analyzing and segmenting holidaymakers and their use of social media. *Journal of Destination Marketing & Management*, 11, 166-182. doi: https://doi.org/10.1016/j.jdmm.2017.07.005

**Fang, B., Ye, Q., Kucukusta, D., & Law, R. (2016).** Analysis of the perceived value of online tourism reviews:

Influence of readability and reviewer characteristics. *Tourism Management*, 52, 498-506.

**Fazzolari, M., & Petrocchi, M. (2018).** A study on online travel reviews through intelligent data analysis. *Information Technology & Tourism*, 20(8), 37-58. doi: 10.1007/s40558-018-0121-z

**Fetscherin, M. (2019).** The five types of brand hate: How they affect consumer behavior. *Journal of Business Research*, 101, 116-127. doi: 10.1016/j.jbusres.2019.04.017

**García-García, J. A., Reding-Bernal, A., & López-Alvarenga, J. C. (2013).** Cálculo del tamaño de la muestra en investigación en educación médica. *Investigación en Educación Médica*, 2(8), 217-224.

**Guo, Y., Barnes, S. J., & Jia, Q. (2017).** Mining meaning from online ratings and reviews: Tourist satisfaction analysis using latent dirichlet allocation. *Tourism Management*, 59, 467-483. doi:10.1016/j.tourman.2016.09.009

**Huang, Q. F., & Lu, Y. Q. (2017).** Generational perspective on consumer behavior: China's potential outbound tourist market. *Tourism Management Perspectives*, 24, 7-15. doi: 10.1016/j.tmp.2017.07.008

**Jiménez-Barreto, J., Rubio, N., & Campo, S. (2020).** Destination brand authenticity: What an experiential simulacrum! A multigroup analysis of its antecedents and outcomes through official online platforms. *Tourism Management*, 77, 310-322. doi: https://doi.org/10.1016/j.tourman.2019.104022

**Kotler, P., & Keller, K. L. (2006).** *Dirección de Marketing* (9na ed.). Mexico: Editorial Prentice-Hall.

**Miah, S. J., Vu, H. Q., Gammack, J., & McGrath, M. (2017).** A Big Data Analytics Method of tourist behavior analysis. *Information and management*, 54, 771-785.

- Miguéis, V., & Nóvoa, H. (2017).** Exploring online travel reviews using data analytics: An exploratory study. *Service Science*, 9(4), 315-323.
- Öter, Z. (2018).** Hospitality marketing and consumer behavior. *Journal of Hospitality and Tourism Management*, 36, 111-112. doi: 10.1016/j.jhtm.2017.10.002
- Oum, S., & Han, D. (2011).** An empirical study of the determinants of the intention to participate in user-created contents (UCC) services. *Expert Systems with Applications*, 38(12), 15110-15121. doi: 10.1016/j.eswa.2011.05.098
- Pantano, E., Priporas, C. V., & Stylos, N. (2017).** You will like it!' using open data to predict tourists' response to a tourist attraction. *Tourism Management*, 60, 430-438.
- Pestana, M. H., Parreira, A., & Moutinho, L. (2020).** Motivations, emotions and satisfaction: The keys to a tourism destination choice. *Journal of Destination Marketing & Management*, 16, 125-134. doi: <https://doi.org/10.1016/j.jdmm.2018.12.006>
- Serra Cantallops, A., & Salvi, F. (2014).** New consumer behavior: A review of research on eWOM and hotels. *International Journal of Hospitality Management*, 36, 41-51. doi: 10.1016/j.ijhm.2013.08.007
- Shin, Y. H., Im, J., Jung, S. E., & Severt, K. (2018).** The theory of planned behavior and the norm activation model approach to consumer behavior regarding organic menus. *International Journal of Hospitality Management*, 69, 21-29.
- Tan, G. W., Lee, V. H., Hew, J. J., Ooi, K.-B., & Wong, L.-W. (2018).** The interactive mobile social media advertising: An imminent approach to advertise tourism products and services? *Telematics and Informatics*, 35(8), 2270-2288. . doi: <https://doi.org/10.1016/j.tele.2018.09.005>
- Tsai, C., & Wang, Y. (2017).** Experiential value in branding food tourism. *Journal of Destination Marketing & Management*, 6(1), 56-65. doi: <https://doi.org/10.1016/j.jdmm.2016.02.003>
- Vu, H. Q., Li, G., Law, R., & Zhang, Y. (2019).** Exploring Tourist Dining Preferences Based on Restaurant Reviews. *Journal of Travel Research*, 58(1), 149-167. doi: 10.1177/0047287517744672
- Wang, W., & Wang, L. (2018).** Merger case analysis of tongcheng and yilong, two online travel agents. *Business & Economy*, 13(5), 93-113.
- Xiang, Z., Schwartz, Z., Gerdes, J., & Uysal, M. (2014).** What can big data and text analytics tell us about hotel guest experience and satisfaction? . *International Journal of Hospitality Management*, 44(7), 120-130. doi: doi:10.1016/j.ijhm.2014.10.013