

RESEARCH

Open Access

Exploring halal tourism tweets on social media



Ali Feizollah¹, Mohamed M. Mostafa², Ainin Sulaiman^{1*} , Zalina Zakaria^{1,3} and Ahmad Firdaus⁴

*Correspondence:

ainins@um.edu.my

¹ UM Halal Research Center,
University of Malaya,
50603 Kuala Lumpur,
Malaysia

Full list of author information
is available at the end of the
article

Abstract

This study explores tweets from Oct 2008 to Oct 2018 related to halal tourism. The tweets were extracted from twitter and underwent various cleaning processes. A total of 33,880 tweets were used for analysis. Analysis intended to (1) identify the topics users tweet about regarding halal tourism, and (2) analyze the emotion-based sentiment of the tweets. To identify and analyze the topics, the study used a word list, concordance graphs, semantic network analysis, and topic-modeling approaches. The NRC emotion lexicon was used to examine the sentiment of the tweets. The analysis illustrated that the word “halal” occurred in the highest number of tweets and was primarily associated with the words “food” and “hotel”. It was also observed that non-Muslim countries such as Japan and Thailand appear to be popular as halal tourist destinations. Sentiment analysis found that there were more positive than negative sentiments among the tweets. The findings have shown that halal tourism is a global market and not only restricted to Muslim countries. Thus, industry players should take the opportunity to use social media to their advantage to promote their halal tourism packages as it is an effective method of communication in this decade.

Keywords: Halal, Halal tourism, Sentiment analysis, Twitter, Topic modeling

Introduction

Halal tourism is considered a sub-field of religious tourism. This type of tourism is based on Islamic Sharia Law, which guides all aspects of a Muslim’s life from birth to death. In general, halal refers to anything that is allowed according to Sharia law and includes matters as diverse as food [1], banking [2], cosmetics [3], pharmaceutical products and vaccines [4], and tourism [5]. Certain jurisdictions have legally defined halal; for example, in the Malaysian Trade Description (Definition of halal) Order 2011, halal is defined as when food or goods have followed the requirements imposed by the Islamic law [6].

Most people associate halal with products, but not many are aware that halal is associated with services like tourism as well. Although there have been some attempts to define halal tourism [7], it should be noted that there is currently no agreed-upon definition, likely because the concept is a multidisciplinary one. For example, Ahmed and Akbaba [8] described halal tourism as any tourism activity permitted in Islam to attract both Muslims and non-Muslims. Jafari and Scott [9] defined it as “the encouragement of tourists likely to meet the requirements of Sharia law”, while Carboni et al. [10] referred

to halal tourism as “tourism in accordance with Islam, involving people of the Muslim faith who are interested in keeping with their personal religious habits whilst travelling”.

Due to demographic factors and rate of spending among Muslims and non-Muslims, the market for halal tourism seems to have strong future potential. While the world population growth rate was 0.7 percent in non-Muslim majority countries, it was 1.5 percent in Muslim-majority countries, and the Muslim population has doubled over the last few decades. Spending by Muslim tourists has also increased exponentially, reaching \$220 billion in 2020 compared to \$156 billion in 2016 [11].

Although halal tourism is considered a niche market, it covers many administrative areas. Halal tourism requires halal food, halal entertainment, gender segregation, and Islamic financial institutions, airlines, and tour packages [12]. The presence of these facilities is essential for Muslims [13]. Many Muslim and non-Muslim countries are competing to become the top destination in this lucrative market despite halal tourism requiring extensive infrastructure and services. According to a recent study by Yagmur et al. [11], 130 countries became halal tourism destinations in 2017, of which only 46 had a Muslim-majority population.

During travel, tourists take pictures and share their thoughts on social media to update friends and family and to share their experiences with business owners, who traditionally ask customers to fill out customer experience surveys. As such, social media has become a rich source of data that can be harvested for research.

The halal tourism industry requires considerable research to reach its true potential and offer the best services to tourists. Thus far, researchers have primarily used qualitative methods to study this industry. Salleh et al. [14] surveyed hotel owners about halal tourism and the challenges they faced with it. In another approach, Muslim tourists were surveyed and asked about their preferences on their trip [15]. The authors then prioritized the essential criteria of halal tourism according to these tourists’ perspectives [15]. Other researchers have studied the status of halal tourism in their own countries, such as Egypt and Indonesia [16–18].

A review of the literature illustrated that most studies on Halal tourism were carried out using conventional approaches such as survey, focus group discussions and interviews. To carry out these approaches, it requires adequate funding and at times, it is a time-consuming task [19]. Moreover, social media allows researchers to collect a bigger volume of data and the speed is faster [19]. In addition, we can get access to the users’ actual perception, whereby they are not afraid to express their comments compared to the conventional approaches [20]. Hence, for this research, we collected data from Twitter and analyzed the topics discussed as well as the sentiment of the tweets to gather detailed information on the opinions of tourists. By examining the sentiment of travelers on social media, this study will add to the existing knowledge on halal tourism. Moreover, using the big data approach, allows us to assess Twitter’s user’s sentiment in relation to halal tourism faster thus the halal tourism industry players can react to the sentiments more efficiently and effectively. In addition, this study uses the big data approach hence adding to existing methodology in the halal tourism area.

The objectives of this paper are:

- (1) To identify the topics users tweet about regarding halal tourism
- (2) To analyze the emotion-based sentiment of the tweets.

This paper is structured as follows: "Related work" section provides an overview of current related research and establishes a context for this paper. "Methodology" section explains the present study's methodology and details each step of the data analysis. "Findings" section then explores the results in detail using graphs. "Discussion" section provides a discussion on the analysis while the conclusion is presented in "Conclusion" section.

Related work

Although halal tourism represents a sizeable segment of the international tourism market [21, 22], research on this topic remains in its early stages [23]. For example, Salleh et al. [14] conducted qualitative research on the implementation of Muslim-friendly hotels in Malaysia. They interviewed hotel managers about their services, food, rooms, entertainment, facilities, and staff. Hotel managers responded by detailing the effects of converting to Muslim-friendly hotels on their brands and capacity management, the process of halal certification, and the difficulties faced by international chain hotels.

Saad et al. [16] studied Sharia-compliant hotels to identify the challenges of their development in Egypt. They refined group opinions and sent an email questionnaire to experts over two rounds. They found that the most significant features of these hotels were gender segregation and the absence of alcohol and pork on-site. Other essential features were having the Quran, prayer mats, and a sign indicating the direction of Mecca in each room. The challenges facing the hotels, as identified in this survey, were competition from non-halal hotels and the lack of specific criteria for international chain hotels to follow.

Ristawati et al. [18] followed a quantitative approach to determine the effect of tourists' experience and the implementation of innovative measures on the island of Lombok, Indonesia. The researchers surveyed 140 tourists and received 126 responses. Using structural equation modeling (SEM), a statistical analysis tool to solve simultaneous multilevel models, they showed the path coefficient and significance level for the collected surveys. Their analysis revealed that customer experience has a high impact on destination choice, followed by the destinations' innovative value. They concluded that variable customer experience and innovative value significantly influence tourists' satisfaction and the image of a halal destination.

Battour et al. [15] explored the impact of a destination's attributes on Muslims' choices of travel destinations. They found that the essential religious attributes required to meet Muslims' needs in a hotel were places of worship, availability of halal food, banned alcohol, gender segregation, and dress code. The results showed that female respondents were more concerned about privacy and segregation in entertainment facilities, while male respondents were more concerned about worship facilities.

Fakir and Erraoui [24] explored the perceptions of halal tourism of tourists in Morocco. The authors conducted questionnaires in two phases. In phase one, they asked tourists in the city of Agadir about their level of satisfaction, the reason for their trip, and halal offers by hotels. In the second phase, hotel managers, also in Agadir, were asked about services they offered, proposed tariffs, and their communication methods for promoting their halal offers. The authors ultimately found that only 3% of the respondents did not know about halal services offered by hotels in Agadir.

Aini and Khudzaeva [25] took a more technical approach and used K-means clustering to identify potential halal tourism destinations. This clustering algorithm groups data based on closeness, or similarity of attributes, in each iteration. The authors collected data in five provinces of Indonesia—DKI Jakarta, West Sumatera, Great Malang, Riau islands, and Aceh—and the data were clustered via the algorithm based on family-friendliness, safety, number of Muslim visitors, dining options, accommodations, and visa requirements. New data based on the criteria were then fed to the algorithm and the resulting clusters were used to suggest potential halal tourism destinations.

Shakona et al. [26] studied the behavior of Muslim travelers in South Carolina and found that their Islamic beliefs influenced their behavior. Muslim travelers were specifically affected by proximity to a mosque, hijab and dress code requirements, the avoidance of alcohol and pork, and the avoidance of travel during the month of fasting. The authors believe that their study can provide guidelines for developers and those in the tourism industry to prepare for Muslim travelers more consciously.

Using a qualitative approach, Jia and Chaozhi [27] explored the gap between halal tourists' needs and tourism practitioners' responses to such needs in China. Results based on thematic analysis revealed that Muslim tourists' needs revolve mainly around halal food, hotel amenities, entertainment facilities, water-friendly toilets, prayer facilities, and service staff. Results also revealed that while tourism practitioners in China generally showed positive attitudes towards halal tourism, they only provide halal food for Muslim tourists. Using a similar methodology, Manaa [28] investigated the influence of halal food availability and travel experience satisfaction on halal tourists' intentions to revisit tourist destinations among UAE respondents. Results showed that halal food availability strongly influences Muslim travelers' intentions to revisit the destination. Availability of halal food was also found to affect both the type of accommodation chosen and the length of stay.

Suhantanto et al. [29] conducted a survey and employed exploratory factor analysis to investigate the influence of halal experience, service quality, perceived value, and satisfaction on tourists' loyalty. Results revealed that the major dimensions of the halal experience included halal amenities, accommodation, and staff. Partial least squares modeling also confirmed the importance of perceived value, perceived experience, and satisfaction on Muslim tourists' loyalty. Similarly, Papastathopoulos et al. [23] examined the relationship between halal tourists' intentions and willingness to pay a premium as a function of halal physical and nonphysical hotel attributes. Using partial least squares modeling techniques, the authors found evidence to group halal tourists into three major categories—"utilitarian Muslim guests", "independent Muslim guests" and "leisure Muslim guests."

From this literature review, we find that previous research on halal tourism has generally adopted a small-scale conceptual or qualitative approach based on few in-depth interviews conducted with travelers and tourism practitioners [22, 27, 30–32], or considered a limited number of stakeholders in quantitative studies by gathering survey data only on Muslim travelers [33–35]. Although social media has made it easier to choose halal-friendly destinations [21, 22], no previous studies have used big data from social media platforms to investigate sentiments and attitudes towards halal tourism. In this study, we aim to fill this gap by using text mining and topic modeling approaches to examine topics and sentiments expressed towards halal tourism.

A review of the literature illustrated that big data has been used by others in different contexts. For example, Prameswari et al. [36] analyzed online hotel reviews to examine the trends of Indonesia’s priority tourist destinations, while Nguyen et al. [37] used Twitter data from users in the USA to investigate the relationship between preterm birth and low birth weight. Philander and Zhong [38] also used Twitter data related to Las Vegas resorts to explore customers’ attitudes and perceptions. Twitter data was also used by Oliveira, Cortez, Areal [39] to predict financial matters such as stock market returns and volatility. This study fills the gap as it explores big data concerning halal tourism which has not been fully explored by previous studies.

Methodology

This section describes the methodology used in this study, with the following subsections outlining the details.

As shown in Fig. 1, depiction of our methodology’s architecture, the first stage of the study was data acquisition, which involved collecting tweets from Twitter using certain keywords. Since there is limited access to Twitter via APIs, we developed a Python script to collect tweets without limitation. Please refer to "Data acquisition and cleaning" section for more detail.

The third stage of the methodology was data analysis, which involved creating a word list ("Word list" section), a concordance graph ("Concordance graph" section), and a semantic network diagram ("Semantic network analysis" section) to better understand the data. These analyses uncovered repetitive words, distinct patterns in the data, and relationships between patterns.

The last stage of the study was emotion-based sentiment analysis. In this step, we calculated the emotion of each tweet and visualized the overall results.

Data acquisition and cleaning

Twitter was chosen as the primary data source because tweets are shorter by design than posts on other social media platforms, facilitating much quicker analysis. Furthermore, it is one of the most-used social media platforms worldwide [40]. Twitter provides an

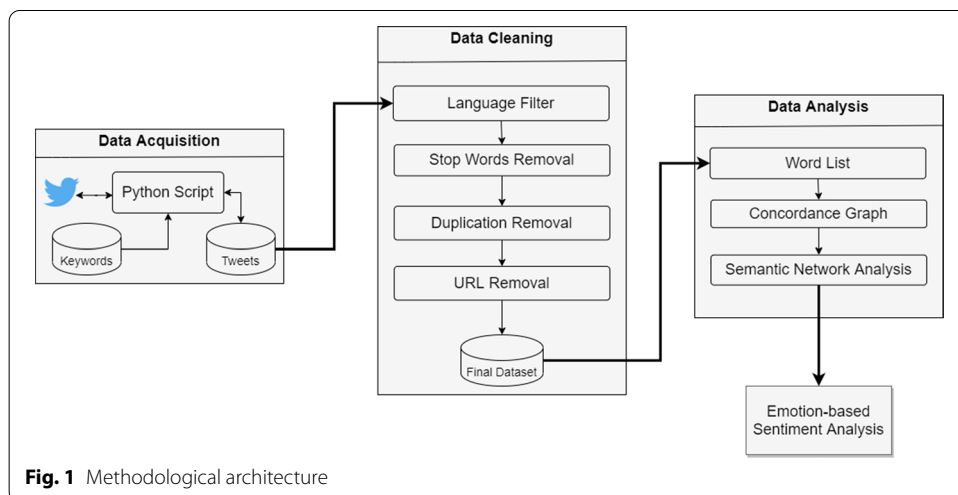


Fig. 1 Methodological architecture

Application Programming Interface (API) to retrieve tweets; however, the API is limited to retrieving tweets from the past seven days and the desired depth of our research required a wider temporal range. We thereby developed a Python script to achieve our goal of analyzing tweets from October 2008 to October 2018. As we started the extraction in November 2018, we decided the cut-off point to be October 2018 and calculating backwards we decided to extract tweets from October 2008 to enable us to get ten years data.

The script uses Scrapy, a Python package, to perform automated Twitter searches according to keywords listed in Table 1, parse the results of each search, and deposit each tweet's text and date posted into a JavaScript Object Notation (JSON) file, which we used as a database. The keywords were finalized after a brain storming session with several Halal experts.

Our data collection, using the listed keywords for the stated dates, resulted in 85,259 tweets, providing us with enough data to produce comprehensive results.

Data cleaning focused on English language tweets as the majority of the gathered tweets were in English. It is worth noting that approximately 50%–72.5% of the language used on social media, including Twitter, is in English [41]. Our searches were intentionally limited to English language tweets; if we had used languages other than English, it would have introduced a linguistic bias in the analysis as the algorithms are tuned for the English language. At this stage, the dataset contained 62,918 tweets.

Following Geetha, Singha, and Sinha [42]'s example, the remaining data cleaning involved removing punctuation, numbers, stop-words, white spaces, special characters, URLs, and emoticons; converting all letters to lower case; and stemming. Similarly, we followed the methodology of Jiao et al. [43] and deleted identical tweets since such posts do not contribute new information and are a form of duplication. The final sample was comprised of 33,880 tweets.

Several packages within the R version 3.6 environment [44] were used to conduct the analysis, including tidytext, plyr, stringr, and ggplot2. The igraph package version 1.2.4 [45] was used for semantic network visualization. This software was selected because it is capable of handling large-scale semantic network graphs.

Topic identification

Word list

The word list generated from the tweets allowed us to identify the main topics mentioned by users. Analyzing the frequency of words in a corpus—a body of text; in this case, the gathered tweets—or comparing the number of repetitions in a word list helps uncover patterns and provide insight into a particular topic. This method is useful for finding the most repeated words in a corpus and perform further analysis. O'Leary [46] explained that even though this approach may appear trivial, it can be used to identify characteristics of the analyzed topic. Similarly, Barlow [47] argued that generating word lists represents “the most radical transformation of a text used in linguistic analysis.”

Table 1 List of keywords used in Twitter searches

Halal travel	Halal tourism	Halal hotel
Halal accommodation	Halal trip	Muslim friendly tourism
Muslim friendly travel	Muslim friendly hotel	Muslim friendly accommodation
Muslim friendly trip	Sharia compliance accommodation	Sharia compliance hotel
Sharia compliance tourism	Sharia compliance travel	Sharia compliance trip

Concordance graph

A concordance graph lists each occurrence of a word or pattern in a corpus and accurately and explicitly shows the different language patterns in context. The primary purpose of the concordance graph is to “place each word back in its original context so that the details of its use and behavior can be properly examined” [48].

Semantic network analysis

Semantic network analysis, a branch of network and graph theory, is used to identify how words are linked to each other in a corpus [49]. In the graph generated from semantic network analysis, each node represents a word, and the edges represent dyadic ties between them, showing the frequency with which a set of words appear in the corpus [50].

Topic modeling

The manual examination and annotation of tweets is ideal, though not feasible in a dataset comprised of thousands of tweets [51]. We instead relied on automation. Topic modeling groups a document into latent topics automatically. A topic is represented by a set of words that can be interpreted as a topic theme [52].

First, in order to get familiar with the data and show them in an understandable form, we apply T-SNE algorithm to the data.

Principal Component Analysis (PCA) [53] and T-distributed Stochastic Neighbor Embedding (T-SNE) are two well-known statistical algorithms used for dimensionality reduction. The T-SNE was introduced by van der Maaten and Hinton in 2008 [54]. While PCA is a linear algorithm, T-SNE is a non-linear algorithm, which makes it a better candidate for dimensionality reduction purpose.

Visualizing data in a 2-D plane or a 3-D plane is a way to see how data is shaped and distributed. However, the collected data has more than three dimensions. The dimensionality reduction algorithms take a data point in high-dimensional space and find a corresponding point in a low-dimensional space.

Emotion-based sentiment analysis

The previous section analyzed the collected tweets in terms of frequency, consistency, relationship, and topics. These analyses familiarized us with the data so we could ensure that the data was relevant to the theme of halal tourism. This section presents an emotion-based sentiment analysis of the collected data. Though the methods used in this work are not novel, the results are unique.

Sentiment analysis

Two primary methods are available for conducting sentiment analysis: the lexicon-based and the corpus-based methods [55]. However, the corpus-based method is rarely used in analyzing sentiment orientation because of its high computational needs. Both methods are based on computing sentiment scores by comparing phrases to an expert-defined dictionary record; however, they differ in their implementation. The lexicon-based approach is based on a pre-defined expert dictionary, while the corpus-based approach is based on a corpus of subjective words.

Several lexicons are available to conduct sentiment analysis, including the General Inquirer [56], the SentiWordNet [57], the LIWC dictionary [58], the Q-WordNet [59], the lexicon of Subjectivity Clues [60], and the Sentiment-based Lexicon [61]. In this study, we use the NRC lexicon [62] because it has been implemented successfully in similar research [1]. In addition to classifying data into positive and negative sentiments, the NRC lexicon, which was assembled manually via crowdsourcing, is capable of classifying emotions into eight basic categories: trust, surprise, fear, sadness, anger, joy, anticipation, and disgust.

Sentiment analysis is the process of identifying and extracting “subjective information from large volumes of unstructured data by combining data mining techniques, machine learning, natural language processing, information retrieval, and knowledge management” [63]. Wu et al. [64] posited that human sentiments could be categorized into specific primary emotions; therefore, we classified the data on halal tourism into the eight emotion classes from the NRC lexicon, listed above.

Koylu et al. [65] noted that sentiment analysis is part of the computational linguistics domain that identifies opinions, emotions, or moods expressed in the text. Oliveira et al. [63] argued that sentiment analysis aims to extract “subjective information from large volumes of unstructured data by combining data mining techniques, machine learning, natural language processing, information retrieval, and knowledge management.” Sentiment analysis classifies texts as positive, negative, or neutral. It is also represented by the range between 1 (highly positive) and -1 (highly negative).

Sentiment analysis [66] is conducted at the sentence [67] or document level [68]. This study, however, analyzed tweets at the word level to identify positive and negative tweets. To find each word’s sentiment, we used the NRC lexicon developed by Mohammad and Turney [62].

Findings

This section describes the findings of the study based on the methodology described in “[Methodology](#)” section.

Topics tweeted

This section presents results related to the topics identified in tweets in various analysis results such as word list, concordance graph, etc.

Word list

Figure 2 shows the most frequent words in the dataset, with “halal”, “travel”, and “tourism” topping the list.

The word “halal” is used the most frequently compared to other words, followed by “travel” and “tourism”. The word “food” is 9th most frequent. This shows that halal tourism includes a variety of services including halal food and that users are interested in halal food as it appeared in many tweets.

The word list also includes “Indonesia” and “japan”. The frequency of reference to Indonesia was somewhat expected since it is the most populous Muslim country. That of Japan, however, shows the commitment of this non-Muslim country to promoting halal tourism. Due to the heavy promotion of halal tourism in Japan, many Muslims are interested in visiting this country, hence this country made it to the word list.

Concordance graph

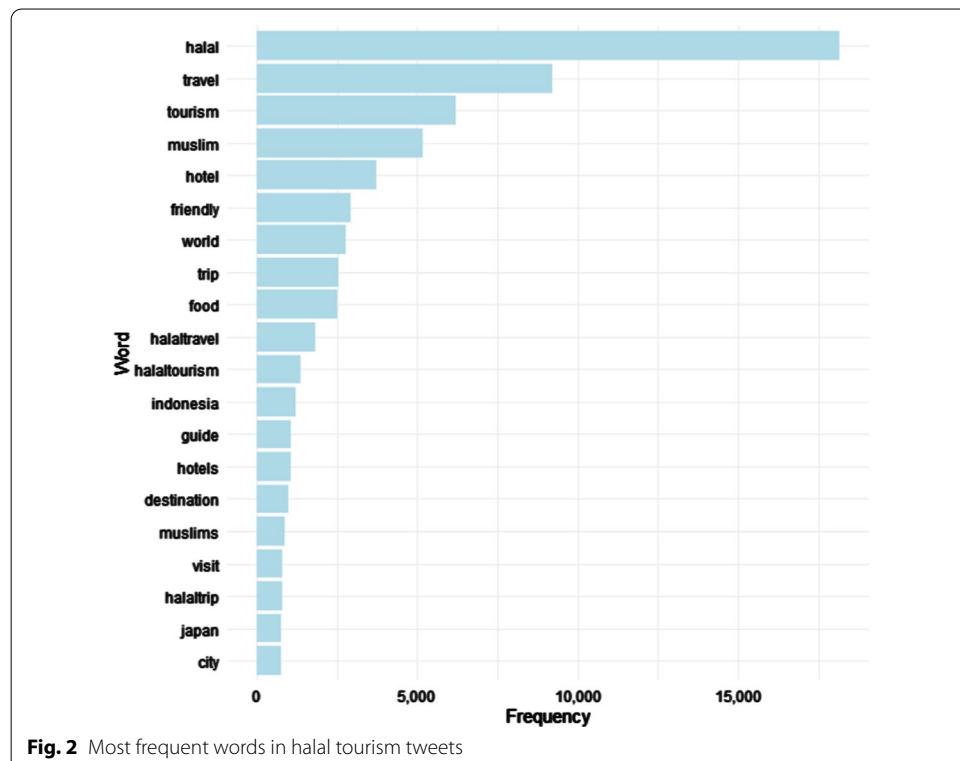
Figure 3 shows the concordance graph for the word “halal” in halal tourism tweets.

Figure 3 shows a concordance graph for the keyword “halal”, generated from the corpus. It displays the keyword, also known as a “node”, centrally, with the context—the rest of each tweet—displayed to the right and left of the keyword. It is worth noting that the graph is read vertically and not horizontally, but patterns can be seen on either side of the keyword as well. Based on Fig. 3, the keyword “halal” is frequently accompanied by the word “accommodation”, as well as words for and about food. Figure 3 also shows that the word “halal” is used consistently and not coincidentally.

Semantic network analysis

The graph generated from semantic network analysis is illustrated in Fig. 4. As mentioned by [50] each node represents a word, and the edges represent dyadic ties between them, showing the frequency with which a set of words appear in the corpus. Since our dataset is a collection of tweets written by different individuals, the semantic network shows a collective cognitive structure among these users [69].

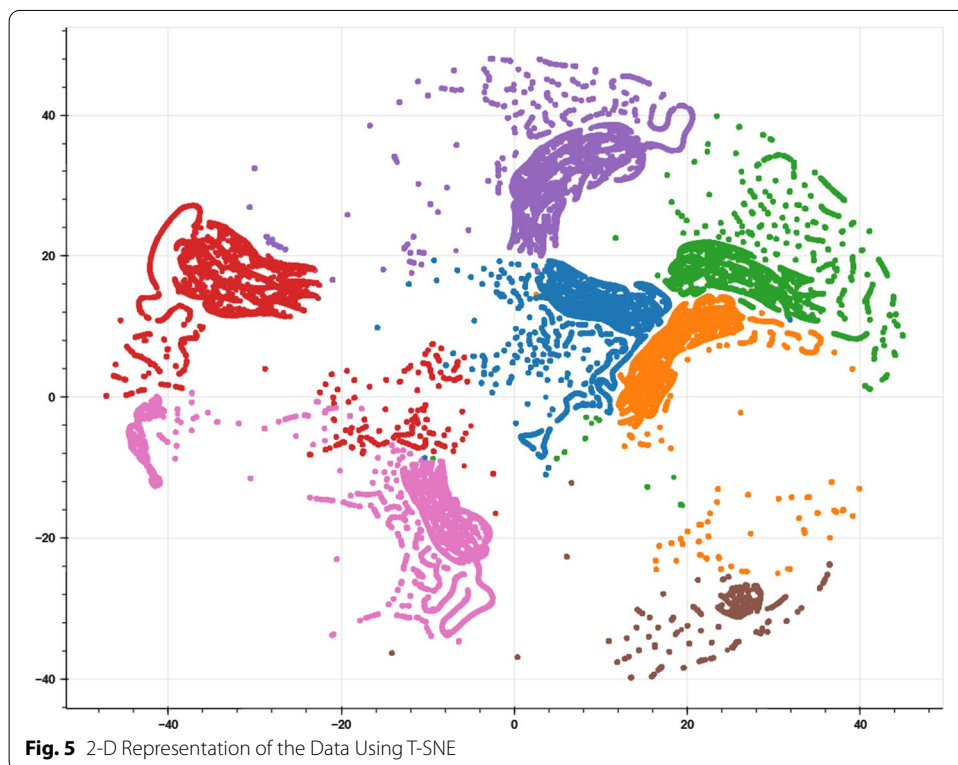
From Fig. 4, we can see that there is a strong link between some word pairs such as “halal” and “meals” (7909 occurrences), “free” and “halal” (7905 occurrences), “halal” and “accommodation” (7852 occurrences), “meals” and “accommodation” (7841 occurrences), and “accommodation” and “tv” (7840 occurrences). The words often accompanying the word “halal” are visible using the semantic network graph. For example, there are strong links between “halal” and “food” and “accommodation”.



data. Then, we tested the algorithm to group the data into 5 topics, which resulted in a vague group of topics. Therefore, trying for 7 topics showed a more promising result. Figure 5 shows the 2-D representation of topics in the tweets.

Figure 5 shows different topics in the data in different colors. The size of each topic represents how large a topic is. There are some points in Fig. 5 that shows color mixing, which represents some common words between topics. To know more about each topic, Fig. 6 shows the word count and weight of each word in each topic. The word count is the number of repetitions for each word. The weight is calculated by T-SNE algorithm, based on which the importance of each word is shown.

The colors in Fig. 6 correspond to colors in Fig. 5, so that we can see details of each topics including the words and word counts. The first topic represents the services offered in the cities. We can see that Dubai is specifically mentioned in Fig. 6. Additionally, the services in the hajj pilgrimage are also a concern for travelers. The second topic shows travelers’ concern in non-Muslim countries like Thailand. The mention of “grow” and “Islamic” make travelers’ wonder about halal friendliness of services in those countries. The third topic is related to the food and accommodation services for halal tourism. It is interesting that travelers mentioned “CrescentRating”, which is the world’s leading authority on halal-friendly travel based in Singapore. This shows that travelers care about their halal destination and look for related information. Topic four mentions hotel as the most repeated words. In this topic, travelers look for the annual world halal excellence awards to see the winners and select their hotel from the top winners in the awards event. The mention of “Japan” and “food” in the fifth topic show the travelers’ enthusiasm to visit Japan and try their food. They look for



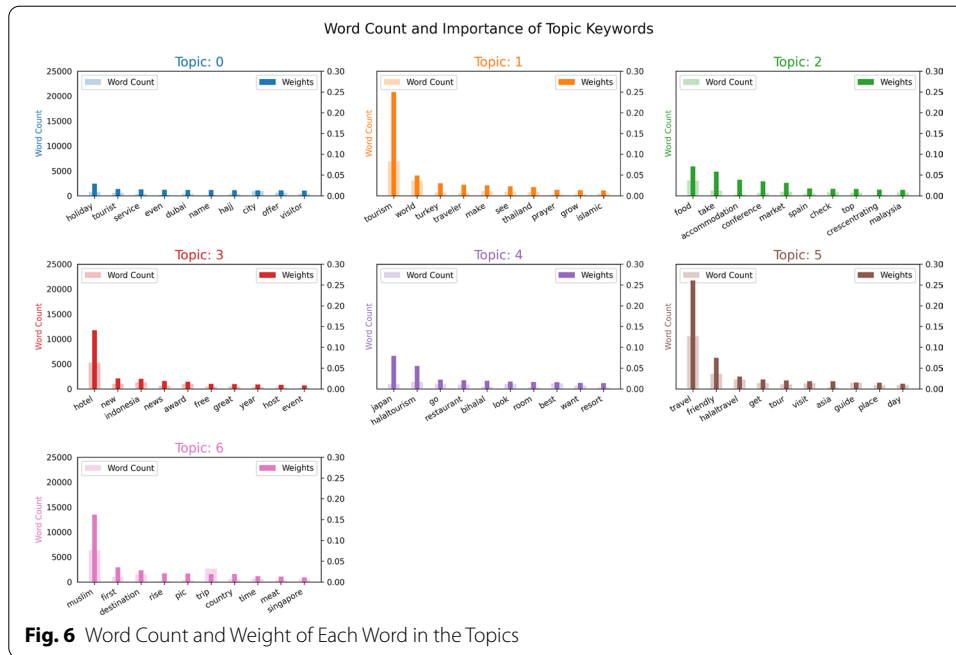


Fig. 6 Word Count and Weight of Each Word in the Topics

restaurants that offer halal food. The sixth topic is about halal tour services in the destinations. Specifically, travelers are interested in the halal tours in Asia to visit points of interests. In the final topic, travelers look to choose their first destination to start their halal travel. They look for pictures and rank of various destinations, among which “Singapore” were mentioned more than other destinations.

We were also interested to see which countries were mentioned more than others in our data. Using the geograpy library in Python, we went through over big data of more than 400,000 data points to identify and rank the countries mentioned in the tweets. Table 2 shows the top 10 countries mentioned in the tweets.

Table 2 shows Indonesia as the top country followed by Japan. The presence of Japan in the second rank shows the interest of travelers in this non-Muslim country. It is worth mentioning that Japan has made an investment over the years to attract Muslim travelers [70].

Emotion-based sentiment analysis

Figure 7 shows the most repeated positive and negative words in tweets. The words “halal” and “tourism” were removed due to their excessive repetition.

The NRC lexicon enabled us to classify the tone or emotion of the text into eight categories to gain more insight into the emotion of the users, as opposed to just positive or negative sentiment. The tweets were classified into fear, disgust, anger, and sadness—the negative emotions—and trust, joy, surprise, and anticipation—the positive emotions. Figure 8 shows the percentage of the corpus that is classified as each emotion.

Trust, joy, and anticipation were the top three emotions found in the tweets, which are considered positive emotions towards halal tourism. The negative emotions in Fig. 8 are disgust, anger, fear, and sadness.

Table 2 The top 10 most repeated countries

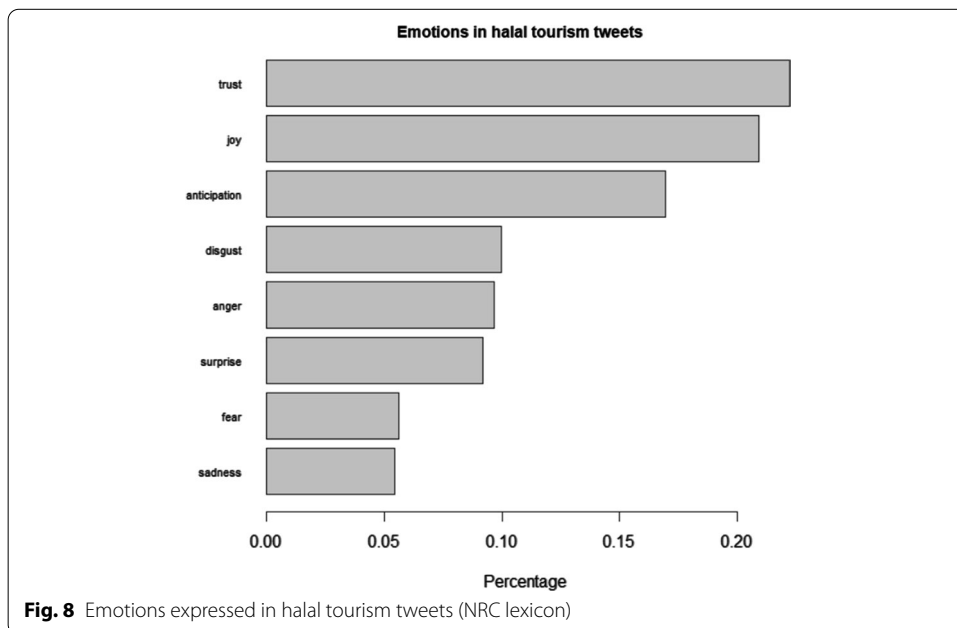
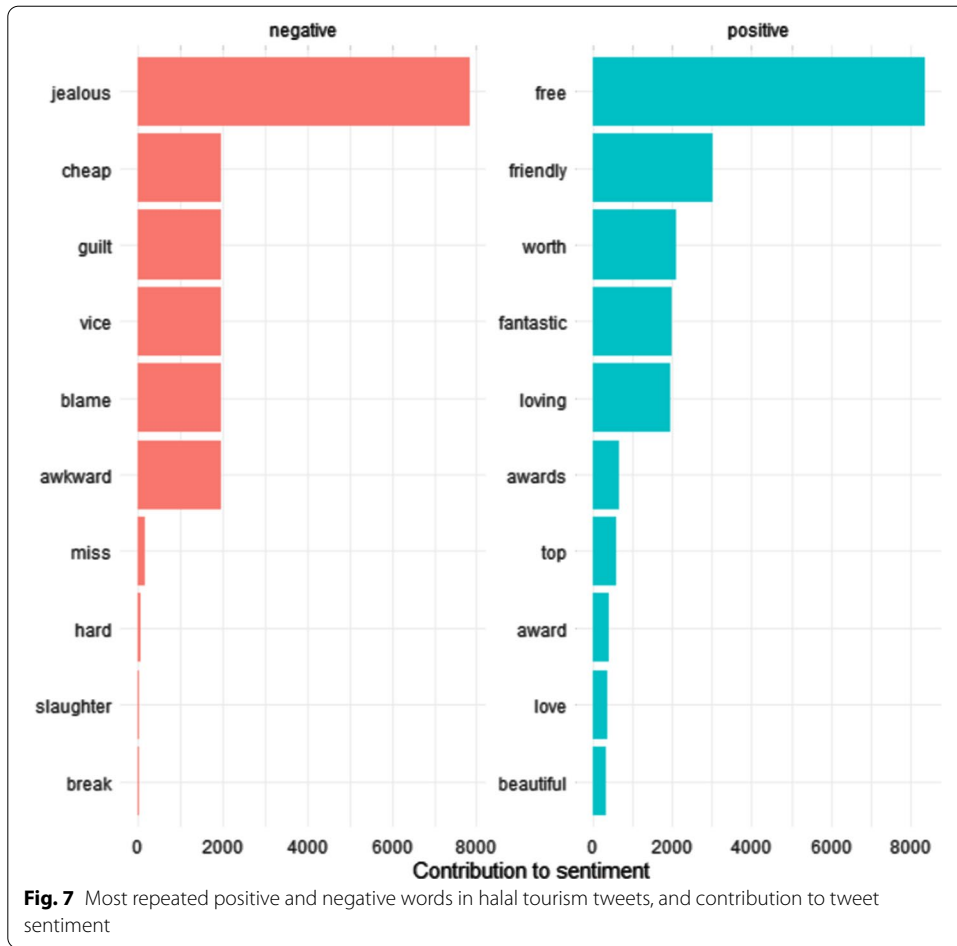
Rank	Country	Count
1	Indonesia	2440
2	Japan	1915
3	Thailand	1135
4	Spain	1065
5	Turkey	1028
6	Malaysia	734
7	Singapore	527
8	United Arab Emirates	459
9	Morocco	379
10	Brazil	355

Discussion

This section elaborates on the results presented in "Findings" section, which highlight several new findings. First, within halal tourism tweets, the word "halal" is the most frequently tweeted followed by "tourism" and "travel". What is interesting to note is that Japan, a non-Muslim country, was tweeted about more frequently than most Muslim countries, apart from Indonesia (from Fig. 2). This could be explained by the Japanese government aggressively promoting halal tourism [70]. This finding is consistent with Yagmur et al. [11] who stated that many countries, regardless of being Muslim or non-Muslim, are competing in the halal tourism market as it is seen to be lucrative. This is further substantiated by the finding, illustrated in Table 2, that several Muslim countries such as Indonesia, UAE, and Malaysia have been regularly mentioned in tweets, as well as non-Muslim countries such as China, Japan, Thailand, and Singapore.

In discussing halal tourism, users tend to associate food and hotels with halal tourism (refer to Figs. 3, 4 and 5). This aligns with the definition of halal tourism by Carboni et al. [10]: "tourism in accordance with Islam, involving people of the Muslim faith who are interested in keeping with their personal religious habits whilst travelling". Such information is vital as it provides us with insight into what consumers expect as halal tourists. Therefore, companies that venture into the halal tourism industry must take this into account when promoting their halal tourism products.

The third major finding relates to the topics discussed within the halal tourism tweets. It was found that topic 4 (trip, can, will, get, know) is the most common topic, followed by topic 6 (tourism, via, japan, market, turkey). This illustrates that tourists are most interested in knowledge discovery and the halal tourism market. For example, users could use feedback from others to discover halal tourism destinations. Topics 1 (travel, halaltravel, guid, destin, halaltour) and 7 (halal, muslim, friend, look, islam) are third and fourth most common where users tweeted about halal travel guides and the concept of halal tourism to spread awareness about it. Other common topics relate to more specific aspects of halal tourism, such as topic 5, about halal food in the halal tourism destinations. Additionally, topics 2 and 3 are about the halal world travel awards in 2018 and halal hotels, respectively, as users tweeted about the winners of the halal world travel awards and their favorite hotels.



Lastly, the findings illustrate that sentiments towards halal tourism are either positive or negative (Fig. 7). Previous studies on halal food presented similar findings [1, 36]. Our analysis to divide the sentiments into emotions as suggested by Mohammad & Turney [62] found that there were more positive emotions such as trust, joy, and anticipation expressed compared to negative emotions such as fear and sadness, as illustrated in Fig. 8. Having trust in the first place sends a strong signal that users trust the whole halal tourism concept and individual services offered in halal tourism like halal food. Joy and anticipation came next, which could be explained by users enjoying halal food and services and expecting new services to emerge. Considering the newness of halal tourism as a concept, users anticipate new initiatives from governments and agencies. With regards to the negative emotions, we can only assume that when halal services are not available in a destination user chose, it elicits a negative emotion such as sadness. Research has shown that users fear that some services such as swimming are banned because of the introduction of halal tourism [71]. This corroborates the appearance of fear among the emotions present in the tweets.

Thus, it can be concluded that Twitter users consider halal tourism to be something that they can trust and look forward to (anticipation) as well as celebrate with joy. Generally, tourists travel because they want to relax and enjoy themselves and halal tourism allows them to do so whilst maintaining their religious practices.

It must be highlighted here that in this study we were able to analyze a large volume of data whereas past studies such as [14, 18] analyzed less than 1000 set of data. The results of this study have shown that the social media data obtained is reliable as although the approach differs the findings are somewhat similar. For example, in this study it was illustrated that halal tourism is a global market and not restricted to Muslim countries alone. This finding supports Yagmur et al. [11]. In addition, this study found that in discussing halal tourism, users associate it to food and accommodation which is the same as what Carboni et al. [10] found.

This study illustrated that there were more positive emotions among the users compared to negative ones, which is consistent with what Al-Ansi & Han, [33] found using a survey.

Conclusion

This work analyzed tweets on halal tourism posted over ten years. One of the contributions of this work, compared to other recent articles, is the use of social media as an alternative to the traditional survey. Millions of tweets are posted on Twitter every day and contain unbiased and rich opinions. Interviews introduce bias as the interviewers could impact responses from the interviewees by expressing their opinion. Conversely, consumers willingly publish their opinions online on Twitter, which is free of external bias. Social media is a powerful platform for shaping marketing strategies, and businesses and hotels should rely on users' opinions and incorporate them into their services. At a higher level, governments can also use guidance from social media to establish their countries as top halal tourism destinations. For instance, Japan and Thailand, which are non-Muslim countries, are highly ranked as popular halal tourism destinations. Businesses and governments can use the semantic network, sentiment, and word list analyses presented in this work to align their strategy on halal tourism going forward. Although this work has analyzed a large

amount of data on halal tourism, it could be expanded upon in the future to complement the results of this work. Researchers should collect data from other social media platforms and combine the results with the analysis done in this paper. We also recommend focusing on specific countries or continents to tailor results for less generalized locations. This study has analyzed sentiments related to halal tourism tweets, with some limitations. First, we used a lexicon-based approach in this work, which may fail to recognize some human expressions like sarcasm and irony. Second, some Twitter accounts may represent vendors in disguise; We did not perform account verification checks to ascertain the authenticity of Twitter accounts used. Thus, future research should attempt to identify and eliminate such accounts from the analysis. Third, we did not compare different database stores and the paradigm for storing them, such as MySQL vs SQLite. In addition, the analysis carried out did not carry out deeper analytics such as a graph to visualize co-occurrence groups. We are suggesting that future research should include these analytics.

Acknowledgements

We would like to thank all the reviewers for their comments.

Authors' contributions

All authors have made substantial contributions to the article. Ali and Firdaus extracted the data and conducted the analysis while Mostafa interpreted the data. Ainin was responsible for the overall publication including discussion, reviewing, and editing of the paper whereas Zalina was involved in the literature and discussion related to Halal Tourism. Ainin and Zalina were also responsible for the acquisition of research funding. The initial drafts of the paper were done by Ali and Mostafa. All authors read and approved the final manuscript.

Funding

We would like to thank the Malaysian Higher Education Consortium of Halal Institute, Ministry of Higher Education for providing the funds to conduct this research. In addition, we thank Universiti Malaya for providing all the research support, infrastructure, and facilities (Grant no: MO 001–2018 and MO 002–2018).

Availability of data and materials

All data generated or analyzed during this study are included in this published article. The datasets generated and/or analyzed during the current study are available from the corresponding author on reasonable request.

Declarations

Ethics approval and consent to participate

Not applicable. This research does not involve any human participants, human data, or human tissue. Data is obtained from Twitter. The data is extracted using Python script.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

Author details

¹UM Halal Research Center, University of Malaya, 50603 Kuala Lumpur, Malaysia. ²Gulf University for Science and Technology, West Mishref, Kuwait. ³Academy of Islamic Studies, Universiti Malaya, 50603 Kuala Lumpur, Malaysia. ⁴Faculty of Computer Systems and Software Engineering, Universiti Malaysia Pahang, 26300 Gambang, Kuantan, Pahang, Malaysia.

Received: 3 November 2020 Accepted: 10 May 2021

Published online: 22 May 2021

References

1. Mostafa MM. Clustering halal food consumers: a Twitter sentiment analysis. *Int J Mark Res.* 2019;61(3):320–37.
2. Rammal HG, Zurbruegg R. Awareness of Islamic banking products among Muslims: the case of Australia. In: *Islamic finance.* 2016; Springer. p. 141–156.
3. The Chemical Engineer, A future in halal cosmetics. 2010 (827): p. 12–24. Available at <https://www.thechemicalengineer.com/magazine/issues/issue-827/>.
4. Bin Abdullah AB. Halal vaccine and the ethical dimension of vaccination programmes. *Islam Civilis Renew.* 2014;274(1713):1–4.
5. Battour MM, Ismail MN, Battor M. Toward a halal tourism market. *Tour Anal.* 2010;15(4):461–70.

6. Zakaria Z, Ismail SZ. Perkembangan pengaturan halal menerusi akta perihal dagangan 2011 di Malaysia. *Jurnal Syariah*. 2015;23(2):189–206.
7. Battour M, Ismail MN. Halal tourism: Concepts, practises, challenges and future. *Tour Manag Perspect*. 2016;19:150–4.
8. Ahmed MJ, Akbaba A. The potential of halal tourism in Ethiopia: opportunities, challenges and prospects. *Uluslararası Güncel Turizm Araştırmaları Dergisi*. 2018;2(1):13–22.
9. Jafari J, Scott N. Muslim world and its tourisms. *Ann Tour Res*. 2014;44(1):1–19.
10. Carboni M, Perelli C, Sistu G. Is Islamic tourism a viable option for Tunisian tourism? Insights from Djerba. *Tour Manag Perspect*. 2014;11(7):1–9.
11. Yagmur Y, Ehtiyar R, Aksu A. Evaluation of halal tourism in terms of bibliometric characteristics. *J Islam Market*. 2019;11(6):1–17.
12. Sahida W et al. The implementation of syariah compliance concept hotel: De Palma Hotel Ampang, Malaysia. In: 2nd international conference on humanities, historical and social sciences. Citeseer; 2011.
13. Islam T, Chandrasekaran U. Halal marketing: growing the pie. *Int J Manag Res Rev*. 2013;3(12):3938.
14. Salleh NZM, et al. Challenges in implementing Islamic friendly hotel in Malaysia. *J Econ Info*. 2019;6(4):15–7.
15. Battour M, Ismail MN, Battor M. The impact of destination attributes on Muslim tourist's choice. *Int J Tour Res*. 2011;13(6):527–40.
16. Saad HE, Ali BN. Sharia-compliant hotels in Egypt: concept and challenges. *Adv Hosp Tour Res*. 2014;2(1):1–15.
17. Bastaman A. Lombok Islamic tourism attractiveness: non-moslem perspectives. *Int J Supply Chain Manag*. 2018;7(2):206–10.
18. Ristawati H, Saufi A, Rinuastuti BH. Effect of customer experience and innovative value on halal destination image and satisfaction of domestic tourists in Lombok Island. *Glob J Manag Bus*. 2019;19(3):1–7.
19. Heikinheimo V, et al. User-generated geographic information for visitor monitoring in a national park: a comparison of social media data and visitor survey. *Int J Geo-Inf*. 2017;6(3):85–100.
20. Feizollah A, et al. Halal products on Twitter: data extraction and sentiment analysis using stack of deep learning algorithms. *IEEE Access*. 2019;7(3):83354–62.
21. Moshin A, Brochado A, Rodrigues H. Halal tourism is traveling fast: Community perceptions and implications. *J Destin Market Manag*. 2020;18(5):100503–20.
22. Mohsin A, Ramli N, Alkhulayfi BA. Halal tourism: emerging opportunities. *Tour Manag Perspect*. 2016;19:137–43.
23. Papastathopoulos A, Kaminakis K, Mertzanis C. What services do Muslim tourists want? Uncovering nonlinear relationships and unobserved heterogeneity. *Tour Manag Perspect*. 2020;35(5):100720.
24. Fakir F, Erraoui E. Moroccan tourist's perceptions toward Halal Tourism. In: 2nd international halal tourism congress. Antalya, Turkey; 2019.
25. Aini Q, Khudzaeva E. Potential halal tourism destinations with applying K-means clustering. *Int J Intell Syst Appl*. 2019;11(9):9–17.
26. Shakona M, et al. Understanding the traveling behavior of Muslims in the United States. *Int J Cult Tour Hosp Res*. 2015;9(1):22–35.
27. Jia X, Chaozhi Z. "Halal tourism": is it the same trend in non-Islamic destinations with Islamic destinations? *Asia Pac J Tour Res*. 2020;25(2):189–204.
28. Mannaa MT. Halal food in the tourist destination and its importance for Muslim travellers. *Curr Issue Tour*. 2020;23(17):2195–206.
29. Suhartanto D, et al. Tourist experience in Halal tourism: what leads to loyalty? *Curr Issue Tour*. 2020;3(2):1–15.
30. Han H, et al. Exploring halal-friendly destination attributes in South Korea: perceptions and behaviors of Muslim travelers toward a non-Muslim destination. *Tour Manage*. 2019;71:151–64.
31. Jaelani A. Halal tourism industry in Indonesia: potential and prospects. *Int Rev Manag Mark*. 2017;7(3):1–19.
32. Yousaf S, Xiucheng F. Halal culinary and tourism marketing strategies on government websites: a preliminary analysis. *Tour Manage*. 2018;68(4):423–43.
33. Al-Ansi A, Han H. Role of halal-friendly destination performances, value, satisfaction, and trust in generating destination image and loyalty. *J Destin Market Manag*. 2019;13(4):51–60.
34. Jeaheng Y, Al-Ansi A, Han H. Halal-friendly hotels: Impact of halal-friendly attributes on guest purchase behaviors in the Thailand hotel industry. *J Travel Tour Mark*. 2019;36(6):729–46.
35. Oktadiana H, Pearce PL, Chon K. Muslim travellers' needs: what don't we know? *Tour Manag Perspect*. 2016;20(3):124–30.
36. Prameswari P, Surjandari I, Laoh E. Mining online reviews in Indonesia's priority tourist destinations using sentiment analysis and text summarization approach. In: 2017 IEEE 8th international conference on awareness science and technology (iCAST), Taichung, Taiwan: IEEE; 2017.
37. Nguyen TT, et al. Twitter-derived measures of sentiment towards minorities (2015–2016) and associations with low birth weight and preterm birth in the United States. *Comput Hum Behav*. 2018;89(3):308–15.
38. Philander K, Zhong Y. Twitter sentiment analysis: capturing sentiment from integrated resort tweets. *Int J Hosp Manag*. 2016;2016(55):16–24.
39. Oliveira N, Cortez P, Areal N. The impact of microblogging data for stock market prediction: using Twitter to predict returns, volatility, trading volume and survey sentiment indices. *Expert Syst Appl*. 2017;73:125–44.
40. Sehl, K. Top Twitter demographics that matter to social media marketers. 2020 [cited 2020 1st November]. <https://blog.hootsuite.com/twitter-demographics/>.
41. Graham M, Hale SA, Gaffney D. Where in the world are you? Geolocation and language identification in Twitter. *Prof Geogr*. 2014;66(4):568–78.
42. Geetha M, Singha P, Sinha S. Relationship between customer sentiment and online customer ratings for hotels-An empirical analysis. *Tour Manage*. 2017;61(8):43–54.
43. Jiao J, Holmes M, Griffin GP. Revisiting image of the city in cyberspace: analysis of spatial Twitter messages during a special event. *J Urban Technol*. 2018;25(3):65–82.

44. R Development Core Team. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria; 2019.
45. Csardi G, Nepusz T. The igraph software package for complex network research. *Int J Complex Syst*. 2006;1695(5):1–9.
46. O'Leary DE. Blog mining-review and extensions: "From each according to his opinion." *Decis Support Syst*. 2011;51(4):821–30.
47. Barlow M. Software for corpus access and analysis. In: *How to use corpora in language teaching*. John Benjamins Publishing; 2004. p. 205–221.
48. Barnbrook G. *Language and computers: a practical introduction to the computer analysis of language*. Edinburgh University Press; 1996. p. 65.
49. Kwon KH, et al. Social media rumors as improvised public opinion: semantic network analyses of twitter discourses during Korean saber rattling 2013. *Asian J Commun*. 2016;26(3):201–22.
50. Krippendorff K. *Content analysis an introduction to its methodology*. Sage publications; 2018.
51. Blei DM. Probabilistic topic models. *Commun ACM*. 2012;55(4):77–84.
52. Fischer-Preßler D, Schwemmer C, Fischbach K. Collective sense-making in times of crisis: connecting terror management theory with Twitter user reactions to the Berlin terrorist attack. *Comput Hum Behav*. 2019;100(1):138–51.
53. Wold S, Esbensen K, Geladi P. Principal component analysis. *Chemom Intell Lab Syst*. 1987;2(1–3):37–52.
54. Van der Maaten L, Hinton G. Visualizing data using t-SNE. *J Mach Learn Res*. 2008;9(11):2579–605.
55. Miao Q, Li Q, Zeng D. Fine-grained opinion mining by integrating multiple review sources. *J Am Soc Inf Sci Technol*. 2010;61(11):2288–99.
56. Stone PJ, Dunphy DC, Smith MS, Ogilvie DM, and Associates. *The General Inquirer: A Computer Approach to Content Analysis*. MIT Press, Cambridge, MA; 1966.
57. Baccianella S, Esuli A, Sebastiani F. Sentiwordnet 3.0: an enhanced lexical resource for sentiment analysis and opinion mining. 7th International Conference on Language Resources and Evaluation (LREC), Malta. 2010.
58. Pennebaker JW, Mehl MR, Niederhoffer KG. Psychological aspects of natural language use: our words, our selves. *Annu Rev Psychol*. 2003;54(1):547–77.
59. Agerri R, García-Serrano A. Q-WordNet: extracting polarity from WordNet Senses. 7th International Conference on Language Resources and Evaluation, Malta. 2010.
60. Wiebe J, et al. Learning subjective language. *Comput Linguist*. 2004;30(3):277–308.
61. Taboada M, et al. Lexicon-based methods for sentiment analysis. *Comput Linguist*. 2011;37(2):267–307.
62. Mohammad SM, Turney PD. Crowdsourcing a word–emotion association lexicon. *Comput Intell*. 2013;29(3):436–65.
63. Oliveira DJS, Bermejo PHS, Dos Santos PA. Can social media reveal the preferences of voters? A comparison between sentiment analysis and traditional opinion polls. *J Inf Technol Polit*. 2017;14(1):34–45.
64. Wu Y, Yang Y, Chiu C-Y. Responses to religious norm defection: the case of Hui Chinese Muslims not following the halal diet. *Int J Intercult Relat*. 2014;39:1–8.
65. Koylu C, et al. CarSenToGram: geovisual text analytics for exploring spatiotemporal variation in public discourse on Twitter. *Cartogr Geogr Inf Sci*. 2019;46(1):57–71.
66. Engonopoulos N et al. ELS: a word-level method for entity-level sentiment analysis. In: *proceedings of the international conference on web intelligence, mining and semantics*. 2011.
67. Farra N et al. Sentence-level and document-level sentiment mining for arabic texts. In: *2010 IEEE international conference on data mining workshops*. IEEE; 2010.
68. Yessenalina A, Yue Y, Cardie C. Multi-level structured models for document-level sentiment classification. In: *Proceedings of the 2010 conference on empirical methods in natural language processing*. Association for Computational Linguistics; 2010.
69. Yuan EJ, Feng M, Danowski JA. "Privacy" in semantic networks on Chinese Social Media: the case of Sina Weibo. *J Commun*. 2013;63(6):1011–31.
70. Japan.halal. *Halal tourism is on the increase in Japan*. 2020 [cited 2020 1st November]. <https://www.japan.travel/en/plan/muslim-travelers/>.
71. Elok Sari, P.N.A.E.A. *Ban on swimming? Myths hinder Indonesia's development in halal tourism*. 2019 [cited 2020 1st December]. <https://halalfocus.net/ban-on-swimming-myths-hinder-indonesias-development-in-halal-tourism/>.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.