Study on Online Travel Review Analysis for Tourism Investigation

LABORATORY OF HARMONIOUS SYSTEMS ENGINEERING
GRADUATE SCHOOL OF INFORMATION SCIENCE AND TECHNOLOGY

Shuang SONG

HOKKAIDO UNIVERSITY
JAPAN

Submitted in partial fulfillment of the requirements for
the Degree of Doctor of Information Science and Technology, Hokkaido University,
February, 2020

Contents

List of figures

List of tables

Publications arising from the thesis

- S. Song, H. Kawamura, J. Uchida, and H. Saito: “Determining tourist satisfaction from travel reviews”, Journal of Information Technology & Tourism, Vol. 21, No. 3, pp. 337-367 (2019)

- S. Song, T. Yamashita, H. Kawamura, J. Uchida, and H. Saito: “Determining the Number of Tourist Arrivals from Travel Reviews A Bivariate Correlation Analysis”. Proceedings of the 20th Asia Pacific Industrial Engineering and Management System Conference, Kanazawa, Japan, pp. 244-248 (2019)

- S. Song, H. Saito, and H. Kawamura: “Content analysis of travel reviews: exploring the needs of tourists from different countries”, Information and communication technologies in tourism 2018 - Proceedings of the International Conference in Jönköping, Sweden, January 24-26, 2018 -, Springer, Cham, pp. 93-105 (2018)

- S. Song, H. Kawamura, J. Uchida, and H. Saito: “Towards a New Investigation Method for Tourists' Needs: from Travel Survey to the Analysis of Travel Reviews”, Proceedings of the 1st International Conference on Digital Practice for Science, Technology, Education, and Management, Sapporo, Japan, pp. 11-16 (2018)

Statement of original authorship

This work contained in this thesis has not been previously submitted to meet requirements for an award at this or any other higher education. To the best of my knowledge and belief, the thesis contains no material previously published or written by another person except where the due reference is made.

Signature:

Date:

Chapter 1 Introduction

1.1 Background

Tourism is a growing industry that brings communities major economic benefits (JTA, 2015). According to the World Tourism Organization (UNWTO), as a worldwide export category, tourism ranked third in 2017 after chemicals and fuels and ahead of automatic products (UNWTO, 2018a). Given its economic impact, tourism has become a key industry in many countries.

To achieve a sustainable growth in tourism, it is important to evaluate the current status of the industry and to monitor and investigate the diverse needs of both international and domestic tourists. Taken Japan as an example, the Japanese government aims to become a tourism-oriented country by achieving the target of 40 million visitors in 2020 and by increasing international and domestic travel consumption (JTA, 2019a). In order to achieve such a vision, appropriate measures are required for the improvement of the creation of attractive tourist areas, for the implementation of Visit Japan promotions, and etc. (JTA, 2013). For decision makings during the development of policies or the engagements of marketing activities, many tourism surveys have been carried out to collect statistical information on a quarterly or annually basis. For example, at the national level, the Japan Tourism Agency (JTA) has been conducting the Accommodation Survey, the National Tourism Survey, and the International Visitor Survey since 2010 (JTA, 2013). At the prefectural level, local governments are conducting surveys based on JTA's *Common Standard for Statistics on Inbound Tourists and Survey Procedure* (JTA, 2019b) to estimate the number of inbound and/or domestic tourists, per-capita tourism consumption, and overall tourism consumption.

In addition to the above mentioned surveys that collect factual statistics, some other surveys are conducted to investigate subjective topics that are related to tourist needs such as travel intentions, expectations, and/or satisfactions. Several examples in Japan are the Survey Concerning Customer Satisfaction implemented by the Hokkaido government (2016), Fukushima Destination Survey by the Fukushima government (2019), and Tourists Satisfaction Survey by the Okayama government (2010).

An ideal survey of tourist's behavior or needs should involve every potential tourist; however, the time and effort needed for such surveys makes them impractical. Consequently, sampling is frequently used in current tourism surveys, where the answers given by a quantity of representative tourists are used to represent all tourists. Nevertheless, acquiring even these samples for interviews or questionnaires is still expensive and time consuming. Besides, although the duration of interview or answering questionnaires is usually controlled, participating in an interview during a trip will inevitably put certain amount of burdens on tourists. That could also be the reason that sometimes the answers could be superficial or extreme. Nevertheless, there is barely a method to detect such biased results except for the comparisons for unusual responses or checking for logical inconsistency. As a result, traditional survey methods, although well-accepted, still have many limitations regarding the cost, response burden, and the uncertainty of the precision of results.

Currently, millions of users post text, photos, or videos to the Internet. Recently, an increasing number of researchers have started to focus on analyzing those online data using either automated or manual analysis. Previous studies in the field of tourism informatics have confirmed the possibility to extract and to interpret travel-related information such as demographic profiling (Wenger, 2008; Fujii et al., 2017), preferences (Marrese-Taylor, 2013), complaints (Maurer & Schaich, 2011; Del Chiappa & Dall'Aglio, 2012; Levy et al., 2013; Liu et al., 2017), and destination images (Murakami & Kawamura, 2011; Serna et al., 2013; Serna et al., 2016; Suzuki & Kurata, 2017) from textual online data. It is also possible to find valuable information in user ratings (Antonio et al., 2018), geographic information (Saeki et al., 2015), and photos (Stepchenkova et al., 2015) as well as in textual data. Indicators such as user ratings values and the frequencies of certain words are often adopted to investigate tourist satisfaction (Lu & Stepchenkova, 2012; Jannach et al., 2014; Zhang & Cole, 2016) or destination images (Schmallegger & Carson, 2009; Dickinger & Költringer, 2011; Tseng et al., 2015; Marine-Roig & Clavé, 2016) while latitudes and longitudes and the names of locations are used to investigate locations that tourists visit (Vu et al., 2015; Saeki et al., 2015).

Making productive use of such copious real-time data potentially makes it possible to significantly reduce the cost of sample collecting and response burden. However, it is notified that online data are not deliberately designed for data analysis and thus, they do not have a well-designed target population, structure and quality (Struijs et al., 2014). In addition, the analysis of online data introduces external variables such as Internet use (Ferrer-Rosell et al., 2017), the choice of online platform (Xiang, 2017), the motivation of posting data (Cheung and Lee, 2012; Yang, 2017), or the choice of words for expression a certain emotion (Fernández et al., 2000). Therefore, it remains unclear that whether online data analysis can generate valid results that reflect tourist behavior or tourist needs.

In this study, the term validity is adopted to describe whether the research truly measures that which it was intended to measure (cited in Golafshani, 2003). It is presented in the agreement between two attempts to measure the same trait through maximally different methods (Campbell & Fiske, 1967; cited in Hammersley, 1987). Examining the validity helps to address the usefulness, as well as to understand the risk and limitation of drawing inference from the results of online data analysis. Nevertheless, only a handful of studies (Daas et al., 2015, Saeki et al., 2015; Xiang et al., 2015; Dickinger & Lalicic, 2016; Brakel et al., 2017; Ferrer-Rosell et al., 2017) were devoted to this fundamental aspect of online data analysis.

Another concern of this work is relevant to the investigations of international tourists. To achieve a sustainable growth of inbound tourism, it is important to understand the differences and similarities among tourists who are coming from different cultures. However, extant cross-culture touristic studies are largely limited to English-speaking regions (Li, 2012), and only a few studies include cross-language analysis using online data. Therefore, more insights and hints are needed to advance the understanding of the nature of online data and the differences and/or similarities of such natures among tourists from various regions with different language backgrounds.

1.2 The statement of problem

As mentioned above, both traditional survey methodologies and online data analysis suffer from certain limitations, which could prevent them from providing valid information for further touristic decision-makings. Besides, the understanding of online data analysis is also insufficient in the cross-region and especially in cross-language context.

To show the validity of online data analysis, the approach adopted by primary studies is to the comparison of its results and the ones collected via those well-accepted survey methodologies. For example, Brakel et al. (2017) compared Facebook and Twitter messages analysis to an administrative survey and found that results from these two methods suffer the same degree of selection bias and measurement bias. While in the context of tourism investigation, Saeki et al. (2015) examined the ordinal correlation between the results of a governmental survey and the results of Twitter data analysis regarding tourist arrivals. Their results showed positive correlations between two rankings but the sample size was statistically insufficient (i.e. n = 10 areas $\times$ 3 months). Plank (2016) found the accuracy of user-generated content on outdoor sports activities such as ski tours is limited by comparing user-reported avalanche danger levels to the ones provided by the Austrian Ministry of the Interior. Dickinger and Lalicic (2016) compared offline and online data to examine how emotions reflected in TripAdvisor reviews linked to destination brand personality. They addressed the differences between the proportions of English review (n = 1,104) and open-ended questionnaires (n = 599) that contain words with different meanings and different emotions.

Because very few previous studies focused on examining the validity of online data analysis, the understanding towards this topic is still insufficient. There are many remaining uncertainties. For example, is online data analysis a valid replacement for traditional tourism surveys? Or can it be used as an assisting tool for finding more detailed information at shorter internals? This work believes that finding out the role of online data analysis is necessary, so valid inference can be drawn from these analyses.

Furthermore, to address the validity of using online data analysis as a proxy for traditional surveys, previous studies focused more on the consistencies in results. This work believes that valuable information could also be revealed from the inconsistencies. For example, inconsistencies could help to recognize the limitations of each method, and hopefully, help to detect certain biases or errors hiding in the results.

Meanwhile, regarding cross-region tourism investigations based on online data analysis, the focuses of previous studies are limited to the comparison of the statistical distributions of the occurrence frequencies of words (He et al., 2012; Hatoh et al., 2013; Buzova et al., 2019; Nakayama and Wan, 2019a; Nakayama and Wan, 2019b) and / or user ratings (Antonio et al., 2018) between tourists from various regions. However, several concerns have been pointed out. For example, the normative system of emotional display rules varies by culture (Fernández et al., 2000; Davis et al., 2012). As a result, emotions (or attitudes) that could be observed from online data are not necessarily equivalent to the internal emotions and thus using comparisons of words or ratings could be an inappropriate measurement of the differences of true feelings among various cultures. This study hopes to find more insights about this issue by comparing the differences in results of online data analysis and the ones of traditional surveys.

1.3 Research aims and objectives

Based on the above, this study aims to establish online data analysis as a valid alternative for tourism investigation. For that purpose, this works shows the validity of employing the analysis of TripAdvisor reviews (see *www.tripadvisor.jp*) for tourist satisfaction and tourist arrivals investigation in cross-region and cross-language context.

The reasons for choosing TripAdvisor, tourist satisfaction investigation, and tourist arrivals investigation are as follows.

TripAdvisor is the largest travel website in the world. Among various data source (e.g., travel websites, travel blogs or SNS) on the Internet, TripAdvisor was chosen as a case study for the following reasons: 1) when compared to SNS, data from travel websites and travel blogs doesn't require extra content filtering to extra travel-related data and allows the identification of tourism destinations and facilities, and 2) among various popular travel websites and travel blogs, TripAdvisor possesses the largest amount of data.

Tourist satisfaction and tourist arrivals investigation are the two major and fundamental type of tourism investigation. Tourist satisfaction refers to the cognitive and emotional reflection of tourist attitudes toward tourism services (Bowen & Clarke, 2002). It is an important subjective evaluation that can be used for identifying problematic aspects of tourism institutions and encouraging more word-of-mouth recommendations (Berezina et al. 2016). However, tourist satisfaction is difficult to measure due to the complexity of the human mind (Tourangeau et al., 2002). Therefore, conducting appropriate investigations requires careful investigation design and staff trainings, which makes satisfaction investigations more costly than the others. On the other hand, online data is known for its potential to reflect tourists' thoughts and feelings (Serna et al, 2013). Therefore, showing it is valid to use online data analysis for tourist satisfaction investigation could allow administrative bodies to easily conduct quicker and worldwide tourist satisfaction investigation, help facility owners to identify their problems quicker, and reduce the burden of tourists for providing information.

The investigation of tourist arrivals, on the other hand, is a universal investigation conducted by governments all over the world. It is used to measure the flows of visitor from a certain area to a tourism destination (UNWTO, 2018; JTA, 2019b). And it could serve as an objective indicator for comparisons among destinations or facilities (JTA, 2019d), help responsive decision-makings (Yagasaki, 2015), and etc. Therefore, showing the validity of using online data analysis for estimating tourist arrivals could bring potential savings for all governments in the world and enable quicker comparisons among destinations or facilities.

The rest of the thesis is organized as following.

Chapter 2 illustrates the current status of tourism surveys and investigations, and introduces related works in the field of survey methodology, tourism informatics, and cross-region and cross-language tourism analyses.

Chapter 3 introduces a method for investigating whether a data source is suitable for tourism investigation by clarifying the types and volumes of available contents. Taken TripAdvisor as an example, this work investigates the structure of its travel website. And then, both manual and automated content analysis is applied to randomly sampled reviews that are written by tourists from six selected countries with various cultures. By that means, this work clarifies the contents of textual reviews using categorizations commonly adopted in traditional surveys, the percentages of reviews including material in each categorization, and the differences and similarities among various countries.

Chapter 4 aims to show the validity of determining tourist satisfaction based on online data analysis. This work introduces a method to conduct comparisons between the attitudes extracted from travel reviews and results recorded in traditional surveys. First, this work reviews scholarly literatures on tourist satisfaction, physiological background of traditional satisfaction surveys, motivation of reviewers, and sentimental analysis. And then, a method for extracting attitudes embedded in travel reviews and a method for investigating the correlation between the attitudes of reviewers and the satisfaction of interviewees in the traditional survey are presented. Moreover, data collection and manual analysis is conducted taking the satisfaction survey conducted by Hokkaido Government (2016) as an example. As a result, this work presents findings that can be incorporated into future automated analysis and that can improve the understanding towards online travel review analysis and traditional survey and towards cross-region and cross-language comparisons.

Chapter 5 introduces a method to show the agreement between the use of online travel review analysis and traditional method for finding tourist arrivals. This work presents a method to calculate the number of reviewers to be compared with the number of tourist arrivals recorded in traditional surveys. Then, taken the Hokkaido Government's statistics on the number of tourist arrivals and accommodations (2019) as an example, a bivariate analysis is conducted to investigate the correlation between the number of reviewers and the number of sightseeing visitors, as well as the correlation between the number of reviewers and the number of overnight travelers by regions of residence. In addition, to address the potential of analyzing tourist preferences, this work also examines the proportions of reviewers-travelers by month and by city.

Chapter 6 provides a summary and draws conclusions. It presents the implications of the study for practical and academic fields, identifies the limitations and indicates directions for future research in the area.

Chapter 2 Literature review

This study involves multiple fields of research including survey methodology, tourism informatics, and cross-region and cross-language comparisons. This chapter is missioned to introduce related works in each field.

2.1 Survey on current tourism surveys and investigations

In order to develop a method that can act as a proxy for traditional survey methods, it is necessary to know the current status of tourism surveys and investigations. This section is a survey that describes the scale, frequency, and/or the content of the world tourism surveys and investigations, and the local ones taken Japan as an example. In the first part, this work only focuses on English-speaking, Japanese-speaking, or Chinese-speaking countries/regions due to language constraints.

2.1.1 World tourism surveys and investigations

Tourism investigations are conducted to collect tourism statistics all over the world. According to the World Tourism Organization (UNWTO), over a hundred countries/regions are taking administrative investigations, whose results have been reported to UNWTO and recorded in the *UNWTO Compendium of Tourism Statistics* (UNWTO, 2019). The publication of this compendium was first started at 1975 and became annually since 1986. It contains 145 internationally comparable basic data series and indicators on six areas including inbound tourism, domestic tourism, outbound tourism, tourism industry, employment, and macro economy indicators. For example, inbound tourism has eight major types of basic data listed as 1) arrivals, 2) arrivals by region, 3) arrivals by main purpose, 4) arrivals by mode of transport, 5) arrivals by forms of organization of the trip, 6) accommodation, 7) expenditure, and 8) expenditure by main purpose of the trip. Meanwhile, the domestic tourism section contains the following five types of basic data: 1) trips, 2) trips by main purpose, 3) trips by mode of transport, 4) trips by form of organization, and 5) accommodation.

Another example of global tourism statistic is the *European Tourism Indicator System* (ETIS) proposed by the European Commission (2016), and is adopted in Spain, Italy, Slovenia, and etc. It contains a set of 43 core indicators and several supplementary indicators used to monitor and manage tourism destinations. Each indicator contains reason for measuring, data requirements, units of measurement, terms in glossary, data collection instructions, method of calculation, frequency data collection, reporting format, international benchmarks, key users, suggested actions, and references. The ETIS scheme includes a visitor survey, a resident survey, an enterprise survey, and a destination management survey. For example, the visitor survey includes the following indicators.

- Average travel (km) by tourists and same day visitors from home to the destination
- Percentage of tourists and same day visitors using different modes of transport to arrive at the destination
- Percentage of tourists and same day visitors using local/soft mobility/public transport services to get around the destination
- Number of tourist nights per month
- Percentage of repeat/return visitors (within 5 years)
- Daily spending per overnight tourist (accommodation, food and drinks, other services)
- Daily spending per same day visitor
- Percentage of visitors that are satisfied with their overall experience in the destination
- Percentage of visitors satisfied with the accessibility of the destination for those with disabilities or specific access requirements

In the United State, the national travel and tourism office (NTTO) (see *travel.trade.gov*) is carrying out the monthly-based Survey of *International Air Travelers Departing the United States* (SIAT) since 2012 (U.S. Department of Transportation, 2019). It is designed to assess the economic impact of international visitation contributed by non-US residents who have visited the country and US residents who was traveling abroad. It adopts the random sampling techniques and contains altogether 32 questions that take an estimate of 15 minute to answer. Questions include the date of travel, the choice of airline (the name of the airline, the flight number, and boarding airport), address of residence, nationality, country of birth, source and media of information when planning the trip, the date of planning the trip, whether receive vaccine or medicine for this trip, media of reservation, payment of insurance, media of reservation for accommodation, main purpose of the trip, travel companion, visited places, contents of travel packages, expenditures, mode of transportations, participated activities, satisfactions towards the airline and the airport, intention of re-visit, overall satisfaction compared to expectations, frequency of visiting the U.S., occupation age, gender, annual household income, ethnicity, race and etc.

Canada has 14 active travel and tourism surveys and statistical programs and 6 inactive ones (Statistics Canada, 2019). For example, the *Travel Survey of Residents of Canada* (TSRC) from 2005 to 2017 was a major source of data used to measure the size and status of Canada's tourism industry. It was a quarterly-based survey involving about 56,000 households in Canada with a response rate of 70 to 75%. It collects information about the volume of trips and expenditures from Canadian residents by trip origin, destination, duration, type of accommodation used, trip reason, mode of travel, etc. (Statistics Canada, 2018). It is interesting that Canada is one of the few countries that conduct travel attitudes and/or travel motivation surveys at the national level. There was a one-time only *Tourism Attitude and Motivation Study* in 1995 (Statistics Canada, 1995), and occasional *Travel Activities and Motivation Survey* (TAMS) until 2006 (Ontario Ministry of Tourism, Culture and Sport, 2006). The TAMS used the random digit dialing sampling technique whereby the telephone numbers are generated randomly by computer. It examined recreational activities with a detailed activity list and travel habits of Canadians and Americans as well as travel motivators (e.g., to see or do something new and different, to seek solitude and isolation, or to gain knowledge of history, other cultures or other places), places visited, type of accommodation used, impressions of Canada.

The Great Britain has a variety of tourism surveys concerning the volume and value, tourism business, consumer behavior, product development, inbound tourism performance, sector-specific researches (e.g., accommodation, culture, or food), and destination competitiveness under its official tourism statistics website, VisitBritain (see *www.visitbritain.org*). For example, the *GB Tourism Survey* (BTA, 2019a) is a monthly-based national consumer survey that measures the number of trips, the number of bed nights, and expenditure of domestic overnight tourism trips while the *GB Day Visits Survey* (BTA, 2019b) focus on the one-day visits. The *England / UK Occupancy Survey* (BTA, 2019c) records the monthly bedroom and bed-space in the accommodation sector. It collects data from a panel of more than 3,000 hotels and other accommodation businesses. The *Annual Survey of Visits to Visitor Attractions* publishes the number of visitors in all tourist attractions in England annually (BTA, 2018).

Australia also has a variety of surveys and tourism-statistics-related publications under the direction of the Tourism Research Australia (TRA) (see *www.tra.gov.au*). The *National Visitor Survey* (NVS) and the International Visitor Survey (IVS) are the two main surveys. The NVS (TRA, 2019a) is an interview-based survey via mobile phone / landline using random digit dialing sampling. Each year, it involves 120,000 residents aged 15 years and over. It contains over 70 questions regarding destination, purpose, transport, travel package, sources to obtain information about the trip, activities, spend, accommodation, travel party, and demographics. Meanwhile, the IVS (TRA, 2019b) samples 40,000 short-term international travelers aged 15 years and over every year. Interviews with those travelers who are departing Australia are conducted at the eight major international airports in Australia. It contains around a hundred questions to collect information about the usual place of residence, repeat visitation, group tours, travel party, sources for obtaining information about Australia, purpose of visit and visited places, transportation and accommodation, activities, expenditure, and demographics. Also, there are other surveys focusing more about motivation and satisfaction such as the *Chinese Satisfaction Survey* in 2014, *Destination Visitor Survey* (DVS), *Visitor Profile and Satisfaction Survey* (VPS), and etc.

In China, administrative tourism statistics is published by the Ministry of Culture and Tourism of the People's Republic of China (MCT) (see *www.mct.gov.cn*). It has a list of hotels by stars, a list of tourism facilities by stars, reports of the number of arrivals and an announcement of the total amount of annual tourism income. However, survey methodologies or detailed results are not provided on the official website.

The Japan Tourism Agency (JTA) (see *www.mlit.go.jp/kankocho*) has been conducting the *Accommodation Survey*, *National Tourism Survey*, and *International Visitor Survey* on a quarterly basis since 2010 (JTA, 2013). The *Accommodation Survey* (JTA, 2019c) samples 18 thousand Japanese Inns, Hotels, Airbnb, and etc. in the country. The *National Tourism Survey* (JTA, 2019d) is designed to estimate the current status of tourism in Japan. It samples 26 thousand residents per season using random sample techniques. It is conducted during April, July, October, and January using mail-based self-administration survey method. The *International Visitor Survey* (JTA, 2019d) is also named the *Consumption Trend Survey for Foreigners Visiting Japan*. It targets at international tourists departing Japan with a targeted sample size of 34,964 per season. It employs the face-to-face interview at 29 airports in Japan. Interviewees need to answer a questionnaire including questions such as the date of travel, type of visa, landing airport, personal profiles (i.e. nationality, place of residence, age, and gender), form of travel (i.e. companies, frequency, former visits, purpose), destinations, accommodations, use of travel package, expenditure, tax-free, satisfaction, intention of re-visit, household income, and etc.

In summary, at the national level, tourism investigations are usually designed to estimate the scale of the tourism industry. Surveys on arrivals are conducted monthly with seasonal or annual report. Commonly adopted methods are face-to-face surveys, mobile/phone surveys, and mail surveys. Certain part of the contents such as the investigation of tourist arrivals, travel purpose, or destinations is common among various countries while the others are distinctive. For example, the SIAT in the U.S. includes questions regarding travel insurance, vaccine, ethnicity, and race; the activities in Canada include detailed categories in the winter sports sector; attitudes towards the wine is listed as an individual category in the Australia’s satisfaction surveys; in Japan, hot springs, the Japanese Inns, the Japanese alcoholics, and attitudes towards tax-free is investigated.

2.1.2 Case study of local tourism investigations: Japan

Apart from national surveys, a large quantity of surveys is conducted by the local governments. Because the research of those local surveys all over the world will take a tremendous amount of time, this work takes the ones undertaken in Japan as a case study to illustrate the current status of local tourism surveys and theirs differences from the national surveys.

In Japan, The Japan Tourism Agency (JTA), a sub-body of the Ministry of Land, Infrastructure, Transport and Tourism (MLIT), holds the publications of national tourism statistics. The JTA’s website includes a set of links directed to all 47 prefectural government’s official website of tourism and travel. Such a clear and strict structure makes it possible to conduct a throughout research over all local government in one country.

The list of local tourism surveys and touristic projects in Japan until Aug. 2018 is attached in Appendix A. Two main purposes of local tourism surveys can be identified from the reports of those surveys: the first one, which is similar to the national investigations, is to estimate the economic impact of local tourism industry, and the second purpose is to provide statistics for decision-making for policies and for the improvement of services to attract more tourists.

According to JTA’s publication, 46 out of 47 local governments except for Osaka are conducting surveys based on JTA’s *Common Standard for Statistics on Inbound Tourists and Survey Procedure* (JTA, 2019b). These surveys consist of two subordinate surveys: 1) on-the-spot survey on the actual number of tourists in each city, and 2) on-the-spot parameter surveys in each prefecture. The former takes the form of interview with the owners or managers at selected tourism facilities or tourism event; then, the results are reported to each prefectural government to be aggregated. The latter samples random tourists at selected tourism facilities and event. It collects information about place of residence, age, gender, one-day visit or overnight visit, accommodation, purpose, travel companions, places of visited, mode of transport, expenditure based on the questionnaire template provided in the document of JTA's Common Standard. In most cases, face-to-face questionnaire or leaving method (i.e. leave the questionnaire at certain locations, let the participants answer them on their own, and collected the questionnaire a few days later) at selected hotels are adopted. The results of those survey, together with the national statistics provided by JTA, enables the estimation of the number of inbound / domestic tourists, per-capita tourism consumption, and overall tourism consumption.

In addition, local governments also carry out distinctive surveys out of their own needs. Frequently, questions such as expectations, satisfactions, or the intention of re-visit are added to JTA's questionnaire template (2019b). At the same time, in prefectures such as Tokyo (Tokyo Metropolitan Government, 2018), Hokkaido (Hokkaido Government, 2016), or Okinawa (Okinawa Prefectural Government, 2019), governmental bodies design their own questionnaire and compose questions regarding satisfactions towards local cultural. Besides, web-based survey is used to investigated travel intention and the level of awareness of the destinations or their specialties in Niigata (Niigata Prefectural Government, 2019) and Shimane (Shimane Prefectural Government, 2019).

Furthermore, in Chiba, Gifu, Aichi, and Ehime, tourism investigation based on the analysis of SNS text and/or geographical information is employed. However, it should be noticed that in the reports of these investigations, only the results are published and the data or the methodology is unclear. For example, in Chiba's *Survey on the Needs and Trends of Foreigner Visitors using SNS* (Chiba Prefectural Government, 2016), an amount of a three-month's text data was collected via SNS in eleven countries such as South Korea or China. It extracted the topics related to travels to Japan, the activities, evaluations, and thoughts towards Chiba, and the volume and tendency about topics related to the Olympics. Meanwhile, in the investigation in Ehime (Ehime Prefecture, 2017), a quantity of 82,871 travel reviews regarding eleven nearby prefectures including Ehime itself were collected. The source of the travel reviews is unknown. Topics were extracted from the title and the comment in the reviews and aggregated by age, by type of companies, and by gender. To be specific, first the patterns of noun-verb or noun-adjective such as "take photos" or "many people" that appears more than 20 times are extracted using Text Mining Studio 5.0. Then, clustering was applied to the destinations and the patterns to extract seven topics using APOSTOOL2. Also, in Gifu and Aichi, geographical information is used to extract travel routings among tourists.

2.2 Online data and its utility in tourism investigation

2.2.1 An overview of the utility of online data

Data, in forms of texts, images, or sound, can be obtained from two main types of sources: online and offline (Orgad, 2009). Orgad (2009) defined online data as data that are obtained using methodologies implemented by and through the Internet, including texts of interviews with research participants that are conducted online, or participant observation in online spaces. The former refers to the data collected through e-mail or the use of web-based surveys (Topp and Pawloski, 2002; Granello and Wheaton, 2004; Lefever et al., 2007). The latter include those data that are provided by Internet-based data sources such as websites or social media platforms, frequently accompanied by data collecting approaches such as data crawling or data scraping (Zanasi. 2000; Pandey et al. 2004; Olston and Najork, 2010). And in this study, online data refers to the latter.

The content of online data, when categorized by the creator, can be grouped into marketer-generated content and user-generated content (Bronner and De Hoog, 2010; Lim et al., 2012; Goh et al., 2013). Marketer-generated content, or firm- or entrepreneurial-generated content, refers to promotional materials provided by marketers such as marketing organizations (Lim et al., 2012) or advertising agencies (Martínez-Navarro and Bigné, 2017). User-generated content such as reviews, photos, or video clips posted via social networking sites (e.g., Facebook, Twitter, or Instagram), websites (e.g., Amazon, TripAdvisor, or Yelp!), or blogs (e.g., Facebook, Twitter, Sina Weibo), on the other hand, is created by the users as a result of the development of Web2.0. User-generated content, with its copious quantities and easily-collectable nature, also plays an important role in the electronic word of mouth, and hence becomes the focus of this study.

The structure of online data can be grouped into three types: structured, semi-structured, and unstructured. Structured data is formatted by data type (e.g., numeric, date) and has restricted input (e.g., limited length or restricted characters). Unstructured data such as texts is unformatted data that may contain dates, numbers, and facts at the same time. Semi-structured data is a type of data that contains semantic tags to describe the hierarchies within the data. For instance, user review is semi-structured data that contains several fields such as review title, comments, user rating, or date of posting. Among those field, user rating and date of posting could be structured data while the title and comments are usually unstructured texts.

Studies on user-generated online data cover a wide range of topics such as its perceived usefulness (Doh and Hwang , 2009; Willemsen et al., 2011; Racherla and Friske, 2012; Fang et al., 2016; Yan et al., 2016; Ismagilova et al., 2019), its impact on purchases (Van der Heijden et al., 2003; Ye et al., 2011; Bahtar and Muda, 2016; Erkan and Evans, 2018), the motivators of providing contents (Hennig-Thurau et al., 2004; Daugherty et al., 2008; Wang and Li, 2014; Poch and Martin, 2015; Crowston and Fagnot, 2018), its credibility (Wathen and Burkell, 2002; Cheung, 2006; Rabjohn et al., 2008; Ayeh et al., 2013; Metzger and Flanagin, 2013; Shan, 2016), or the retrieval of embedded sentiments (Simmons et al., 2011; Ravi and Ravi, 2015; Hu et al., 2017; Sun et al., 2019).

The perceived usefulness or helpfulness has positive influence on purchase decisions. Fang et al. (2016) found that text readability and the characteristics of the information provider affect the perceived value. Racherla and Friske (2012) suggest that the combination of the characteristics of the information provider and the receiver significantly contributes to the perceived usefulness. Erkan and Evans (2016) suggests that information quantity, information readiness, detailed information, and dedicated information are the four key factors that make anonymous reviews to be more influential than friends' recommendations.

As for the motivations to provide information, it is acknowledged that altruism, collectivism, and egoism significantly contributed to the intentions of consumers to post reviews (Cheung and Lee, 2012; Wang and Li, 2014; Poch and Martin, 2015; Yang, 2017). In addition, extrinsic rewards such as economic incentives (Poch and Martin, 2015) or perceived trustworthiness (Wang and Li, 2014) also contribute to the inclination to post reviews.

Moreover, creditability is the perceived trustworthiness and expertise of the information source by the information receiver (Hovland et al., 1953; Metzger and Flanagin, 2013). Metzger (2007) proposed five criteria to evaluate the creditability of online data, namely accuracy, authority (i.e. the credentials and qualifications of the author), objectivity (e.g., the purpose of providing the information), currency (i.e. how up-to-date the information is) and coverage or scope of the information and/or its source (i.e. the comprehensiveness or depth of the information). For instance, regarding authority, "top reviewers" whose credential is system-generated on a review website are recognized to have greater expertise and trustworthiness than laypeople by their peers (Shan, 2016).

Nevertheless, the utility of online data also has its limitations. For example, according to the survey of electoral perdition studies using Twitter data, Gayo-Avello (2013) points out the following challenges: 1) methods used in sentiment analysis fail to catch the subtleties, 2) the data (i.e. tweets) are assumed to be trustworthy when it is not the case, 3) demographics bias is neglected even when it is well known that social media is not a random sample of the population, 4) self-selection bias is simply ignored. For example, in the case of electoral perdition, tweets are more likely to be produced by the politically active people, and etc.

2.2.2 Online data and tourism analysis

There are various channels of travel-related data on the Internet. Table 2.1 shows several examples of travel-related data sources and their features.

Table 2.1 Examples of travel-related data sources and their features

Features	Travel website	Travel blog	SNS
Contents	Facility information, travel reviews, and Q & A	Travel reviews	Personal account: travel reviews, Official account: promotional information
Collecting Methods	Manual collection, data crawling		Manual collection, API
User Info.	Username, [address]		Username, language, [address]
Post Info.	Posting date, visiting place, [visiting date]		Posting date, [posting place]
Data type	Website structure, texts, [images, videos]		[texts, images, videos]
Volume	TripAdvisor: 490m unique users per month, 760m reviews, over 8.3m facilities (Jul. 2019) Expedia: max. of 236.5m unique user per month, 0.75m properties (Sep. 2018) Yelp: 37m mobile app, 77m mobile web, 62m desktop unique users per month, 192m reviews (Jun. 2019)	Travelerspoint: over 0.32m users Travel Blog: over 3,864 blogs TravelArk: over 2000 users (Oct. 2017)	Facebook: 1.47b users (Jul. 2018) Twitter: 326m active users per month, 500m tweets per day (Jun. 2018) Instagram: 1b active users per month, 500m active users per day (Mar. 2019) Line: 187m active users per month (Jun. 2019)

*Information in brackets is optional

For example, Twitter is usually an instant reflection of a user's comments in the form of short messages (Saito, 2011); thus, the instantaneity of tweets could be considered as an advantage in the aspect of tourism investigation. However, because tweets contain a wide range of topics about daily social life, it is difficult to isolate travel-related tweets (Nakajima et al., 2013). Besides, because tweets posted in a certain area could only be extracted with specified longitude and latitude coordinates, geographic meshing and address mapping are required to acquire tweets in a certain destination. It is also noteworthy that the location of posting the tweets is not necessarily equals to the location of the visited places, which introduces an error that cannot be precisely measured into facility-level investigation.

Meanwhile, travel-related websites provide travel blogs and travel reviews. On most travel websites, destinations and tourism facilities are classified according to the administrative districts and the reviews or blogs are related to each destinations or tourism facilities, enabling the identification of the destination (i.e. prefecture or city) and the facility. Further, compared to travel blogs, the quantity of data is larger on travel review websites. However, travel reviews, reflect a delay between the time of travel and the posting of the review. For example, the average delay between travel and posting for 1200 randomly selected reviews written in languages other than Japanese posted in Hokkaido from TripAdvisor was over 1.47 month. Still, over a half of the restaurant and attractions reviews were posted within a month after the travel.

Table 2.2 The average delay between the date of travel and the date of posting the review (month)

Category	**Mean**	**S.D.**	**Median**	***n***
Hotel	1.47	2.52	1	400
Restaurant	1.52	2.79	0	400
Attractions	2.06	3.20	0	400

Serna et al. (2013) found that Facebook and Twitter had more relevance for the public image projected by the destination, while TripAdvisor reviews were more about tourists' thoughts and feelings. Xiang (2017) compared the amount, length, ratings, and sentiments of hotel reviews collected from three travel websites (i.e. TripAdvisor, Expedia, and Yelp!) and found that even in the same destination, the content of the website and reviewers' behaviors vary across different platforms.

The literature in the field of tourism information technology has matured in late 1990s (Cser, 2008). According to Cser (2008), the *Journal of Information Technology and Tourism* (JITT) affiliated with the International Federation of Information Technology and Tourism (IFITT) and the *Journal of Hospitality Information Technology* (JHIT) affiliated are two of the early academic publications in this field. The latter ceased its publication in 2009, which makes JITT the major international publication in this field.

Taken JITT as an example, the focus of online data analysis was official/promotional tourism websites in 1999. Jung (1999) conducted an analysis of the demographic profiles of the users of an official tourism website. Weeks and Crouch (1999) analyzed the structure and contents of twenty Australian-based tourism and hospitality web sites. Procaccino and Miller (1999) compared altogether 345 websites of US-based firms and French-based firms. Holt (2002) compared Chinese language touristic websites and the English ones. Govers and Go (2004) analyzed the pictures and texts in twenty Dubai-based websites to study projected destination image. At that time, other discussed topics that may also relate to online data analysis are Internet-based destination management systems (Fesenmaier et al., 1999), questionnaire-based bookings prediction (Kliček, 2000; Morrisonn, 2001), personalized services (Loh, 2003; González et al., 2003; Gretzel et al., 2004), and etc.

The leading publications relevant to online social networks and online travel reviews in JITT started in 2008 (Chung and Buhalis, 2008; Yoo and Gretzel, 2008; Dippelreiter et al., 2008). For example, Yoo and Gretzel (2008) conducted an online survey to investigate what motivates consumers to write travel reviews. In 2009, Schmallegger and Carson compared destination images from blogs, review sites, and forums with destination-marketing-organization projected images. Zhang et al. (2009) investigated the sources (i.e. peer travelers, third parties, and travel companies) of online travel reviews and how can travel reviews be utilized for recommendations. Lee and Tussyadiah (2010) studied travelers' preferences for particular forms of information and found that textual data combined with images and/or videos have a greater influence on motivation to travel than text-only information. Dickinger and Költringer (2011) applied automated content analysis to thousands of blog entries to extract destination image to be compared with the images from a conventional image study. And their study suggests that blogs are a valuable additional source of information for destination marketing organizations. Inversini and Eynard (2011) employed the metadata of user-generated images collected from Flickr to decide the similarities between destinations. Jannach et al. (2014) analyzed the multi-criteria user ratings from hotel reviews on TripAdvisor and explored the utility of such data for the deployment of online recommendation services. Wang and Morais (2014) examined 69 blog entries posted by Chinese tourists about their experiences of visiting a Chinese minority group. Taylor et al. (2015) investigated the differences in the use of Twitter by two separate lodging segments (e.g., middle-class and luxury) and by Generation Y.

Concerns for online data tourism analysis in JITT kept rising since 2016. Huertas and Marine-Roig (2016) studied users' reaction to promotional contents on Facebook and found that the most identifying or destination-specific themes / attributes are the ones that trigger the most reactions. Dickinger and Lalicic (2016) compared responses of open-ended questionnaires about the perceptions of destination branding of Vienna, Austria with texts in 1,104 English travel reviews (about restaurants, attractions, and hotels) about actual travel experience on TripAdvisor. Dictionary-based text mining was applied to extract words in five categorizations namely competence (e.g., reliable, hard-working), excitement (e.g., daring, trendy), ruggedness (e.g., outdoorsy, masculine), sincerity (e.g., down-to-earth, family-oriented), and sophistication (e.g., upper-class, good-looking), and in six types of emotions namely anger, sadness, joy, fear, disgust, and surprise. They found that the proportions of words in each category used to describe the perceptions and actual experience are significantly different. Marine-Roig and Clavé (2016) proposed a semi-automatic method for downloading, arrangement, cleaning, debugging, and analysis of large-scale travel blogs and online travel data and used over 130,000 trip diaries to study the destination image of tourists who visited Catalonia between 2004 and 2014. Rossetti et al. (2016) proposed several topic models to analyze the contents of travel reviews and their application scenarios. Krawczyk and Xiang (2016) examined hotel reviews to find out how hotel brands are perceived. García-Pablos et al. (2016) used open source and free natural language processing tools to analyze hotel reviews collected from Zoover and HolidayCheck. Neidhardt et al. (2017) examine the sentiment scores in comments collected from one online review website. Amaro and Duarte (2017) studied the use of social media in Portugal and in the UK in different phase of travel. And so on.

The literatures of touristic online data analysis studies are also published in other tourism, computer science and management information system journals (Cser, 2008). For example, Tourism Management, as the leading international journal that covers a multitude of topics concerning the planning and management of travel and tourism, has been publishing researches that observed online behavior since 2002 (e.g., Wan, 2002; Beldona, 2005; Lee et al., 2006; Litvin et al., 2008; Kim et al., 2009; Qu and Lee, 2011) and that adopted user-generated online data analysis since 2009 (e.g., Vermeulen and Seegers, 2009; Xiang and Gretzel, 2010; Sparks and Browning, 2011; Lu and Stepchenkova, 2012; Chaves et al., 2012; Boley et al., 2013; Sparks et al., 2013; Vu et al., 2015; Tseng et al., 2015) and presented an increase in the quantity of such researches since 2016 (e.g., Fang et al., 2016; Zhang and Cole, 2016; Banerjee and Chua, 2016; Baka, 2016; Mariani et al. 2016; Plank, 2016). Other examples are Journal of Travel Research, Annals of Tourism Research, or Tourism Analysis.

As a newly emerged field of study, the variety of methodologies implemented in previous literatures is relatively limited. Common procedures of touristic online data analysis including data collection and storage, data cleaning, data sampling, data processing to add metadata in purpose to retrieve knowledge/information, and results aggregation and interpretation.

Data collection and storage is to gather a quantity of user-generated data in forms of texts, images and/or videos such as webpages that contain travel reviews or data entries that contain a variety of attributes such as username, postdate, or the urls of an image file, and then store these data in local or online data storage and management environment such as a database for future processing. As observed in previous studies, user-generated data can be collected via four types of methods: 1) manual scraping, 2) the use of API provided in certain platforms, 3) the use of data crawling tools, and 4) the use of official dataset provided by travel-related institutions or platforms. Considering the time and effort that would take to acquire the entire data within a particular platform, in most cases, collected data are generated during a specific period of time, in specific areas, and/or by a specific group of users.

Data cleaning, usually not a required procedure is to re-organize and remove noise data from the collected data. For example, when collect reviews in the form of webpage, it is necessary to delete metadata such as html tags and unrelated contents. Another example is to remove/replace certain punctuations such as a sequence of exclamations marks or emoji, in which case, dictionary-based removal / replacement is usually adopted.

Data sampling is to choose a set of data as representative of a large quantity of data that could take tremendous time to process. It is frequently used with manual analysis. In most case, random sampling is adopted, but the particular method of acquiring a random sample is barely clarified in the literature. Data sampling can also be achieved during the phase of data collection when choosing a specific destination or user groups instead of collecting the entire data. In the situation such as in destinations that only received a limited amount of reviewing, further sampling may seem unnecessary.

Data processing, as the most sensitive procedure, is to add extra information to the data or partials of the data so the data can be further categorized by certain criteria such as emotion (e.g., happiness, anger, sadness, or surprise) or sentiment (e.g., food, sights, or accommodations). And it can be achieved by two types of methods: 1) manual analysis and 2) automated analysis. Manual analysis could involve an individual, or a group of individuals, sometimes with one extra person as a tie breaker when the headcount is an even number. Automated analysis is conducted with one or a set of natural language processing or image identification programs / software / services. Further elaboration on manual and automated analysis of textual data will be presented in Chapter 4.

Results aggregation is to calculate the amount or proportion of annotated entries in each category to produce numbers for statistical tests and/or comparisons. And then by manual interpretation, knowledge / information that are useful to certain destinations, organizations and field of researches could be extracted.

2.3 Cross-region and cross-language tourism analysis

Cross-cultural analysis is one of the major research areas in tourism studies (Cohen et al., 2014; Bakir et al., 2017) focusing on the comparisons among tourists who are coming from various cultures started in late 1980s by Sheldon and Fox (1988) and Richardson and Crompton (1988). Tourists from different cultures are commonly segmented by nationalities, regions of residence, ethnicity groups, or language groups. This study takes the interests in cross-region and cross-language aspects.

Kim et al. (2002) concluded four techniques of conducting cross country comparisons: 1) comparing tourists of different nationalities through the eyes of tour guides, 2) comparison between tourist groups from different countries, 3) tourists-host comparison, and 4) organizational behavior in the hospitality industry. In the context of travel motivators, Soldatenko and Backer (2019) reviewed 71 publications from 1988 to 2017 and found that 40 studies conducted comparisons between Eastern and Western countries, 20 studies compared among Western countries, and seven studies among Asian countries.

Also, according to a meta-review by Li (2014), among the 91 surveyed publications, questionnaire surveys were the most widely used means of data collection, followed by secondary data, experiment and interview. Even nowadays, with the development of web2.0, questionnaire surveys (Wang et al., 2017; Özdemir and Yolal, 2017; Jung et al., 2018) and the use of secondary data (Lu and Chen, 2014; Huang and Crotts, 2019) are the two major approaches of data collecting. Meanwhile, the use of online data (Hatoh et al, 2013; Serna et al, 2013; Stepchenkova et al., 2014; Saeki et al, 2015; Kim & Stepchenkova, 2017; Liu et al., 2017) is increasing.

One major form of data is text such as comments in travel reviews or in tweets. However, despite of the fact that the quantity of such data is massive or influence of such data as information source for travel planning has increased, the approaches or techniques developed for cross-language comparisons are comparatively limited due to language equivalency problems. The most basic and simple solution is manual translation. For example, Hatoh et al. (2013) studied the difference between the viewpoints of Japanese and Chinese tourists using travel reviews from TripAdvisor. They used text mining techniques to compare the coherence of keywords with the highest occurrence frequency from Japanese and Chinese travel reviews, and then conducted manual comparison for sentimental analysis. Zhang and Li (2018) used manual classification of high frequency words in Chinese and English reviews. Meanwhile, with the recent development of machine-learning based natural language processing software, more and more researchers started to employ such instruments. For example, Nakayama and Wan (2019a; 2019b) used a tool named IBM Watson Explorer Content Analytics 11.0.1 for the comparisons of the top 50 sentiment expressions in Japanese and English restaurant reviews. Such an approach could be convenient, but extra works is needed to validate the precision and reliability of the conclusions.

It is also possible to avoid cross-language analysis when studying the needs of tourists from different countries by selecting data that are written in one commonly used language (Kim & Stepchenkova, 2017), by adopting quantitative indicators such as the length of the reviews (Mariani et al., 2018), or by using non-textual data such as photos (Stepchenkova et.al, 2015), geographical information (Saeki et al, 2015), or user rating (Liu et al., 2017).

All publications confirms that demographic profiles, tourist motives, perceptions of a destination, satisfaction levels, and tourist activities vary by culture (Li, 2014; Soldatenko and Backer, 2019). Because culture encompasses elements as shared values, beliefs, and norms (Li, 2014), it is also important to examine this topic from the perspective of social psychology, politics, economics, or etc. For example, from a psychological perspective, it is noteworthy that the normative system of emotional display rules also varies by culture (Fernández et al., 2000; Davis et al., 2012). That is, even when two groups of people possess similar internal emotion, how they express such emotion could be different via facial, body, verbal, and/or written languages, which bring the inequivalences into data analysis.

Chapter 3 The structure of travel website and content analysis of its travel reviews

3.1 Introduction

To show that travel reviews is a potential data source for tourism investigation, this study first undertook an investigation of the world's largest travel website, TripAdvisor. This chapter explains the structure of the website and what do people normally write in travel reviews with the help of both manual and automated analysis. Besides, for hints and insights of developing a method to analyses tourists with different geographical and linguistic backgrounds, this study analyzed and compared the contents of travel reviews posted by tourists from six different regions written in English, Japanese, and Chinese.

3.2 Case study: TripAdvisor

TripAdvisor, founded in 2000, is the largest travel website in the world for hotel booking, travel reviewing and etc. The structure of the website is under continuous revision. Table 3.1 shows its webpage structure until Nov. 2018.

3.2.1 Webpage classification

TripAdvisor contains information about several categories such as hotels, attractions, restaurants, flights, and the others. Because the focus of this study is travel reviews in each destination, this work only investigates the first three categories. Each of these three categories contains a list of facilities. Each page of facilities contains information about three sub categories: information about the facility provided by the owner, a list of travel reviews, and a list of Q&A. There are also a statistical summary of the reviews and options to refine the list by traveler rating (as in excellent, very good, average, poor, and terrible), by traveler type (as in families, couples, solo, business, friends), by season, by language, or by keywords.

Table 3.1 The webpage structure of TripAdvisor

Category	Sub category	Details
Hotels	Hotel Info.	Average Rating， the amount of reviews, address, amenities, hotel classification, and etc.
	Review	Reviewer name（username）, rating, posting date, review title, review text, [Google Translate button， period of stay, visiting date, photos]*
	Q&A	Username, text, posting date, [the amount of answers, username of the answerer, text of the answer, Google Translate button, helpful button]*
Things to do (Attractions)	Attraction Info.	Average Rating， the amount of reviews, Categorise of Attractions, Address, Opening hours, [Average period of stay, intro. Of the facility] *
	Review	id.
	Q&A	id.
Restaurants	Restaurant Info.	Average rating, the amount of reviews, budget range, category of food, address and etc.
	Review	id.
	Q&A	id.
Others categories: flights, forum, airline review, blogs		

*Inside of brackets is optional

3.2.2 Facility classification

The list of facilities can be refined by destination and by category. Destinations are determined according to the administrative districts in each country. For example, in ".jp", the categories are continent, country, district, prefecture, city (also including national parks or islands), and ward hierarchically. Categories use the multi-label classification and include two levels: category and tag. That is, each category contains multiple tags. Each facility was labeled with multiple tags. Therefore, one facility can belong to multiple categories. Table 3.2 is an example of categories and tags used on Attractions.

3.2.3 Review model

Travel reviews are consisted of the following information: username of the reviewer, review title, review rating in the scale of five, date of posting, comment, and in some cases, data of visiting and photos. Also, review written in foreign languages contains a Google Translate button. Whether the language is foreign or not is decided by country code top level domain of the website and the official language in each country. For example, in ".jp" (*i.e.* in Japan), language that is not Japanese is considered as foreign languages. Besides, hotel reviews also contain ratings towards certain sub-aspects such as service or cleanness, while restaurant reviews allow the rating towards service or food.

Table 3.2 Example of attractions categories and tags

Category	**Tag**
Sights & Landmarks	Historic sites, Sacred & Religious sites…
Nature Parks	Parks, Mountains…
Museums	Art Galleries, History Museums…
Shopping	Gift & Specialty Shops, Shopping Malls…
Spas & Wellness	Spas, Onsen Resorts…
Outdoor Activities	Golf Courses, Ski & Snowboard Area…

Table 3.3 Information about the user

Location	Details
Facility Page	username， [address]*， the amount of reviews， [the amount of users who click the Thank button]*
Mouse over the username	In addition to the above, user level, years of membership, the number of visited destinations, the number of posted photos, the number of reviewers by ratings, [customized badges (i.e. tags such as age group)]*
User Profile	In addition to the above, the number of followers, a list of posted reviews, badges, a map of the destinations based on the destinations from posted reviews

*Inside of brackets is optional

3.2.4 User model

Table 3.3 shows the information about a user on TripAdvisor. There are at least three locations that show user information, listed as followed: in the facility page next to the review, mouse over the username, and user profile page. In the first situation, the information is a brief summary about the user, including the username, customized profile picture, volunteered address, the amount of posts, and the number of users who think the review is helpful. More detailed information such as the distribution of ratings of the reviews written by this user or years of memberships can be found when mouse over or click the username or click his/her profile image. Further, clicking the username (presented as a link) will direct to the user profile page where the history of the user's reviews and the track of destinations that the user has been to can be found.

```
https://www.tripadvisor.jp/{Hotel | Attraction | Restaurant}-g{City id}(-Activities)-c{Facility category id}-t{Tag id}-{City name}_{prefecture name}.html
```

Fig. 3.1 The URL of the facility list under each destination

```
https://www.tripadvisor.jp/ShowUserReviews-g{city id}-d{facility id}-r{review id}-{facility name}-{city name}_{prefecture name}.html
```

Fig. 3.2 The URL of the review

3.2.5 URL structure

Pattern can be observed from the URLs of webpages on TripAdvisor. Fig. 3.1 and Fig. 3.2 are two examples. Each destination, category, tag, and review is assigned with a distinctive number. Destination id starts with the letter g, category c. tag t, and review r. Also, the English names of the facilities, the destinations can be isolated from the urls of a facility.

3.3 Methodology

This section explains the method used to analyze reviews, which includes two main steps: collection and categorization of reviews, and categorization of review contents. The first step explains the data that is needed for the cross-region comparison. In the second step, two approaches (manual analysis and text-mining techniques) are used. The manual analysis session explains a content categorization model, developed from a pre-analysis of contents of 100 randomly selected reviews. The text-mining section explains a semantic categorization model for nouns.

3.3.1 Data collection

Ideally, target data should be travel-related texts that are: 1) posted by people from different region for a single attraction allowing cross-region comparison, 2) posted in various types of attraction spots, considering tourists’ different preferences in selecting destinations, 3) posted in attraction spots located in different regions, considering tourists' behaviour may differ between domestic and overseas travel.

Regions and attractions are selected as follows: regions to be studied are chosen by the following two standards: 1) regions that vary with regards to social culture and geographic conditions, 2) regions whose native language include English, Chinese or Japanese. As a result, United States, Australia, Great Britain, China, Japan and Singapore are selected.

There are over 23.1 million reviews posted in the six selected countries on TripAdvisor. Ideally, all reviews should be collected and analysed, but due to the experimental constraint, to study attractions we narrowed them down to those that 1) are located in the top 20 cities with the highest amount of attractions in each country except for Singapore, 2) belong to the 7 attraction types that are frequently investigated in national travel surveys and that also exist in TripAdvisor, namely Sights & Landmarks, Nature & Parks, Shopping, Museums, Zoos & Aquariums, Water & Amusement Parks, and Food & Drink, 3) have are among the top 30 with regards the amount of reviews for each attraction type in every city, and 4) have more than a hundred reviews.

Finally, the latest 1,000 reviews, a sampling size with a confidence level of 95% and an error of approximately 3%, are collected in each language (English, Chinese and Japanese) from each attraction.

3.3.2 Data categorization

The reviews need to be categorized according to the region of a tourist. On TripAdvisor, a user's location information is provided along with the review when the user has filled in this information. However, location is a type of unstructured textual address that may or may not include the name of the region of a location. Therefore, we first need to extract the region of a location from the address. Fig. 3.3 shows the steps taken in this study.

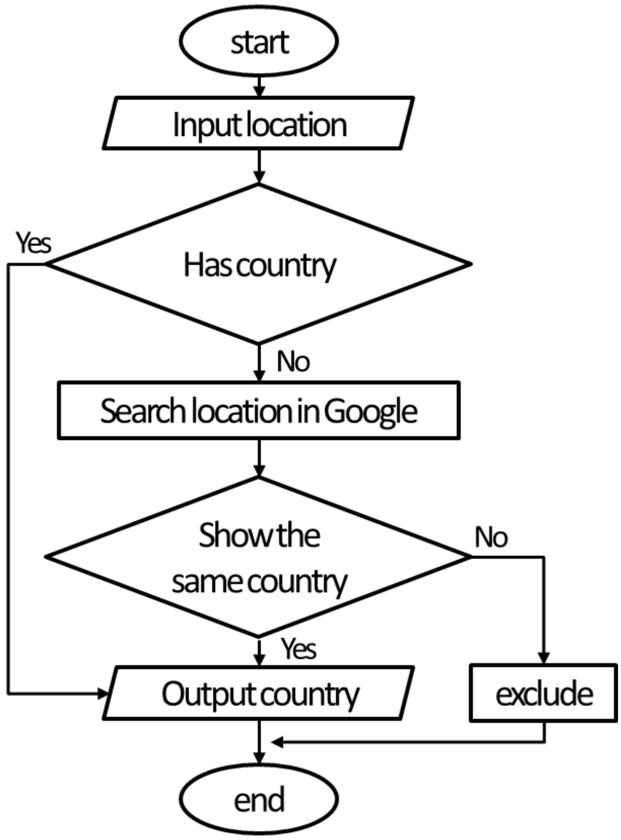

Fig. 3.3 Steps to extract country info. from the location

Ideally, the region of a tourist should be the region where he/she was raised, received an education, or has been living for a very long time to have resulted in an influence on his/her preference and attitudes from a social, culture or geographic perspective. However, TripAdvisor only provides a user's current location information. Thus, as shown in Fig. 3.4, to rule out those who move abroad for a middle or long term stay for the purpose of a job, education, or others; a tourist's region will be the region of the location the tourist claims to be only if the native language of the region of the location is consistent with the language used in the review. For example, if a tourist filled the location information as Beijing, China, he/she will be categorized as a Chinese if he/she writes reviews in Chinese.

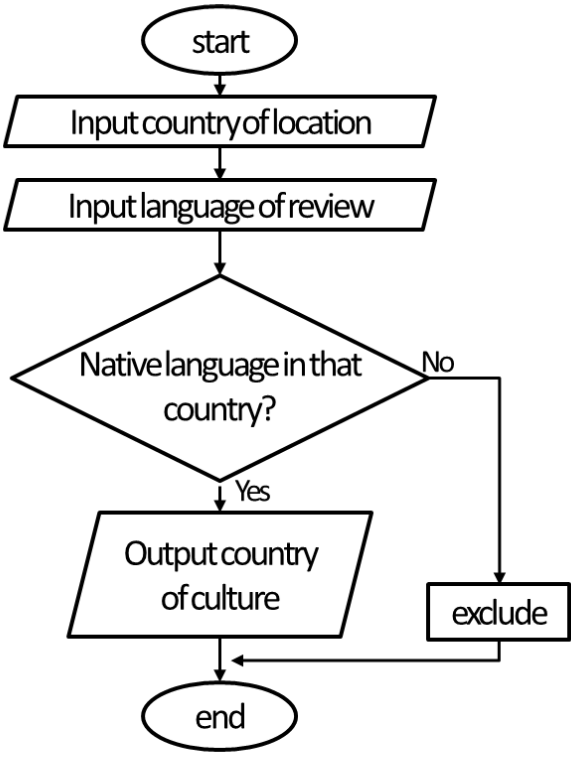

Fig. 3.4 Steps of Review categorization

Data was collected from 1/2/2017 to 4/8/2017 using a data crawling tool. As a result, approx. 290,000 English reviews posted by American tourists, 210,000 English reviews by British, 110,000 English reviews by Australian, 16,000 English reviews by Singaporean, 170,000 Japanese reviews by Japanese and 90,000 Chinese reviews by Chinese were collected.

Table 3.4 Reviews in dataset 1

Destinations	American	Australian	British	Chinese	Japanese	Singaporean	Total
U.S.	69	7	30	3	32	0	141
Australia	8	53	10	6	18	5	100
Great Britain	37	18	88	13	28	3	187
China	28	32	17	50	26	27	180
Japan	55	39	33	93	87	53	360
Singapore	27	63	50	34	46	112	332
Total	224	212	228	199	237	200	1300

Three Data Sets are created from the collected data for the following purpose:

- Dataset 1 for manual content analysis: 1,300 reviews randomly selected from 6,000 reviews (1,000 reviews per region) randomly selected from 69 attractions with the most reviews in the six selected regions, covering various attraction types namely parks, religious locations, historic locations, museums, observations, zoos, botanic gardens, shopping areas, bridges, water areas, mountains, and amusement parks.

- Dataset 2 for pre-analysis of the content of reviews: 100 randomly selected reviews from Dataset 1.

- Dataset 3 for automated analysis: 30,000 reviews (5,000 reviews per region), randomly selected from the collected data.

3.3.3 Content categorization model: a pre-analysis

Textual data is unstructured data that need to be grouped into categories to enable knowledge to be extracted (Marine-Roig & Clavé, 2015). Categories can be pre-established or discerned from the text itself (Stepchenkova, 2012). Nakajima and Ohta (2013) identified the content of travel blogs as tourist spots information (name, location and introduction of attractions), their related information (tourists' actions, tourists' emotions, tourists' description of the attraction, and origin or background knowledge of the attraction) and unrelated information (a third-person's experience, tourists' past experience, and information of unrelated attractions). Kurashima and Ukuda (2009) classify tourists' experience into location, time, activity and emotion. Nakajima et al. (2013) categorized Twitter content into food, view and action. Their categories are later expanded to be applied to tourists' actions (shopping, eating, experiencing, staying and seeing) (cited in Iinuma et al., 2017).

However, there is no categorization model to describe the contents of textual data from the perspective of tourism investigations. Therefore, in this session, a pre-analysis is conducted to generate such a model using Dataset 2. It includes the following two steps.

STEP 1 - content extraction. Content is identified phrase by phrase and is named using expressions adopted from primary studies and surveys. Taking the review shown in Fig. 3.5 as an example, three types of content (visit date/time, see, and comment) can be extracted. Results of the content of 100 reviews are shown in table 3.5.

STEP 2 - content categorization. Content is categorized into 10 sub-categories and further into 3 main categories using KJ method, an affinity diagram helps to categorize and organize a large number of fragmented uncertain information into logical cohesive groups (Kawakita, 1986). During the process, categories from existing studies (Nakajima & Ohta, 2013; Iinuma et al., 2017) and traditional surveys are also taken into consideration. Explanations of the 3 main categories are as follows:

We visited at night (*visit date/time*), what a great atmosphere (*comment: good atmosphere*) and prices were pretty reasonable (*comment: good price*). Lots to see and a big range of different restaurants to choose from (*see: restaurant*). Not a 'touristy' as we'd expected, which was good (*comment: good not touristy*).

Fig. 3.5 Example of the content analysis of a review

Table 3.5 Content of 100 reviews

Category	Sub Category	Content	Review Amount
Information of the attractions	Introduction of the attraction	Resource description	43
		Background	16
		Location	13
		Ticket	9
		Access method	8
		Shop	5
		Price	9
		Atmosphere	4
		Activity	2
		Open/Close time	2
	Introduction of Surrounding attractions	Surrounding attractions	2
		Surrounding events	2
		Surrounding facilities	1
	Recommendation	Recommend point	44
		Watch out issues	10
		Other attractions	8
		Photo point	2
	Information about the people	People	16
		Waiting duration/Queue	5
		Staff/ Guide	3

Table 3.5 Content of 100 reviews (Continued)

Information of the tourist	Past experience	Times of Visits	4
		Experience from past visits	2
		Personal history	1
	Travel background	Visit date/time	15
		Purpose/ motivation	7
		Weather	6
		Companion	5
		Stay duration	4
		Visit form	3
	See	-	10
	Do	Take photo	10
		Eat/ Drink	9
		Attend activity	8
		Walk around	4
		Accidental event	1
		Shopping	1
Comment	On this travel	-	72
	On next travel	-	3

Information about the attractions: Descriptive content about an attraction or otherwise related attractions. This category also includes recommendations and information about people in/around the attraction.

Information about the tourist: Personal experience before and during the travel. To distinguish personal experience from descriptive content, reviews with information about visiting date or weather, or reviews involving expressions such as “we saw …”, “we went there with our friends …” are counted.

Comment: Tourists’ positive or negative feelings and opinions towards content involved in the information of an attraction or the tourist.

Altogether 1,300 reviews (Dataset1) are manually analyzed. The content of these reviews is transformed into categories accordingly.

3.3.4 Semantic categorization and text mining

Text-mining techniques are used as a second approach of content analysis to confirm the conclusions acquired using manual analysis. The general idea here is to firstly break a text into separated words using a morphological parsing technique. Next one finds the most written words, then these words are categorized based on their semantics, and finally the amount of reviews included in each semantic category is calculated.

It is worthwhile to point out that the result of morphological parsing depends on the dictionary embedded in different tools (Murakami and Kawamura, 2011). Thus, to process reviews written in different language fairly, it is important to use the same tool (esp. the same dictionary). Besides, among the various kinds of parts of speech, nouns are considered as the most informative ones in search for content in reviews (Liu et al., 2017). Therefore, the process of semantic categorization can be listed as the following 3 steps. The amount of reviews included in each semantic category will be calculated.

1) Review translation: Chinese and Japanese reviews are translated into English using Google Translation (see translate.google.co.jp) to be processed by the same dictionary.

2) Morphological parsing. A tool called TreeTagger is used to annotate a text and each word with part-of-speech and lemma information (Schmid, 1994; Schmid, 1995).

3) Semantic categorization of nouns. A variety of 1,420 nouns which occurred more than 50 times in Dataset 3 is manually categorized according to their semantics, as shown in table 3.6.

Table 3.6 Categorization of nouns

Semantics	Variety of nouns
View(nature/ artificial)	269
Culture	92
Food	73
Access/transportation	62
Activity	58
Shopping	47
Atmosphere	40
Infrastructure	38
Price	22
Service/staff	15
Safety	6
Sanitary	2
Not travel related	691
Total	1,420

3.4 Results and findings

3.4.1 Results of manual analysis

This session shows the results of the manual analysis of 1,300 reviews. Table 3.7 shows the amount of reviews with content included in each main category. Information about the tourist and comments are considered as useful information for tourism investigation. When the numbers are divided by the total review amount in each column, we can see that over 34% of reviews contain useful information. Meanwhile, the percentage differs among regions. Thus, the smaller the percentage, the bigger the sample size should be in a sampling study.

Table 3.7 Amount of reviews with content included in each main category

Category	**US**	**AU**	**GB**	**CN**	**JP**	**SG**
1.Information about the attractions	186 (83%)	176 (83%)	204 (86%)	165 (83%)	178 (78%)	182 (91%)
2.Information about the tourist	87 (39%)	87 (41%)	92 (39%)	80 (40%)	137 (60%)	68 (34%)
3.Comment	177 (79%)	180 (85%)	204 (86%)	98 (49%)	119 (52%)	120 (60%)
Total reviews	224	212	228	199	237	200

US: American, AU: Australian, GB: British

CN: Chinese, JP: Japanese, SG: Singaporean

Table 3.8 Percentage of reviews with content included in the 10 sub categories

Sub category	**US**	**AU**	**GB**	**CN**	**JP**	**SG**
1.Introduction of the attraction	72%	67%	72%	69%	58%	80%
2.Introduction of Surrounding attractions	7%	9%	7%	11%	8%	9%
3.Recommendation	26%	32%	31%	37%	20%	39%
4.Information about the people	21%	20%	21%	19%	22%	12%
5.Past experience	2%	3%	3%	1%	5%	7%
6.Travel background	23%	21%	20%	19%	25%	20%
7.See	25%	31%	26%	31%	44%	18%
8.Do	13%	11%	11%	11%	18%	11%
9.Comment on this travel	78%	84%	86%	47%	51%	58%
10. Comment on next travel	2%	4%	3%	4%	2%	4%

Table 3.8 shows that percentage (amount of reviews divided by the total review amount in each column in table 3.7) of reviews with content included in each sub category. Less than 7% of the reviews contain information about tourists' past experience and comments on the next travel. Detailed statistics of sub categories 6, 7, 8 and 9 are shown below.

Table 3.9 Amount of reviews with travel background

Travel background	US	AU	GB	CN	JP	SG
Visit date/time	23	21	20	15	27	18
Companion (family/couple/friends)	10	12	17	8	14	10
Weather	16	10	9	8	24	9
Stay duration	5	8	7	4	0	10
Travel formality	4	5	2	2	0	1

Table 3.10 Amount of reviews with things the tourist saw

see	US	AU	GB	CN	JP	SG
tree/flower/garden/plant/lawn	15	20	17	15	25	13
animal/fish/birds	3	8	3	7	8	3
Night view/lights	3	3	2	6	10	2
building/architecture/design	5	4	3	3	6	3
exhibits	6	1	2	2	6	3
lake	1	6	3	0	4	3
sea/beach	2	0	2	7	3	1
waterfall	1	4	1	1	4	1

Table 3.11 Amount of reviews with activities the tourist attended

do	US	AU	GB	CN	JP	SG
walking/jogging/cycling/hiking/climbing	10	6	4	2	14	8
take photo	5	2	3	10	10	2
boat trip	2	4	9	4	5	7
food	5	8	10	7	9	5
shopping	2	3	1	1	0	0
guide tour	3	0	1	0	1	0

Table 3.9 shows the amount of reviews with travel background that can be useful for fact-finding investigation. Also, among the 1,300 reviews, only 33 reviews (2.5%) have clearly identified their travel purposes or motivations, such as “nearby the hotel” (6 reviews), “because it is famous” (3 reviews), “was recommended by a friend” (2 reviews), celebration of a birthday, or desire to see a certain view or show.

Table 3.10 shows the amount of reviews about things the tourist saw, which can be useful for investigating tourists’ viewpoints. Of 489 reviews about things the tourists saw, only 72 reviews (14.7%) use specific words instead of general words, such as azaleas instead of flower, cherry blossom instead of tree.

Table 3.11 shows the amount of reviews with things the tourist did, which can be useful for investigating tourists’ actions. American and Japanese tended to write about doing more exercises. Also, Chinese and Japanese wrote about taking photos.

Table 3.12 shows the most written comments, which can be useful for analyzing tourists’ complaints and compliments. Comments on views are the most, followed by general comments about the whole travel experience (worth a visit or enjoyable). American, Australian and British tourists gave comments on the staff and guide. American and British cared if the culture is interesting or not. Australian and Singaporean commented more on food. British tourists commented more on price.

Table 3.12 Amount of reviews with the most written comments on this travel

Comments	**US**	**AU**	**GB**	**CN**	**JP**	**SG**
good view	94	92	76	35	55	62
worth a visit	14	22	37	17	7	13
must visit	19	31	17	8	6	15
enjoyable	17	18	17	6	26	8
good activity	14	13	16	5	8	6
good staff	10	7	10	1	2	0
good culture	10	4	9	0	1	4
good food	3	7	5	3	1	8
good access	7	4	1	3	2	2
good infra	6	3	6	1	0	2
nothing special	2	4	2	3	1	3
bad crowds	3	2	2	3	2	3
bad price	0	2	7	1	1	1
good price	1	4	4	1	1	1
bad infra	4	2	3	1	1	0
good shop	2	5	1	1	0	1
good guide	4	1	4	0	0	0
good building	2	0	3	0	2	1
worthy(money)	2	0	5	0	0	0

Table 3.13 Amount of reviews with different semantics divided by 5,000

Semantics	**US**	**AU**	**GB**	**CN**	**JP**	**SG**
view	84%	83%	81%	93%	88%	83%
activity	54%	57%	58%	38%	37%	52%
culture	46%	38%	41%	47%	32%	42%
access	29%	32%	27%	31%	36%	36%
food	31%	40%	35%	21%	28%	37%
shopping	24%	27%	25%	21%	33%	34%
atmosphere	23%	19%	20%	31%	23%	20%
infrastructure	20%	20%	23%	10%	16%	19%
price	15%	16%	19%	17%	15%	19%
service	9%	12%	15%	3%	5%	9%
safety	2%	1%	2%	3%	2%	2%
sanitary	0%	0%	0%	0%	0%	1%

3.4.2 Results of text-mining

Table 3.13 shows the results of text-mining. There are results consistent with the manual analysis. For example, most reviews contain words about the views and activities. Americans wrote more about cultures; Australian and Singaporean wrote more about food; British wrote more about price and services, and little about access; Chinese wrote more about views and less about activities. However, it is difficult to exclude the descriptive information from the needs-related information (information of the tourist and comments) without further analysis.

3.5 Conclusion

In this chapter, contents of travel reviews collected from TripAdvisor posted by tourists from six countries written in English, Japanese, or Chinese, are analyzed. During the manual analysis, a content categorization model with three main categories and ten sub categories (see table 3.5) is developed based on the KJ method. By transforming review content into these categories, information that is useful for finding travel facts, tourists' viewpoints, actions, complaints and compliments can be identified apart from descriptive information. Over 34% of reviews contain information of the tourist and comments, which can be useful for extracting tourists' needs. This percentage is higher than the average response rate of questionnaire (i.e. 30%) according to the statistics provided by Research Works (2017). This result suggests that travel reviews from TripAdvisor can serve as a possible data source for tourism investigation. Furthermore, since the percentage of useful information differs between regions, sample size should be selected accordingly. Moreover, in the automated analysis based on basic text-mining techniques, results consistent with manual analysis are found. This finding also suggests that it is possible to use automated analysis instead of manual analysis.

Limitations. Reviews used in this research are limited to six regions and three languages. Also, reviews are collected from a single data source. Therefore, the results may not be able to be generalized to other data sources. Besides, reviews are clustered by the region of a tourist in this research, while other factors such as tourists' age, gender or occupation, travel season or destination, activities taken part in and experience can also be used for clustering. Moreover, the amount of reviews analyzed using the manual analysis is limited. As a result, the overall amount of reviews is rather small when looking at detailed content and each country which prohibits from providing a representative conclusion concerning tourists' needs. Automated content analysis is needed to process massive data. But basic text-mining techniques used in this research are unable to distinguish tourism-investigation-related information from descriptive information.

Future research. To enable the analysis of a huge amount of reviews, a method that automatically transforms the content of reviews into the three main categories and the ten sub categories is needed. Meanwhile, information about the attraction is considered as unrelated information for investigating tourists, but it may be helpful for the analysis of the characteristics of the destinations.

Chapter 4 Sentiment analysis for investigating tourist satisfaction

4.1 Introduction

Based on the analysis of the comments in 1,300 online travel reviews to determine what people write in their reviews in Chapter 3. This study found that over 49% of the analyzed reviews contained information about tourist satisfaction, over 34% contained travel facts, and 2.5% contained travel intentions. From these results, it appeared that travel reviews on TripAdvisor can be considered a potential data source for investigating tourist needs.

However, the analysis of online data to investigate tourist needs introduces external variables such as Internet use (Ferrer-Rosell et al., 2017) and the choice of online platform (Xiang, 2017). Therefore, although conclusions from data mining can be suggestive, the conclusions yielded by such data can accurately represent tourist needs is just an assumption.

This chapter aimed to determine the validity of identifying tourist needs from travel reviews through the use of text data mining. Travel reviews from TripAdvisor were used as an example of online textual data and employed the results of a traditional sampling survey as comparison data. The authors developed a method to extract attitudes from textual comments in online travel reviews for comparison with the results of the traditional survey. Because of the high possibility of finding information on tourist satisfaction, the main concern in this chapter was satisfaction. Also, this study was focused on Hokkaido, one of the best-known tourist destinations in Japan. For comparison, a guest survey implemented by the Hokkaido government, the Survey Concerning Customer Satisfaction (2016) was used. The three main tasks in this chapter were the following.

1) Creating a set of rules to extract answers to survey questions from travel reviews.
2) Creating a method to compare aggregate answers from reviews to those of a traditional survey.
3) Exploring the differences and similarities between reviews written in different languages.

4.2 Tourist satisfaction

The analysis of tourist satisfaction is a rich area of research. Tourist satisfaction is a cognitive and emotional reflection of tourist attitudes toward tourism service (Bowen & Clarke, 2002) and can be affected by many variables such as expectation and performance (Oliver, 1980; Aksu et al., 2010; Berezina et al., 2016). For example, tourists can be dissatisfied if expected components of service are not provided or are improperly delivered. Nevertheless, Aksu et al. (2010) found a strong positive correlation between pre-trip expectations and tourist satisfaction. Moreover, identification of satisfied and unsatisfied tourists positively impacts tourism development. For example, higher satisfaction leads to higher revisit intentions (Omar et al, 2017) and more word-of-mouth recommendations. Conversely, dissatisfied tourists help identify problematic aspects of tourism institutions (Berezina et al., 2016).

Tourist satisfaction has been investigated through tourism surveys or through the analysis of online data available in forms such as travel reviews. Below, section 4.2.1 discusses the advantages and limitations of traditional survey methods. And section 4.2.2 introduces travel reviews with a discussion of their advantages and disadvantages and shows research that has used travel reviews for measuring satisfaction.

4.2.1 Traditional tourism surveys

The most common traditional survey methods include face-to-face interviews and self-administered surveys such as mail surveys and Internet surveys. Face-to-face interviews provide more truthful personal information because the answers can be screened by interviewers (Opdenakker, 2006). The presence of interviewers can increase the inclination of respondents to cooperate, but it can also introduce interviewer-related errors attributable to specific interviewer characteristics such as race or gender (Fowler & Mangione, 1990). Also, the personnel cost of such interviews is the highest of all survey methods. As illustrated in Chapter 2, interview-based surveys usually take about three months to a year. Interviews mostly take place at the destination; thus, sample size will be probabilistically limited if the arrivals of the tourists from a certain overseas area is limited. Meanwhile, mail and Internet surveys can reach more respondents from more areas in a comparatively shorter period of time. However, producing and distributing questionnaires and incentives can be costly. Another problem they face is their low response rate. According to Research Works (2017), the response rate for mail surveys averages about 30%. Shibutani et al. (2015) reported on a case study involving a single sample in which the response rate for mail surveys was about 50%, while that for Internet survey was about 22%.

Apart from the implementation challenges and limitations outlined above, the exact manner in which the human mind reacts to a question and produces an answer is also extremely complex, as Tourangeau et al. (2002) have indicated. From the psychology point of view, these authors divided the process of answering a survey question into the following four major components: comprehension, retrieval, judgment, and response. The first component, comprehension, is the process of understating the survey question. At this stage, problems such as semantic difficulties due to the use of ambiguous words or because of the respondent’s unfamiliarity with the topic may arise. The second component, retrieval, is the process of organizing information from relevant memories. The accuracy of answers to survey questions is hence influenced by the structure of memory and how it is searched. The third component, judgment, can be affected by contexts such as question order or the language used in other questions. The final component, response, is the mapping of judgments to answers that are allowed on the survey form. During this process, responses may be prejudiced by factors such as the social desirability bias or the extreme response bias. The social desirability bias is a tendency of respondents to hold back negative opinions in order to present themselves as socially desirable (Fisher, 1993). The extreme response bias is a predisposition to only select the most extreme options even if respondents do not actually maintain an immoderate opinion (Yüksel, 2017).

Because each destination has its own local specialties and characteristics, their surveys include different questions despite sharing a common topic—tourist satisfaction. Furthermore, they lack standardized guidelines for measurement. For example, in the Consumption Trend Survey, the Japan National Tourist Organization (JTA, 2019b) composed 20 questions on tourist satisfaction, some of which were destination-specific (e.g., bathing in a hot spring, staying in a Japanese-style inn, or drinking Japanese alcoholic beverages). These questions could be answered with two options, satisfied or not satisfied. Another example is the Survey Concerning Customer Satisfaction by the Hokkaido government (2016). This survey was carried out from Jun. 1st, 2016 to Feb. 28th, 2017. In total, 1,709 participants were randomly selected and interviewed at airports, docks, or information centers in Hokkaido. The interviewees were international tourists from China, Taiwan, Korea, Hong Kong, Thailand, Malaysia, Singapore, America, Australia, Europe, and other regions. The Hokkaido government composed 11 questions to investigate how satisfied the tourists were. The government asked the participants to answer 11 questions by selecting one of seven options: (1) very satisfied, (2) satisfied, (3) fairly satisfied, (4) it was ok, (5) not very satisfied, (6) not satisfied, and (7) did not use. The government then calculated regional satisfaction rates by adding up the numbers of the participants from each region who answered (1), (2), or (3) (as shown in table.4.1). For example, of the 628 participants coming from Taiwan, 94.8% were considered satisfied with their meals at each tourist destination in Hokkaido.

Table 4.1 Satisfaction rate from the guest survey (%)

#	Items	TW	CN	HK	SG	AU	US	EU
1	for the entire trip and sightseeing	95.9	95.7	94.7	96.6	91.7	89.5	95.2
2	meals at each tourist destination	94.8	93.5	96.3	96.6	92.0	100	90.5
3	souvenirs	92.7	93.4	91.8	89.3	58.3	52.6	61.9
4	accommodations	95.2	92.7	92.5	96.3	92.0	94.7	81.0
5	tourist attractions	95.3	94.2	96.2	93.1	80.0	84.2	76.2
6	Wi-Fi accessibility	75.1	80.9	76.9	83.3	88.0	78.9	76.2
7	multilingual informational signs at tourist destinations	83.2	84.5	84.2	90.0	87.0	84.2	71.4
8	local staff's linguistic abilities	80.3	76.8	79.9	86.7	88.0	78.9	66.7
9	transportation system	93.8	94.0	96.2	76.7	88.0	89.5	81.0
10	customer service	97.2	96.2	97.0	93.1	88.0	100	95.2
11	scenery	98.1	95.6	95.5	96.7	92.0	100	95.2
n	628	411	148	32	27	22	27	27

Original source: Hokkaido Government, 2016;
TW: Taiwan, CN: mainland China, HK: Hong Kong, SG: Singapore,
AU: Australia, US: America, EU: Europe

4.2.2 Online travel reviews

In general, the utility of online reviews still faces many challenges. The first concern is the under-reporting bias, which signifies that consumers who are greatly satisfied or intensely dissatisfied are more likely to post reviews. Koh (2011) also suggested that the under-reporting bias may vary across cultures by evincing that Chinese online reviews reflected a film’s quality as perceived by the general population more accurately than reviews in the U.S. This concern leads to the necessity of examining consumer motives in providing reviews. In the context of restaurant review websites, Cheung and Lee (2012) found that altruism, collectivism, and egoism significantly contributed to the intentions of consumers to post reviews. In addition, Yang (2017) suggested that individuals who were more incentivized to help other users or to support the company were more inclined to provide reviews. Meanwhile, Fu et al. (2015) examined the motivation of posting positive and negative reviews about the online shopping experience. They suggested that consumers who intended to post positive reviews were more driven by underlying attitudinal factors, whereas those who intended to post negative reviews were more driven by social pressure.

Another concern is the credibility of online reviews. Lee et al. (2011) suggested that active users who provide their age, gender, and location information tend to be more trustworthy. Kusumasondjaja et al. (2012) suggested that negative reviews are more credible than positive reviews. Another problem of travel reviews is the acquisition bias. In fact, this leaning is also a problem that presents in traditional surveys. The acquisition bias is a condition that only consumers who have a favorable attitude toward a product will acquire said product (Koh, 2011). This situation creates a bias toward a greater number of positive reviews or responses.

When using online travel reviews to measure satisfaction, several approaches have been used to identify satisfied and unsatisfied reviewers. For example, Berezina et al. (2016) identified satisfied and unsatisfied reviewers according to whether the reviewer recommended the property to others. In addition, comments expressing positive emotions were used to indicate satisfaction while those expressing negative emotions were employed to denote dissatisfaction (Zhou et al., 2014; Shao et al., 2017). This approach is supported by the outcome of the study conducted by Xiang et al. (2015), who examined the associations between user ratings and guest experiences in hotel reviews and found that customers tend to use particular words to describe their experiences when they are happy or unhappy about the hotel.

The use of travel reviews introduces many external variables to the measurement of satisfaction. Ferrer-Rosell et al. (2017) examined the relationship between actual pre-trip Internet use and overall trip satisfaction using survey responses provided by an official travel agency. They suggested that pre-trip Internet use would not affect overall trip satisfaction. On the other hand, reviewers' behaviors vary across different platforms (Xiang, 2017). In addition, the expression of satisfaction varies by language. For example, Antonio et al. (2018) applied data analysis tools to English, Spanish, and Portuguese reviews. They suggest that reviews written in English reflected higher ratings than the others.

4.3 Sentiment analysis

Sentiment analysis, also known as opinion mining, is a thriving research area that aims to extract attitudes from given documents (Pang & Lee, 2008). Sentiment analysis can be performed on different levels: document level, sentence level, or topic level. Document-level analysis examines the overall attitude shown in the reviews, in which the 5-point scale for ratings can be a useful indicator of the overall attitude (Berezina et al., 2016). However, one review may contain information on several topics and reviewers may show different attitudes toward each of these topics. For example, in a hotel review, the reviewer may be satisfied with the amenities, but dissatisfied with the service. Extracting such information requires the reviews to be examined at a finer-grained level (Lu et al., 2011).

Topic-level analysis addresses two tasks (Schouten & Frasincar, 2016): 1) topic extraction and 2) sentiment classification.

Topics can be either generated from the reviews or pre-defined. The 11 questions from the guest survey, for example, were 11 pre-defined topics of tourist satisfaction. Although TripAdvisor provides rating data for some topics, because the questions were assigned by survey conductors, the presence of corresponding rating data was not guaranteed. Consequently, we need to extract the underlying attitudes toward each topic from the textual comments.

Sentiments are usually classified by polarities (positive and negative). In addition to positive and negative sentiments, a neutral category is usually included in reviews to indicate equally positive and negative sentiments (Pang & Lee, 2008), or simply a factual account (i.e., description or explanation) of a place or a facility (Murakami & Kawamura, 2013). It is interesting that although facts are not opinions, the behavior of mentioning facts can be emotional (Cabral & Hortacsu, 2010). Conversely, sentiments can be classified from the perspective of human emotions (Plutchik, 1960; Parrott, 2001). For example, Plutchik (1960) divides human emotions into eight categories: *Joy*, *Trust*, *Fear*, *Surprise*, *Sadness*, *Disgust*, *Anger*, and *Anticipation*. Human emotions can be further categorized into *positive*, *neutral*, *negative*, and *ambiguous* emotions (Strapparava & Valitutti, 2004). For example, Kurashima et al. (2009) suggest that *joy* is a positive emotion, *surprise* cannot be identified as either positive or a negative without context, and the others (i.e., *confusion*, *anger*, *sadness*, *tiredness*, *fear*, and *dissatisfaction*) are negative emotions.

Automated analysis can generate a general statistical description of a large amount of data in a very short time. Many algorithms, from linguistic approaches to machine-learning-based approaches, have been proposed for topic-level sentiment analysis. According to Schouten and Frasinca's survey (2016), many researchers have used accuracy, precision, recall, and F_1 value as indicators to measure quantitative performance, where the values of those indicators can be promising. However, many algorithms only apply to data in specific domains (e.g., movie reviews, financial news, or product reviews, for instance). Nevertheless, the results of most machine-learning-based approaches are difficult to duplicate because the performance of the classifiers largely depend on unavailable training data.

Meanwhile, manual analysis can generate specific and detailed results, as well as provide explanations that help interpret the results of automated analysis (Murakami & Kawamura, 2013). However, manual analysis may introduce error and bias, which could result in disagreements in outcomes (Pang et al., 2002; Tokuhisa et al., 2015). Therefore, guidelines are needed for manual analysis. For example, in the Gold Standard, one record will be analyzed by two or more expert annotators and will be adopted as high-quality labeled data when independent annotators reach some level of agreement (Petrillo & Baycroft, 2010). However, the creation of a gold standard dataset is costly. To reduce the cost, Wissler et al. (2014) suggested several possible methods such as reducing the number of annotation tags and using non-expert annotators. Ukpabi et al. (2018) used computer-assisted manual analysis when conflicts between two annotators were flagged by the program, and a third annotator acted as the tie-breaker.

4.4 Methodology

This section explains the methodology employed to use sentiment analysis to relate the attitudes expressed in travel reviews with the tourist satisfaction recorded in the guest survey as shown in Fig. 4.1.

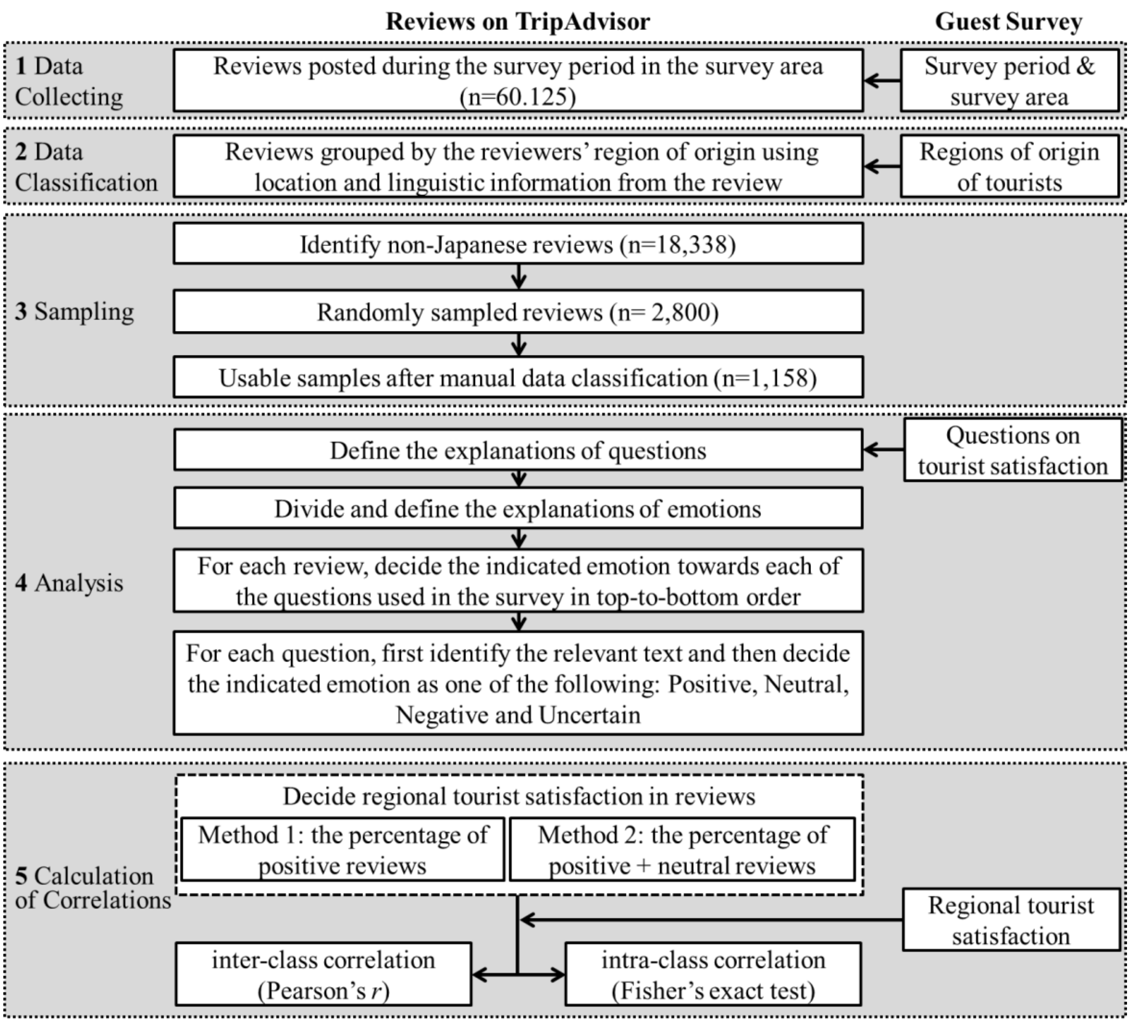

Fig. 4.1 Methodology flow

The remainder of this section is organized accordingly. First, the steps taken to prepare data are explained from section 4.4.1 to section 4.4.3. Next, a method for extracting answers to the survey questions from reviews is developed in section 4.4.4. Finally, section 4.4.5 explains the comparison between the attitudes articulated in reviews and the tourist satisfaction registered in the guest survey. Two aggregation methods are used to determine the expression of tourist satisfaction in the reviews.

4.4.1 Data collecting

Travel reviews collected for this study were to be comparable to the data used in the guest survey. In other words, reviews of the survey area had to have been posted during the survey period. These reviews were collected from TripAdvisor using a data-crawling tool. There were a total of 60,125 reviews of hotels, restaurants, and attractions posted from Jun. 1st, 2016 to Feb. 28th, 2017 in Hokkaido. All those reviews were stored in MySQL. Each review record included the attributes listed in table 4.2.

Table 4.2 Attributes of the reviews

#	**Attributes**	**Explanation**
1	review_id	Id assigned by TripAdvisor, extracted from review_url
2	review_url	Urls of the review
3	attraction_type	Hotel, restaurant, or attraction
4	attraction_name	The name of the hotel, restaurant, or attraction
5	visit_date	Date of visit, inputted by the user, can be empty
6	post_date	Date of posting the review, recorded by TripAdvisor
7	review_title	Title of the review
8	review_text	Content of the review in its original language
9	user_name	Name of the user
10	user_location	Location of the user, unformatted address inputted by the user, can be empty
11	has_translate_button	A value of 0 or 1 indicating whether a review is written in the native language in the website's country domain (e.g., Japanese review in.jp)

4.4.2 Data classification

Tourists are usually grouped by region of residence or nationality in traditional surveys. Thus, travel reviews were also grouped by the reviewers’ country/region of origin. In this study, a heuristic method was used to identify reviewers’ home regions using the location and linguistic information from the review. That is, a reviewer’s home region will be the region in the location information only if the language used in the review is (one of) the native language(s) of that location. For example, if a reviewer filled the location information as Beijing, China, the reviewer’s region will be assumed to be China if he/she writes the review in Simplified Chinese. This technique enables the removal of those who moved abroad for long-term purposes such as jobs or education.

4.4.3 Sampling

The identification of reviewers’ country of origin and language was needed for review classification. Location information is provided in reviews if the reviewer input it in advance. Most addressees are unformatted; thus, they needed to be mapped to their corresponding countries. Moreover, no language tags are provided with the reviews. For this study, reviews were collected from TripAdvisor.jp, which means reviews not written in Japanese appear with a Google Translate Button. This mechanism helped to identify the 18,338 non-Japanese reviews. However, the authors could not further clarify the linguistic information without looking into the reviews.

It would be ideal to study tourists from all ten regions represented in the guest survey. However, due to language constraints, this study could only use reviews written in English, Simplified Chinese, or Traditional Chinese. The tourists studied will therefore be limited to those from seven regions: America, Singapore, Australia, China, Taiwan, Hong Kong, and Britain.

Table 4.3 Numbers of the samples

Country of origin	English	Simplified Chinese	Traditional Chinese	Total
Singapore	332	0	-	332
Hong Kong	122	0	127	249
Australia	190	-	-	190
America	168	-	-	168
Taiwan	-	-	109	109
China	-	75	-	75
Britain	35	-	-	35
Total	847	75	236	1,158

For these reasons, this study first sampled 2,800 reviews from the 18,338 non-Japanese reviews, assuming 400 reviews for each region to provide a sampling size with a confidence level of 95% and an error level of less than 5%. These 2,800 reviews were randomly selected using the following command in MySQL: *select * from review order by rand() limit 0, 2800*, where *review* is the name of the table that stores 18,338 non-Japanese reviews. Then, reviews written in other languages or posted by tourists from other regions were manually deleted, yielding 1,158 usable samples (see table 4.3).

4.4.4 Manual analysis

4.4.4.1 Pre-experiment

Unlike traditional surveys, where targeted information can be acquired with designated questions, travel reviews comprise mainly unformatted, user-generated data. Thus, rules are needed to extract answers to survey question from travel reviews. This section introduces the rules for manual analysis, which were created and developed through trial-and-error experiments.

In this study, manual techniques were used for the following two reason: first, machines have a limited capacity to decide whether a word (or phrase) is relevant to the reviewers (e.g., “we heard people say that this place is great” does not necessarily mean the reviewer thinks the attraction is great); second, without an integrated tourism database/software program, it is impossible for machines to determine whether a word is related to a resource inside the survey area. In addition, in chapter 3, this study found that 79%~91% of the reviews contained non-needs-related information and an average of 9% of the reviews mentioned information about other destinations. Based on these percentages, the limitations of machine-based natural language processes were expected to introduce unavoidable noise into the results. Thus, this study considered it more appropriate to perform manual analysis as a first, necessary step toward an advanced automated analysis.

Experiment purposes. The purposes of the pre-experiments can be specified as 1) creating of a set of rules to help non-expert annotators manually tag reviews using 5 options (i.e., *positive*, *neutral*, *negative*, *ambiguous*, or *unrelated*) related to 11 questions from the guest survey, and 2) confirming the consistency of the results provided by each pair of annotators.

Experiment design. First, survey reports (including the questionnaire) may sometimes fail to provide a clear definition of the questions (see Items in table 4.1), which increases the possibility of confusion and misunderstanding. Therefore, explanations were created by adopting entries from published dictionaries and rewriting them to suit the survey area (e.g., *meals* can be rephrased as *food, or eating food in the survey area*). Second, the five options on the scale described above were used to increase the possibility of higher consistency. The attitudes were reduced to three choices (*positive*, *neutral*, or *negative*); and when no relevant information appeared, the review was tagged as *unrelated*. Additionally, *uncertain* was added as a temporary notation, since ambiguity is inevitable for human subjects.

Procedures. A trial-and-error approach was used to perfect the consistency of the results between individuals. Each circle including four steps: 1) rule creation, 2) the experiment, 3) interviews, and 4) rule modification. Ideally, these rules would produce identical results despite the difference in experience and knowledge between the annotators. Because the analysis was manual, annotators' subjectivity and average time spent per review were also used for the evaluation. Two circles were conducted. The rules in each circle were as follows.

- **Circle 1:** The rules were based on topic-level annotation, where annotators were asked to identify phrases related to each question and identify the attitude expressed by each phrase. Ideally, this can produce a set of annotated data that can be re-used for automated topic-level sentiment analysis in the future.

- **Circle 2:** Rules were based on document-level annotation, where annotators were asked to find text related to each question and identify the overall attitude toward that question.

During the interviews, intensive discussions were conducted with each annotator about each difference in judgment between the annotator and the first author. At the end of each discussion, the annotator was asked whether he/she agreed to change his/her answers. Through this technique helped to obtain the highest possible consistency.

Results. A total of seven non-expert annotators were involved (three in circle 1 and four in circle 2), not including the first author. In the 1st circle, all three annotators complained about the difficulty of identifying and separating related phrases, especially when multiple topics were mentioned in one sentence. It was also difficult to identify attitudes on a phrase level because the contexts were limited. By adopting the 2nd set of rules, it was possible to cut down by half the time needed to process a single review. One annotator from the 1st circle was asked to test the 2nd set of rules to see if any improvement resulted. The feedback that the modified rules were much simpler was received. In the 2nd circle, we found an overall consistency of 86.7%~91.1% between the judgments of the annotators and the first author (see table 4.4). And the consistency can be slightly increased to 90.9%~91.5% with intensive discussion. This suggests that if a third person can fully understand the rules, he or she can possibly obtain around 90% of the same judgments as this study's.

Table 4.4 Consistency (P_0) and Kappa-value of results between each pair of annotators

language	judge1	judge2	pre-interview			post-interview		
			P_0	**Kappa**	**p**	**P_0**	**Kappa**	**p**
CN	D	1st author	0.867	0.622	.000	0.909	0.742	.000
ENG	F	1st author	0.911	0.776	.000	0.915	0.786	.000
JP	E	1st author	0.893	0.601	.000	0.909	0.655	.000
JP	G	1st author	0.879	0.557	.000	-	-	-
JP	E_pre	G	0.885	0.539	.000	-	-	-

4.4.4.2 Final rules

Fig. 4.2 is an example of a travel review. Reviews were output to Microsoft Excel 2010, with the name of the hotel/restaurant/attraction and text of the comment provided. To avoid omissions or errors, this study 1) created a *STATUS* column containing a function indicating whether an answer had been input rather than omitted or doubly input, and 2) limited the legal input to the *POSITIVE*, *NEGATIVE*, *NEUTRAL*, and *UNCERTAIN* columns to 0 and 1.

NAME	REVIEW	STATUS	Questions	POSITIVE	NEGATIVE	NEUTRAL	UNCERTAIN
Odori Park	This event was fantastic! Twelve blocks of snow and ice sculptures, light shows, and food and beverage stands combine for a wonderful festival that makes the most of a long, snowy winter in northern Japan.	ok	for the entire trip and sightseeing	0			
		ok	meals at each tourist destination	0			
		ok	souvenirs	0			
		ok	accommodations	0			
		ok	tourist attractions	0			
		ok	wifi accessibility	0			
		ok	multilingual informational signs at tourist destinations	0			
		ok	local staff's linguistic abilities	0			
		ok	transportation system	0			
		ok	customer service	0			
		ok	scenery	1			

input numbers here
0：not a match
1：a match

Fig. 4.2 Example of analyzing a travel review

The final rules were (1) for each review, fill in the blank cells with the number 0 or 1 for each of the 11 questions (see table 4.5) in top-to-bottom order, (2) for each question, first identify the relevant text from the textual comment and then judge the emotion (see table 4.6) expressed by the related text. To be specific, input 1 in the *POSITIVE* cell if the related text is clearly positive, in the *NEGATIVE* cell if obviously negative, or in the *NEUTRAL* cell if neither positive nor negative or both positive and negative, and *UNCERTAIN* if not clear. Input 0 in any cell if the comment is irrelevant.

Explanations for each question (see table 4.5) were adopted from dictionaries (3rd edition of the Daijirin Japanese Dictionary and Oxford English Dictionary Online). In addition, explanations and examples of the emotions are provided in table 4.6.

All 1,158 sample reviews were manually analyzed following the rules outlined above. The numbers of positive, neutral, and negative reviews in each region are shown in Appendix B.

Table 4.5 Explanations of questions

Items	Explanation
for the entire trip and sightseeing	Comment on the entire travel to Hokkaido.
meals at each tourist destination	Eating food in Hokkaido. Or, food in Hokkaido.
souvenirs	Things or products brought from Hokkaido to home.
accommodations	Rooms, buildings or facilities for staying or living in Hokkaido.
tourist attractions	A place of interest in Hokkaido where tourists visit, typically for its inherent or exhibited natural or cultural value, historical significance, natural or built beauty, offering leisure, adventure and amusement.
transportation system	Facilities (road, bridge, boat, railway, etc.) and vehicles for the movement of passengers or goods (cars, boats, airplanes, etc.) inside Hokkaido.
customer service	The assistance and advice provided by a company to those people who buy or use its products or services during the travel in Hokkaido.
scenery	Natural or artificial view in Hokkaido that are seen through eyes, especially when picturesque.

Table 4.6 Explanations of emotions

Emotions	Explanation	Example
positive	Expressing or implying affirmation, agreement, or permission.	great, good, like, love, beautiful, fascinating, fun…
negative	Not desirable nor optimistic.	terrible, bad, boring, disgusting, disappointing …
neutral	Neither positive nor negative; describing in an objective and non-judgmental way.	new/old, size, amount, fame, time, distance, there are sth…

4.4.5 Calculation of correlations

In most survey reports, statistics are provided at an aggregate level (e.g., the amount or percentage of respondents who select a specific answer to a question), which limits the possible comparisons to the following: 1) inter-class correlation in answering a set of questions. For example, assuming survey respondents were more satisfied with meals than with souvenirs, reviewers showing the same tendency would also be more satisfied with meals than souvenirs; and 2) intra-class correlation in answering each question: For example, assuming 90% of survey respondents were satisfied with meals and 10% of them were unsatisfied, reviewers sharing the same pattern would show the same distribution of satisfied and unsatisfied users.

For inter-class correlation, we can use the Pearson correlation coefficient (or simply Pearson's r) and Spearman's ρ. The Pearson's r is the most widely used method of measuring the association between two sets of data (Lee Rodgers & Nicewander, 1988; Choi et al, 2010). Pearson's r indicates a linear correlation with a value ranging from -1 to 1, where a value of -1 suggests a total negative linear correlation between x and y with all data points lying on a line for which y decreases as x increase, a value of 0 suggests no linear correlation, and a value of 1 suggests a total positive linear correlation for which y increases as x increases. Pearson's r is a parametric measure, meaning that precision decreases with the presence of outlier values. Spearman's ρ is a nonparametric measure of the degree of similarity between two set of ranked order data. Choi et al. (2010) suggested that Spearman's ρ can be used with a sample size as small as 10; however, it is not recommended when the data contain many tied values.

For intra-class correlation, we can use a chi-square test of independence or a Fisher's exact test to show whether the difference between two variables in a contingency table is statistically significant. For example, assuming 595 out of 628 (or 94.8%) survey respondents were satisfied with their meals while 57 out of 59 (or 96.6%) reviewers were satisfied with their meals, the null hypothesis can be that the proportion of satisfied survey respondents is the same as the proportion of satisfied reviewers. The chi-square test and Fisher's exact test can produce a p-value, which indicates whether it is appropriate to reject the null hypothesis at a certain significant level. In addition, the Fisher's exact test is more accurate than the chi-square test when the total sample size is smaller than 1000 (McDonald, 2014).

By manually analyzing the reviews, we can find the numbers of the positive, neutral, and negative reviews in each region (see table Numbers of positive (P), neutral (E) and negative (N) reviews in Appendix B). However, the positive-neutral-negative division of attitudes is not equivalent to that used in the guest survey. Therefore, the two following types of aggregation methods are used to indicate tourist satisfaction in reviews:

- **Method 1: the percentage of positive reviews** for each question in each region, calculated by dividing the numbers of positive reviews (i.e., values of P in the table in Appendix B) by a non-zero sum (i.e., values of Total in the table in Appendix B) of positive, neutral, and negative reviews.

- **Method 2: the percentage of positive + neutral reviews** (i.e., values of P and E in the table in Appendix B), considering that the behavior of mentioning of facts of can be emotional (either positive or negative).

Pearson's r was selected as the measurement for inter-class correlation because many tied values are observed either in the above percentages or in the satisfaction rates. The Fisher's exact test was selected for intra-class correlation because sample sizes were limited.

Calculation of Pearson's *r*. For method 1, considering the percentages of positive reviews in a region as vector x, and corresponding elements in the satisfaction rates in that region as vector y, the Pearson's r can be calculated using formula 4.1, where n is the number of elements in vector x or vector y, x_i and y_i are the elements indexed with i, $\bar{x}$ is the mean of all elements in vector x, s_x is the sample standard deviation of all elements in vector x, and similarly for $\bar{y}$ and s_y. For method 2, Pearson's r can be likewise be calculated by considering the percentages of positive + neutral reviews in a region as vector x. In addition, a t-test is applied to calculate the p-value.

$$r = \frac{(\sum_{i=1}^{n} x_i y_i) - n\bar{x}\bar{y}}{(n-1)s_x s_y} \quad (4.1)$$

Fisher's exact test for each question was conducted using R**.** Using responses to *meals at each tourist destination* in Taiwan as an example, we found that 94.8% of 628 participants were satisfied. Based on this, we can calculate the number of satisfied participants (595) and others (33). In the sampled reviews, we found 39 positive reviews, 18 neutral reviews, and 2 negative reviews. For method 1, where positive was considered equivalent to satisfied, we calculated the number of satisfied reviewers (39) and others (18+2 = 20). By using the following command in R, we found a two-tailed p-value indicating whether the difference between 94.8% (595 out of 628) and 66.1% (39 out of 59) was significant. For method 2, where positive and neutral were both considered equivalent to satisfied, a p-value could similarly be found by changing the number of satisfied reviewers to 39 + 18 = 57 and the other to 2.

Fisher.test (matrix (c(*595*,*33*,*39*,*20*), nrow = 2, byrow = T))

4.5 Results

Table 4.7 shows the percentages of positive reviews in each region. Table 4.8 shows the combined percentages of positive and neutral reviews in each region. Those tables show that the percentages of positive reviews differed from the satisfaction rates in the guest survey; by adding the neutral reviews, however, percentages can approach the satisfaction rates.

Table 4.7 Percentages of positive reviews in each region with p-value from fisher's exact test (%)

#	Items	TW	CN	HK (CN)	HK (ENG)	SG	AU	US	GB
1	for the entire trip	-	0.0 *	-	-	-	-	66.7 **	-
2	meals	66.1 **	50.0 **	56.8 **	67.1 **	65.9 **	77.7	70.5 **	82.6
3	souvenirs	28.6 **	28.6 **	37.5 **	37.5 **	14.7 **	35.7	13.3 *	-
4	accommodations	73.7 **	66.7 **	77.3 **	63.6 **	70.8 **	72.7	63.0 **	55.6
5	tourist attractions	29.0 **	43.5 **	34.8 **	25.0 **	46.8 **	50.0 **	51.9 **	55.6
6	Wi-Fi accessibility	0.0 **	-	0.0	100	22.2 **	75.0	50.0	0.0
7	multilingual signs	0.0	-	100	33.3	33.3 **	12.5 **	18.2 **	50.0
8	linguistic abilities	33.3	0.0 *	25.0 **	44.4 *	22.2 **	27.3 **	28.6	0.0
9	transportation	28.8 **	34.5 **	60.9 **	40.7 **	42.9 **	46.4 **	31.3 **	44.4
10	customer service	89.5	50.0 **	78.6 **	74.0 **	67.0 **	69.8	66.7 **	72.2 *
11	scenery	76.7 **	63.0 **	66.7 **	68.4 **	80.0 *	86.0	78.4 *	100

(* $p < 0.05$, ** $p < 0.01$)

Table 4.8 Percentages of (positive + neutral) reviews in each region with fisher's exact test's p-value (%)

#	Items	TW	CN	HK (CN)	HK (ENG)	SG	AU	US	GB
1	for the entire trip	-	100	-	-	-	-	100	-
2	meals	96.6	83.3**	95.9	95.7	95.7	95.9	93.3	95.7
3	souvenirs	85.7	100	100	100	100	100**	93.3*	-
4	accommodations	97.4	100	97.7	95.5	93.4	93.2	82.6	77.8
5	tourist attractions	90.3	91.3	100	91.7	96.1	91.7	94.2	88.9
6	Wi-Fi accessibility	50.0	-	0.0	100	77.8	100	50.0	0.0
7	multilingual signs	0.0	-	100	33.3	83.3	87.5	100	50.0
8	linguistic abilities	66.7	33.3	62.5	77.8	75.0	72.7	85.7	50.0
9	transportation	98.1	93.1	95.7	98.1	93.2*	92.9	96.9	88.9
10	customer service	94.7	87.5	89.3	84.0**	93.9	92.1	93.7	88.9
11	scenery	100	100	100	97.4	97.8	96.0	100	100

(* $p < 0.05$, ** $p < 0.01$)

Table 4.9 shows the Pearson's *r* for each region. First, except for Australia, (strong) positive correlations were found between the attitudes in reviews and the satisfaction rates in the guest survey. However, we must note that the number of samples was small in the guest surveys in Australia (n = 27), America (n = 22), Europe (n = 27), and Singapore (n = 32). Therefore, Pearson's *r* may differ with a greater number of samples. Meanwhile, we noticed different levels of correlation between Chinese-oriented regions and English-oriented regions. For Hong Kong, a correlation was found with Traditional Chinese reviews, but none was found with English reviews. The 122 English reviews and 127 Traditional Chinese reviews should be enough to provide a sample with an error rate of less than 10% at a confidence level of 95%; however, the linguistic preferences of the participants in the guest survey were unclear.

Table 4.9 Pearson's *r* between the attitudes in reviews and the satisfaction rates in guest survey

Region	Positive		Positive + Neutral		Sample Size	
	r	p	r	p	Review	Survey
TW (n=10)	0.757*	.011	0.746*	.013	109	628
CN (n=9)	0.522	.149	0.935**	.000	75	411
HK_CN (n=10)	0.473	.167	0.774**	.009	127	148
HK_ENG (n=10)	-0.232	.520	0.339	.338	122	148
SG (n=10)	0.648*	.043	0.437	.207	332	32
AU (n=10)	0.423	.223	-0.265	.460	190	27
US (n=11)	0.788**	.004	0.206	.543	168	22
GB (n=9)	0.818**	.007	0.627	.071	35	27(EU)

Table 4.10 Response rate of each question in each region (%)

#	Items	TW	CN	HK	SG	AU	US	GB	Mean
1	for the entire trip	0.0	1.3	0.0	0.0	0.0	1.8	0.0	0.4
2	meals	54.1	48.0	57.8	63.6	63.7	62.5	65.7	59.3
3	souvenirs	12.8	9.3	9.6	10.2	7.4	8.9	0.0	8.3
4	accommodations	34.9	20.0	35.3	31.9	23.2	27.4	25.7	28.3
5	tourist attractions	28.4	30.7	18.9	23.2	31.6	31.0	25.7	27.1
6	Wi-Fi accessibility	3.7	0.0	1.2	2.7	2.1	1.2	2.9	2.0
7	multilingual signs	0.9	0.0	2.0	3.6	4.2	6.5	5.7	3.3
8	linguistic abilities	2.8	4.0	6.8	10.8	5.8	8.3	5.7	6.3
9	transportation	47.7	38.7	40.2	40.1	29.5	38.1	25.7	37.1
10	customer service	17.4	32.0	31.3	34.6	33.2	37.5	51.4	33.9
11	scenery	39.4	36.0	24.9	27.1	26.3	30.4	20.0	29.2
	n	109	75	244	332	190	168	35	-

Table 4.10 is the response rate for each question in the reviews, calculated by dividing the total number (the values in “Total” in Appendix B) in each row by the number of sample sizes. We can see that the response rates were low for “Wi-Fi accessibility,” “multilingual informational signs,” “local staff's linguistic abilities,” and especially low in “the entire trip and sightseeing” (0.4% on average), which means more samples of travel reviews are required to investigate these questions.

4.6 Discussion

This section discusses the results, addresses the limitations, and highlights the contributions of this chapter to research and practice.

4.6.1 General discussion

From our findings, this study would like to believe that with an appropriate aggregate method, it is possible to find similar patterns between reviewers and survey respondents. Hence, it is possible to use the analysis of online travel reviews as a low-cost substitute for traditional surveys. In other words, for investigations attempting to determine specific areas in which tourists are more (or less) satisfied, tourist satisfaction can be predicted using the percentages of positive and neutral reviews from reviewers from Chinese-oriented regions, and the percentages of positive reviews from reviewers from English-oriented regions.

The result that tourists from diverse language backgrounds expressed their satisfaction in different ways is very interesting. From a psychological perspective, the finding that Chinese-speaking reviewers may use both positive and neutral emotions to express their satisfaction is consistent with previous observations made in scholarly literature (Fernández et al., 2000; Davis et al., 2012). For example, Fernández et al. (2000) examined the verbal and non-verbal expressions of joy, anger, and sadness across difference countries. The results of their investigation revealed that Asians (Japanese and Chinese) have a stronger normative system of emotional display rules than other groups (Europeans, Americans, and Latin-Americans). In the tourism context, however, extant cross-cultural studies based on data mining have been limited to the comparison of the statistical distributions of sentimental words (He et al., 2012; Buzova et al., 2019) and / or user ratings (Antonio et al., 2018) between several countries. Although the presence of emotional words or higher ratings may indicate tourist satisfaction within a particular country or language group, we must note that the differences recorded in the statistical distributions may not be equivalent to the disparity in the true feelings of tourists because the expression of satisfaction differs from language to language.

Next, the percentages of combined positive and neutral reviews are numerically similar to the satisfaction rates in the guest survey. Still, doubt remains as to whether those percentages can be directly used to represent satisfaction rates, because most of the values of satisfaction rates in the guest survey are close to 100%; thus, adding the percentage of neutral reviews to the percentage of positive reviews should make the two values equivalent. On the other hand, this result suggests that neutrality may be included as satisfaction in survey results. This incorporation may raise concerns about the social desirability bias as survey respondents may shape their answers to please interviewers. It may also elicit the extreme response bias as survey respondents may only select the most extreme options.

4.6.2 Limitations and future research directions

Nevertheless, the findings outlined above leave us with a question: Where exactly is the split point between the satisfied and unsatisfied in reviews? In this methodology, three possible attitudes were assumed in reviews (*positive*, *neutral*, and *negative*). From the results, we can assume that the split point between satisfied and dissatisfied English-speaking tourists may be closer to the positive value, while for Chinese-speaking tourists, it may be close to neutral. However, if we could further divide the attitudes between positive and negative into *k* degrees, we may be able to identify more precise split points by examining *k*-1 patterns of aggregate method and the coefficients.

This study has many other limitations. First, because reviews are collected only from TripAdvisor, their conclusions cannot be generalized to other data sources. In addition, reviews are grouped by location and linguistic information, which may not be sufficient to identify a tourist's place of residence or nationality. Other features, such as gender or age, also can be taken into consideration (Fujii et al., 2017). Also, due to the linguistic constraints, reviews are limited to seven regions. To overcome this limitation, translation service such as Google Translate may be adopted. In addition, due to the experimental constraints of data collecting, the scale of the guest survey was rather small; however, the method itself can still be considered transferable. Thus, this method should be applied to a larger-scaled survey, a national survey for example, to further validate our conclusions. Moreover, the development of an automated analysis method is needed to overcome the limitations of manual analysis. Furthermore, it may be possible to find some interesting results related to tourist satisfaction if we followed the time sequence of several travel reviews posted by individual tourists using the expectation disconfirmation theory (Oliver, 1980). Finally, the credibility of the travel reviews is an inevitable problem in online data mining; thus, methods are needed to rule out those unreliable data.

4.6.3 Implications

With regard to the investigation of tourist satisfaction, both traditional tourism survey and the analysis of online data suffer from certain limitations such as non-response bias or the response biases mentioned in the literature. Nevertheless, as a commonly accepted method, traditional surveys continue to be conducted by destination marketing organizations. Meanwhile, more focus is placed on improvements in the accuracy and efficiency of extracting and interpreting useful information from online data. However, the general assumption that the conclusions yielded by online data can accurately represent tourist needs has not been adequately tested.

This study contributes in several ways to the appreciation of the issues pertaining to the analysis of travel reviews and to an understanding of the value of such online user appraisals to the investigation of tourist satisfaction.

First, a method was devised to conduct an appropriate comparison between the attitudes expressed in travel reviews (textual data) and the tourist satisfaction recorded in traditional tourism surveys (quantitative data). This method is based on statistics, social psychology literature, and information technology techniques and was developed through trial-and-error experiments. By relating online travel reviews and traditional surveys, the existence of cross-cultural differences in display rules was confirmed in reviews, and the possible influence of the social desirability bias and the extreme response bias in traditional survey was noted. These biases, though obvious and self-evident, are difficult to detect if the two forms of ascertaining tourist satisfaction are examined independently. Also, the manners in which these biases are likely to influence the results are not very clearly illustrated in extant literature. Further, from the practical point of view, if a generalized model of how attitudes in online travel reviews and results in traditional surveys relate to each other could be established, the former could be used as a proxy for the latter. Even more, travel reviews could be employed as a cross-checking tool to detect errors or faked results in traditional surveys.

Additionally, as shown in table 4.4, this study identified one of the limitations of sentiment analysis attributed to the human factor during the manual annotation. The consistency between independent annotators is usually reported to assess the reliability of the results of manual annotation (Pang et al., 2002; Tokuhisa et al., 2015; Kim & Stepchenkova, 2017). However, little has been accomplished to show the highest possible consistency. This study found that consistency can be increased slightly through intensive discussion, but discovered that reaching 100% agreement is difficult. Tourangeau et al. (2002) suggest that human attitudes contain existing evaluations, vague impressions, general values, and relevant feelings and beliefs. Therefore, even when two individuals fully understand each other’s logic and reasoning, they would still not agree with each other because of attitudinal differentials. This fact incorporates inevitable bias into any sentiment analysis which involves manually annotated data.

In addition, several of our findings can be incorporated into the advancement of future automated analysis. First, apart from precision, recall, and F_1 value, the consistency and kappa-value in table 4.4 can be used to evaluate the results of an automated analysis.

That is, if a program can reach similar or higher consistency and kappa-value, we can consider it an individual intelligence and conclude that its performance is as good as that of a non-expert annotator. Secondly, because identification of related text is the first step in automated analysis, methods are needed to enable machines to relate text in reviews with questions in surveys. However, questions differ in each survey (meaning that one labeled dataset may not be applicable for another survey), and the creation of labeled data is costly and difficult, especially at the level of specific topics. Those problems limit the adaptability of automated analysis to dictionary-based approaches (Serna et al., 2016) or non-supervised machining learning approaches (Lu et al., 2011). However, reviews contain many regional terms, most of which are not included in existing lexicons (e.g., local foods and local brands). Thus, should a dictionary-based approach be adopted, methods will be needed to extract and classify unknown words. Furthermore, it will be important to identify texts related to resources in the survey area (e.g., attitudes toward Tokyo's foods should not be counted as attitudes toward Hokkaido's foods). To do this, we should also create a knowledge database of the local characteristics of each destination (Kim & Stepchenkova, 2017; Suzuki & Kurata, 2017).

4.7 Conclusion

This chapter aimed to determine tourist satisfaction from travel reviews through the use of text data mining. A method was presented to manually extract tourist attitudes from reviews to be compared to the satisfaction rates in a traditional survey, using travel reviews from TripAdvisor and the Hokkaido survey as an example. The main findings are as follows:

- By calculating Pearson's *r*, we found (strong) correlations between the attitudes in reviews and the satisfaction rates recorded in the guest survey in six out of seven regions. This finding suggests that it is possible to use the analysis of online travel reviews as a low-cost solution to replace traditional surveys.

- The percentages of positive reviews differ from the satisfaction rates recorded in the guest surveys, whereas the percentages of combined positive and neutral reviews are numerically similar to these satisfaction rates. This finding suggests that survey results can be more positive than they should be, possibly because of the influence of the social desirability bias and the extreme response bias.

- Tourists with different language backgrounds expressed satisfaction differently, which suggests that when using textual data to study satisfaction in different regions, reviewers' native languages should be taken into account.

However, this study has a number of limitations, such as limited data sources and limited sample sizes. Thus, the conclusions should be further validated with a larger-scaled survey, more review samples, and the support of a developed automated analysis method.

Chapter 5 Analyzing the number of reviewers for investigating tourist arrivals

5.1 Introduction

Apart from tourist satisfaction, the number of tourist arrivals is another fundamental statistical indicator. It enables comparisons of current industry statuses and trends among regions (JTA, 2019b), responsive decision-makings according to seasonal surge/drop in demand (Yagasaki, 2015), and monitoring places of interest currently favored by tourists (Saeki et al., 2015). According to JTA's Common Standard (2019b), in Japan, current approaches of data collecting consist of interviews with managers and questionnaires with sampled visitors at selected sightseeing spots, festivals and events in a limited period of time. Not only do these types of investigations have the potential to be time-consuming and expensive, they may also produce inaccurate results since they are only conducted on a limited number of days every three months. In addition to the time needed for conducting a survey, an intervening delay is also common between the date of the investigation and the date of the publication of the statistics. This issue increases the difficulty for prompt comparisons among destinations.

Recently, case studies such as Saeki et al. (2015) or JTA (2017), as well as application development such as Nightley Inc. (2019) have been conducted to analyze the visited places of SNS users. Making productive use of online data potentially makes it possible to monitoring travel flow quicker and at a finer level compared to traditional surveys.

The idea of monitoring the flow of tourist via the use of online data analysis is based on the hypothesis that the volume of the creators of online data is positively correlated to the volume of tourist arrivals; thus, as the volume of tourist arrivals increases, the volume of online data creators also increases. However, as mentioned in section 1.2, the validation of the relationship between these two variables is very limited. For example, Saeki et al. (2015) compared the results of a governmental survey and the results of the analysis of Twitter data to show that the latter can act as a low-cost proxy for the former. To be specific, they collected tweets in Tokyo during a period of three months and in four different languages and calculated the ordinal correlation between the rankings of the number of Twitter users in ten sightseeing areas and the rankings of the number of visitors in corresponding areas provided by the Tokyo government. But the sample size in their study was statistically insufficient (i.e. n = 10 areas x 3 months for each language group) and the conclusions were drawn only in ordinal level. Furthermore, other factors such as seasonal change of the volume of tourist arrivals and the difference of the characteristics among various destinations could influence the relationship between the two variables. Thus, this hypothesis should also be tested across a longer period of time and across various destinations.

Additionally, because previous studies focused mainly on SNS data, the potential usefulness and limitations of travel reviews from travel websites has not been fully explored. For example, destinations and tourism facilities are classified according to the administrative districts on travel websites. And the reviews are directly related to each tourism facilities. This structure enables the identification of the destination (i.e. prefecture or city) and the facility directly from online travel reviews, which can be utilized to estimate the number of tourist arrivals at a facility level. Furthermore, as shown in Chapter 4, the content of online travel reviews can potentially reflect tourist satisfaction. Thus, analyzing online travel reviews can be utilized to create an integrated framework from identifying the problematic districts/facilities based on estimating the number of arrivals, to further identifying the problems based on content analysis.

Based on the above, this chapter aimed to show the validity of estimating the number of tourist arrivals from online travel reviews. For this purpose, a method was created to aggregate the number of reviewers from travel reviews on TripAdvisor for comparisons with the number of arrivals from a traditional survey. This chapter focused on the following two tasks.

1) Find the relationship between the number of arrivals and the number of reviewers in the context of all the travelers (i.e. domestic and international)
2) Further examine the relationship by month, by city, and by region in the context of inbound tourism.

5.2 Methodology

This section explains the methodology employed to find the relationship between the number of tourist arrivals and the number of reviewers.

5.2.1 Choosing a survey on the number of tourist arrivals

To examine the relationship between the number of tourist arrivals and the number of reviewers, reliable statistics of number of tourist arrivals are needed. This study narrows extant published statistics down to administrative statistics in Japan, as introduced in section 2.1.2.

In Japan, 46 out of 47 prefectures except for Osaka are conducting the survey on Inbound Domestic and Foreign Tourists based on JTA's Common Standard (JTA, 2019b). Based on the survey in section 2.1.2, published administrative statistics on the number of international arrivals (or overnight stays) take the following forms:

1. total number of all international arrivals in a whole year in the whole prefecture,
2. total number of all international arrivals in each month in the whole prefecture,
3. total number of all international arrivals in a whole year in each area/city,
4. total number of arrivals in each month by regions of residence in the whole prefecture,
5. total number of arrivals in a whole year in each city by regions of residence.

And among those prefectures, only Hokkaido government published both 4 and 5. Thus, to enable the comparisons by regions of residence, by cities, and by months, survey reports published by Hokkaido government (2019) were adopted in this study.

Hokkaido Survey on tourist arrivals has been carried out quarterly since 2010 based on Hokkaido's survey guideline, a localized revision of JTA's Common Standard (2019). In the Hokkaido survey, a year is divided from Apr. 1st to next year's Mar. 31st. The latest annual report is about year 2017 (from Apr. 2017 to Mar. 2018) in Sep. 2019. As noticed, the latest report is one and a half years behind time.

There are two types of estimated statistics about tourist arrivals in this report. The first type is the number of sightseeing visitors in each month in each city. It is an estimated value based on the results from the on-spot surveys on the actual number of inbound tourists in each city, on-spot parameter surveys (e.g., the percentage of tourists with each attribute, the average number of spots visited) in Hokkaido, and the national statistics provided by JTA. And the number of sightseeing visitors includes 1) the number of tourists from other countries and other prefectures, 2) the number of tourists from Hokkaido, 3) the number of one-day sightseeing visitors, 4) the number of overnight sightseeing visitors, and 5) the sum of 1 and 2 (or 3 and 4). This study adopted the summed number in each month in each city to be compared to the number of reviewers in each month in each city, because reviews are posted by tourists that fall into all first four aforementioned types.

The second type is the number of overnight travelers. It includes 1) the sum of international and domestic overnight travelers, and 2) the number of international overnight travelers by region of residence. The latter is published in two subordinary forms: a) in each month in the whole Hokkaido and b) in each city during one year). It is considered that the number of international overnight travelers is approximately equally to the number of international sightseeing visitors, because one-day international traveller is few.

Similar to the number of sightseeing visitors, the number of international overnight travelers is estimated based on JTA's statistics and JTA's parameter surveys. Furthermore, two methods are used to aggregate the number of international overnight travelers as following: 1) the total number and 2) the cumulative total number. For example, if one traveller stays in a hotel for five days, he/she will be counted as one traveller in the first method but five travelers in the second method. This study adopted the numbers calculated by the first method to be compared with the number of international reviewers by region of residence, because the lengths of staying are not provided in online travel reviews.

Table 5.1 Attributes of reviews

Attribute	Explanation
Review URL	Used to extract Review ID, City ID and Facility ID
Username	Include Anonymous Users
Address	Include Anonymous Users; Approx. 34.4% with no address; In the form of unformatted text
Date of posting	Year/ Month/ Day

5.2.2 Data collecting

Travel reviews are categorized into *Hotel*, *Restaurant*, *Attractions*, *Flights* and *Others* on TripAdvisor. To collect reviews relative to each destination, this study focused on the first three categories.

Then, to collect corresponding reviews, all reviews posted to all 179 cities in Hokkaido during Jan. 1st, 2000 to Mar. 31th, 2018 were collected using a data colleting tool. Altogether 272,117 reviews were collected. The attributes of reviews are shown in table 5.1.

5.2.3 Data classification and aggregation

To be compared to the statistics registered in the guest survey, the number of reviewers that includes both international and domestic reviewers in each moth in each city, as well as the number international reviewers by region of residence in each month in the whole Hokkaido and in each city during one year is needed. Following is the method to calculate those numbers.

First, the reviews were grouped by year using their date of posting. A year is divided from Apr. 1st to next year’s Mar. 31st using the same division used in the guest survey. Next, all the reviews were further grouped by month, and then by city using the City ID extracted from their URLs.

For comparisons of all travelers, the number of usernames with different spellings in each month in each city in all the reviews was used as the number of monthly reviewers in each city.

Meanwhile, for comparisons by region of residence, country of residence was extracted from reviews that contain location information using the following steps.

1. Compose a list of addresses with different spellings from all the reviews.
2. Manually create a mapping list that maps each location to a specific region of residence.
    - For location with region information, the region will be the one indicated in the address. For example, “Sapporo, Japan” is mapped to “Japan”.
    - For the location without such information, the regions were determined by search results via Google. To be specific, if the search results of the location showed the same region, that region would be judged as the region of that location. If the results showed multiple regions, the location would be excluded.
3. Map the location information in reviews to regions of residence using the mapping list.
4. Exclude the reviews that are not from the 18 regions surveyed in Hokkaido.

Then, the number of usernames with different spellings in each month in each city in each residential region was used as the number of regional monthly reviewers in each city. Finally, the total number of annual regional reviewers in each city and the total number of regional monthly reviewers in the whole Hokkaido were calculated to be compared with the results registered in guest survey.

5.2.4 Calculation of correlations

To investigate the correlation relationship between the two variables, the most commonly adopted approaches namely scatter plot and Pearson’s product moment correlation coefficient (i.e. Pearson’s r) were used. Scatter plot uses dots to represent values for two different numeric variables to help visual observation of the relationship between them. Pearson’s r, as introduced in chapter 4, indicates a linear correlation with a value ranging from -1 to 1, where a value of -1 suggests a total negative linear correlation between x and y with all data points lying on a line for which y decreases as

x increase, a value of 0 suggests no linear correlation, and a value of 1 suggests a total positive linear correlation for which y increases as x increases. Also, as the absolute value of Pearson's r increases, the relationship between x and y becomes closer to a regression line $y = ax + b$. This means that the value of x or y can be determined by the other using a linear equation. In this study, x represents the number of reviewers and y represents the number of tourist arrivals in order to produce a regression line for the estimation of y based on x.

5.3 Results

This study adopted the lasted two years' reports as comparison. Table 5.2 shows the number of reviews in 2016 and in 2017

5.3.1 Overall correlation

Fig. 5.1 is the scatter plot where x is the number of monthly reviewers in each cities and y is the number of monthly sightseeing visitors in each cities. The Pearson's r shows a strong positive correlation between the two variables at the significant level of 0.01 in 2016 (r = 0.807, n = 2148) and in 2017 (r = 0.838, n = 2148). This result suggests it is possible to use the number of reviewers as a quick reference of the number of sightseeing visitors. Because the tendency in 2016 and 2017 is similar, this result is possibly reproducible in the future. However, outlier values exist in the two scatter plots. This result suggests that estimation based on basic linear regression would bring errors in certain cities or months..

Table 5.2 Number of reviews

Year	**Amount of reviews**		
	All reviews	**Reviews with location**	**Reviews from 18 regions**
2016	61,696 (100.0%)	48,143 (78.0%)	15,406 (25.0%)
2017	57,786 (100.0%)	43,793 (75.8%)	14,874 (25.7%)

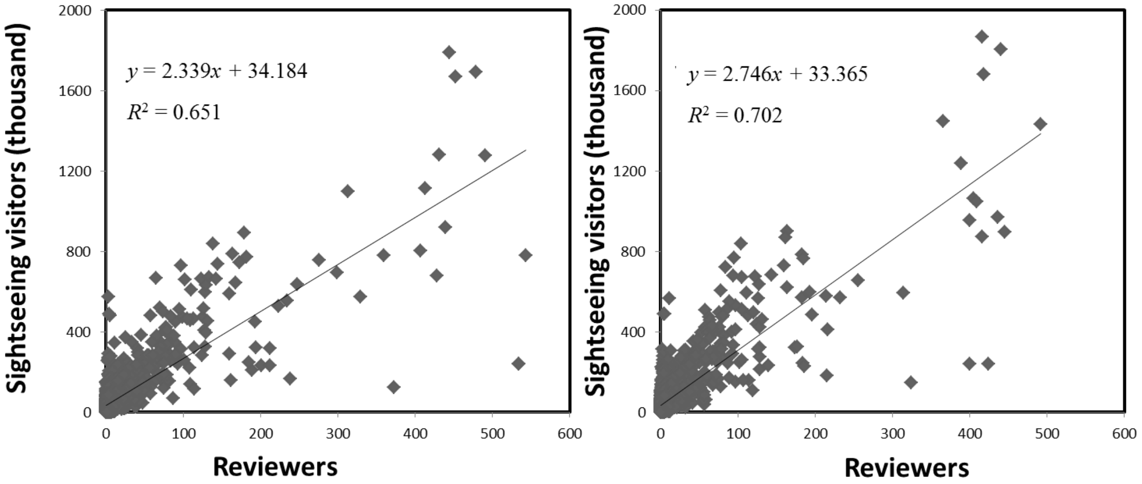

Fig. 5.1 Correlation between the number of sightseeing visitors and the number of reviewers in each city in each month (left: year 2016, right: year 2017)

5.3.2 Correlation by parameter

This section presents the results of the correlations when further examined by city, by month, and by region of residence.

5.3.2.1 Correlations by city

Fig. 5.2 shows the values in distinctive locations. Only seven of out 179 cities are marked because space is limited. Different colors are assigned to each city.

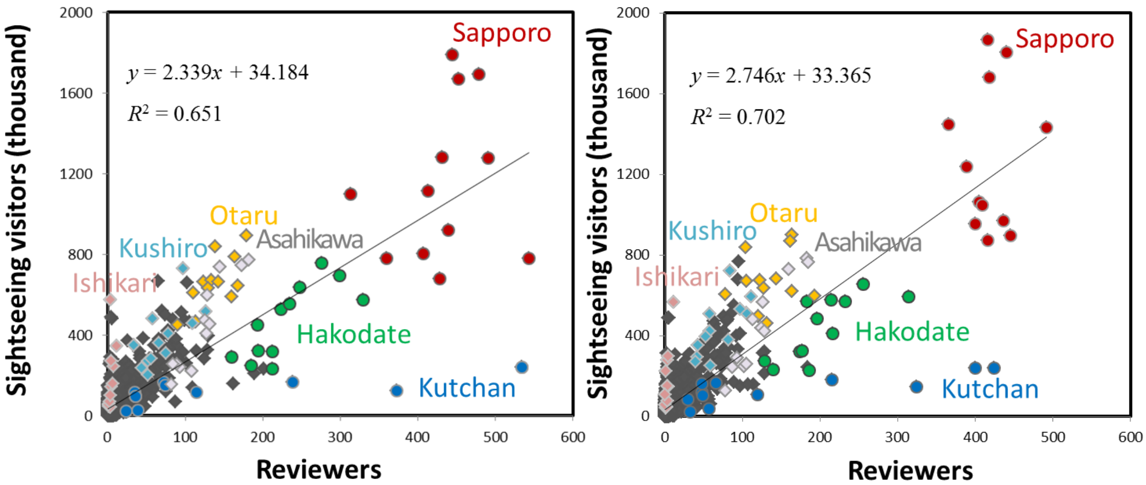

Fig. 5.2 Scatter plots in major cities (left: year 2016, right: year 2017)

In Fig. 5.2, under the regression line lie the values from Kutchan and Hakodate. Meanwhile, values from Otaru, Asahikawa, Kushiro, and Ishikari are above the regression line. Values from Sapporo are divided. Sapporo, Kutchan, and Hakodate are the three major tourism destinations in Hokkaido. A larger proportion of sightseeing visitors posted reviews in these three cities when compared to the proportions in other cities. Sapporo, as the capital city of Hokkaido, receives more sightseeing visitors and reviewers than other cities. Kutchan, one of the famous skiing resorts, shows extremely high reviewer-to-sightseeing-visitor ratios, especially in winter. Hakodate is another famous tourism destination in Hokkaido. Among the data points of Hakodate, the ones in spring and summer are on the right side.

Correlations in these seven cities are shown in table 5.3. Meanwhile, a full list of the correlations in each one of the 179 cities is attached in Appendix C. Because the sample size is small, the following conclusions drawn from table 5.3 could be biased. The correlations in destinations such as Kutchan, Hakodate, Asahikawa, or Kushiro are positive over two years. However, the tendency shifted in Otaru. Moreover, the distribution of values in Sapporo and Ishikari is almost vertical, resulting in severe change of parameters in the two years. That means linear regression may not be the best solution in these destinations.

Table 5.3 Correlations in seven cities in 2016 and 2017 (n = 12)

179 cities	**2016**			**2017**		
	Pearson's *r*	**Slope *a***	**Intercept *b***	**Pearson's *r***	**Slope *a***	**Intercept *b***
Kutchan	0.783**	2.230	-115.4	0.781**	1.698	-77.4
Hakodate	0.777**	0.213	130.7	0.758**	0.244	94.6
Asahikawa	0.937**	0.138	61.3	0.959**	0.152	55.8
Kushiro	0.673*	0.127	21.9	0.714**	0.128	13.3
Otaru	0.708*	0.142	44.0	0.170	0.040	107.8
Ishikari	0.480	0.009	2.3	0.861**	0.017	-0.1
Sapporo	0.204	0.031	397.4	0.080	0.007	410.4

5.3.2.2 Correlations by month

Table 5.4 shows the correlations in each month in 2016 and 2017. Positive correlations are found in each month at the significance level of 0.01. In addition, the values of correlation increased from Apr. to Dec. when compared to the overall correlation in each year. This result suggests that creating independent regression lines in each month from Apr. to Dec. could provide results with higher precision. Meanwhile, from Jan. to Mar., the correlation could be lowered by the diversity among various destinations. Moreover, seasonal change of the values of slope can be observed. The tendency is similar in these two years (r = 0.955**, n = 12), reaching the top in Aug. and the bottom in Jan. Thus, it is possible that the values of slope are predictable.

Table 5.4 Correlations by month in 2016 and 2017 (n = 179)

Month	2016			2017		
	Pearson's r	Slope a	Intercept b	Pearson's r	Slope a	Intercept b
Apr.	0.887**	2.971	19.696	0.914**	2.784	19.396
May	0.858**	2.429	46.341	0.882**	2.643	44.716
Jun.	0.927**	3.001	32.156	0.942**	3.786	28.816
Jul.	0.923**	3.511	46.58	0.934**	3.89	44.97
Aug.	0.895**	3.655	50.681	0.894**	4.024	53.015
Sep.	0.922**	3.308	37.047	0.928**	3.709	34.008
Oct.	0.877**	2.505	34.98	0.889**	2.837	36.702
Nov.	0.910**	2.271	14.754	0.888**	2.261	19.565
Dec.	0.809**	1.64	20.304	0.868**	2.029	18.728
Jan.	0.699**	1.225	27.963	0.759**	1.583	26.315
Feb.	0.698**	1.23	30.846	0.725**	1.876	29.452
Mar.	0.772**	1.805	21.594	0.785**	1.955	23.772

5.3.2.3 Correlations by region of residence

Table 5.5 shows the regional correlations where *x* is the number of annual regional reviewers and *y* is the number of annual regional overnight travelers in each city. Sample size *n* is the number of cities in Hokkaido recorded with at least one overnight traveller from the corresponding region. As a result, strong positive correlations were found in all eighteen regions of residence at the significance level of 0.01. Besides, the values of slope *a* in table 5.5 vary by region, which means the tendency to provide information on TripAdvisor differs by region.

Table 5.5 Correlations by region of residence in each city in one year

Region	2016				2017			
	r	***a***	***b***	**n**	***r***	***a***	***b***	**n**
Great Britain	0.924	50.876	-16.0	80	0.962	65.812	11.1	90
France	0.926	60.077	44.0	78	0.958	79.977	24.0	84
Germany	0.930	93.165	18.0	76	0.979	85.474	20.2	78
Canada	0.971	79.382	-41.3	76	0.965	95.382	-12.1	81
Australia	0.966	70.339	4.5	97	0.936	104.907	39.7	92
U.S.	0.959	120.659	-171.8	109	0.979	136.141	-142.6	118
Vietnam	0.959	96.589	11.9	53	0.950	149.122	21.1	61
Singapore	0.968	131.705	36.3	94	0.984	157.286	-218.3	103
India	0.639	139.936	57.0	42	0.956	174.604	-8.4	44
Philippines	0.993	127.453	14.8	51	0.986	196.999	-39.0	48
Indonesia	0.986	176.555	44.3	57	0.987	298.321	-36.5	56
Russia	0.967	174.125	-9.3	55	0.981	305.997	-21.3	56
Malaysia	0.965	270.437	526.3	78	0.978	395.567	-155.7	82
Hong Kong	0.960	373.407	-494.7	107	0.981	421.351	-560.3	124
Thailand	0.970	625.155	-561.3	88	0.981	735.372	-37.5	90
South Korea	0.974	1496.277	323.9	114	0.979	1639.547	44.4	122
Taiwan	0.948	1478.467	1053.8	129	0.939	1742.455	990.8	130
Mainland China	0.957	4095.889	-1060.1	127	0.964	5813.324	-931.6	132

5.4 Discussion

5.4.1 Potential of determining the number of tourist arrivals from the number of reviewers

Because strong positive linear correlation was found between the number of monthly tourist arrivals in each city and the number of monthly reviewers in each city as the significance level of 0.01, we would like to believe that it is possible to estimate the former based on the latter. In other words, form the perspective of bivariate correlation, analyzing the quantity of reviews can potentially serve as a low-cost and quicker substitute for traditional surveys.

However, it should be noted that outlier values were observed in Fig. 5.1 such as the ones in Sapporo, Kutchan, and Hakodate. This result suggests that estimation based on basic linear regression would bring errors in certain cities or seasons. Regarding this phenomenon, advanced regression techniques or parameter tuning can be applied to improve the precision. For example, Fig. 5.2 shows that data points from Sapporo, Kutchan, and Hakodate are at similar locations in 2016 and 2017, suggesting that the locations of certain cities in certain months can be distinctive. Therefore, the network of these features among various cities and/or months could be utilized to reduce the errors in future. Also, those values should be separated from the values in other cities using individual regression lines for each city to increase the precision of the estimation. Besides, considering these cities are the top tourism destinations in Hokkaido, it is possible that this phenomenon also exists in other major tourism destinations in other areas. The difference between major destinations and the others implies that conclusions acquired in major destinations may not be appliable to other destinations. It is possible that the difference could be caused by the e-word-of-mouth effect that positive reviews attract more tourists which in turn boosts the reviews.

Regarding the seasonal variations, as an inevitable phenomenon in tourism industry (Oi, 2012), is confirmed in the relationships between the number of tourist arrivals and the number of reviewers in each city in this study. Because a similar pattern is found in table 5.4, it is possible that seasonal parameter tuning could benefits the estimation precision by month.

5.4.2 Potential of analyzing tourist preferences by region of residence based on posting rate

In the comparisons by regions of residence, the findings showed the tendency that as the number of overnight travelers increases, the number of reviewers increases. Besides, the proportions of reviewers in TripAdvisor to overnight travelers differ by regions of residence. In this study, this proportion is defined as posting rate as shown in formula 5.1. In this section, we explored the regional posting rate in each month and in each city for hints in analyzing tourist coming from different regions.

$$\text{Posting rate of region } i = \frac{\text{the number of reviewers with location information from region } i}{\text{the number of overnight travelers from region } i} \quad (5.1)$$

Fig. 5.3 shows the annual regional posting rate in Hokkaido in 2017, where *n* is the number of reviewers who provided their location information. Posting rate of the British recorded the highest among all regions, followed by the Canadian, Australian, France, German, American and etc. According to Hokkaido's report, international arrivals from Western countries are far less than those from Asian regions. Because the numbers of arrivals are estimated values based on the results from on-spot surveys and JTA's Common Standards, another doubt may arise, which is that tourism survey may suffer from insufficient on-spot sampling due to limited amounts of visitors from those countries. Consequently, it may lead to the problem of failing to provide statistically significant investigation results. In this particular context, analyzing the quantity of travel reviews may service as a second opinion for cross checking the reliability of traditional survey results. Furthermore, because the westerns are more inclined to provide their information on TripAdvisor than Asians, analyzing their travel reviews may compensate the problem of insufficient samples in analyzing the preferences of tourists coming from certain regions.

Controversially, posting rates in Asian regions are comparatively low. In that case, TripAdvisor may not be the major social media channel in those countries/regions. Therefore, employing TripAdvisor's travel reviews to analyze tourists in those countries/regions could suffer from the risk of biased samples that are restricted to certain residential areas, age groups, or etc. In addition to the penetration rate of TripAdvisor in different countries/regions, the proportions of foreigners in mid/long term stay for job or education purpose is another factor that should be taken into account. For this matter, language recognition could be used along with the address to determine the nationality of the reviewer during the data classification. Moreover, collecting reviews from major travel websites in each country might be an adoptable solution for cross-region comparisons. In that case, the approach must be validated with cautions because the structures and contents of travel reviews can be influenced by the design of a website (Xiang et al., 2017).

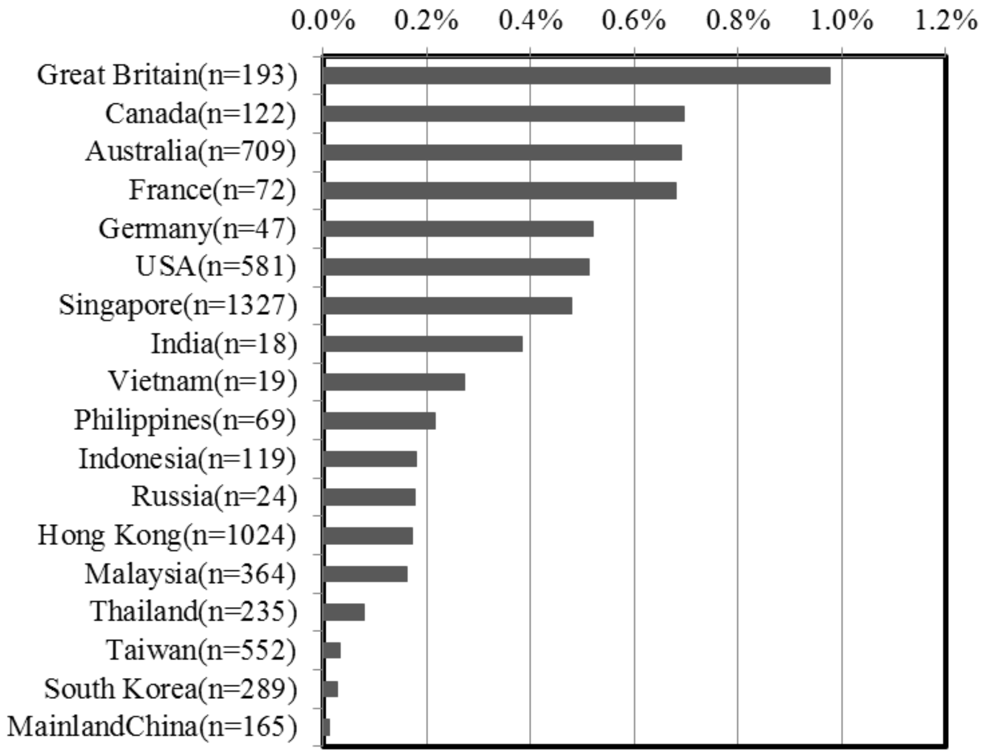

Fig. 5.3 Annual regional posting rate in 2017

5.4.2.1 Regional posting rate by month

Fig. 5.4 shows the regional posting rate in each month in 2017. For each region of residence (i.e. each row), the color is darker when the posting rate is higher. As can be seen, regional posting rate further differs by month. Also, the results are interesting in the perspective that the tendency in most regions is controversial to the results in table 5.4. In other words, opposite to the posting tendency among all travelers (i.e. domestic and international), most regional posting rate is high in spring and summer, but comparatively low in winter. For example, according to Hokkaido’s report, except for Vietnam, the numbers of overnight travelers from Southeast Asia are higher in winter; however, their posting rates are higher in June or August. Countries in Southeast Asia are known for their hot climate. Considering their less chance of having snowfalls in winter, high amount of travel flow is expected from those countries to Hokkaido, explaining the results of high overnight travelers in winter.

Regions	Apr.	May	Jun.	Jul.	Aug.	Sep.	Oct.	Nov.	Dec.	Jan.	Feb.	Mar.
Great Britain	2.061%	2.643%	2.976%	0.651%	1.321%	1.139%	1.218%	1.426%	0.668%	1.054%	1.393%	2.178%
Canada	1.242%	0.922%	0.827%	0.556%	0.903%	1.964%	1.243%	1.329%	1.149%	1.124%	1.182%	1.913%
Australia	1.518%	2.178%	0.824%	0.419%	1.000%	0.338%	2.142%	0.000%	2.778%	1.069%	0.508%	1.515%
France	0.459%	2.525%	2.070%	1.738%	0.600%	0.533%	0.000%	0.794%	0.000%	1.190%	0.683%	1.460%
Germany	1.004%	2.358%	1.538%	1.468%	3.353%	2.168%	1.458%	1.382%	0.432%	0.740%	0.640%	0.929%
USA	1.202%	0.740%	0.761%	0.697%	1.141%	1.162%	0.806%	1.242%	0.834%	0.957%	0.542%	1.042%
Singapore	0.000%	1.520%	1.776%	0.543%	0.467%	0.000%	0.350%	0.000%	0.252%	0.000%	0.251%	3.409%
India	0.833%	0.577%	0.781%	0.802%	1.509%	1.222%	0.545%	0.557%	0.418%	1.270%	0.734%	0.674%
Vietnam	0.313%	1.170%	1.230%	0.308%	1.129%	1.231%	0.533%	0.904%	0.692%	0.633%	0.140%	0.256%
Philippines	0.645%	0.888%	1.942%	0.367%	0.736%	0.974%	0.476%	0.348%	0.296%	0.691%	0.550%	0.286%
Indonesia	0.271%	0.900%	0.591%	0.639%	0.263%	0.330%	0.077%	0.113%	0.236%	0.234%	0.558%	0.150%
Russia	0.732%	0.375%	0.213%	0.470%	1.442%	0.572%	0.332%	0.254%	0.142%	0.339%	0.538%	0.639%
Hong Kong	0.353%	0.328%	0.294%	0.339%	0.667%	0.317%	0.385%	0.153%	0.144%	0.231%	0.161%	0.294%
Malaysia	0.311%	0.226%	0.248%	0.289%	0.303%	0.250%	0.204%	0.297%	0.213%	0.371%	0.201%	0.320%
Thailand	0.085%	0.196%	0.157%	0.215%	0.403%	0.172%	0.103%	0.174%	0.163%	0.092%	0.076%	0.115%
Taiwan	0.074%	0.085%	0.056%	0.044%	0.028%	0.076%	0.061%	0.147%	0.073%	0.035%	0.061%	0.055%
South Korea	0.094%	0.028%	0.049%	0.037%	0.063%	0.077%	0.049%	0.050%	0.067%	0.061%	0.038%	0.053%
MainlandChina	0.021%	0.034%	0.020%	0.014%	0.020%	0.013%	0.022%	0.021%	0.017%	0.013%	0.017%	0.032%

Fig. 5.4 Regional posting rate in each month in 2017

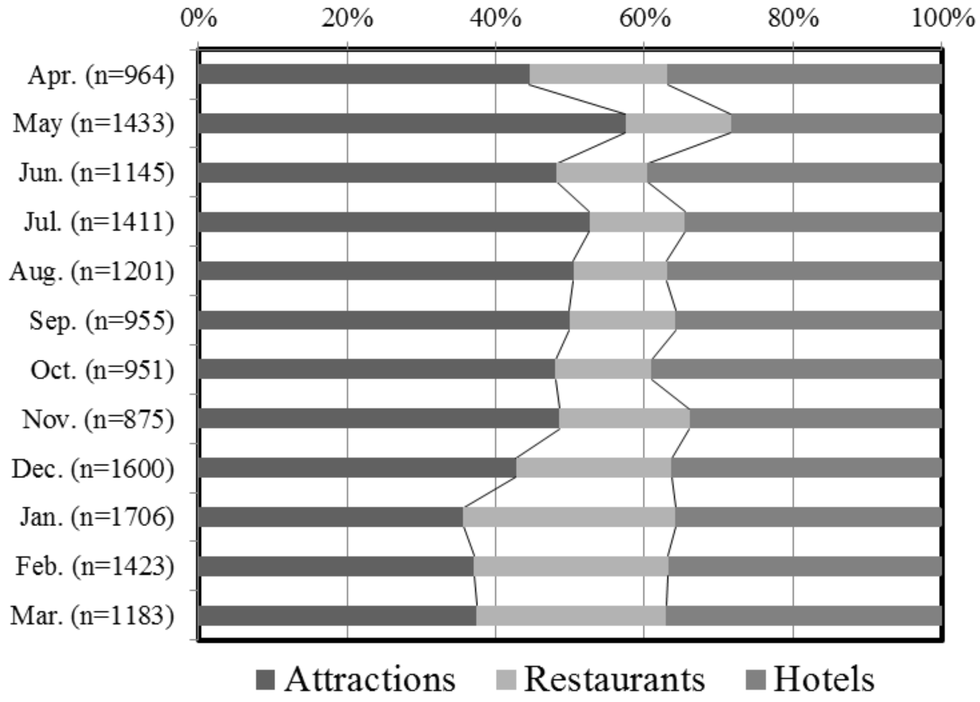

Fig. 5.5 Monthly amount of travel reviews by category in 2017

On the other hand, the disagreement between tourist arrivals and posting rate may raise the following discussions. First, based on the expectation disconfirmation theory (Oliver, 1980), one hypothesis is that the satisfactions of the travel could be mediocre compared to the expectations, resulting in insufficient motivations to post reviews. One clue can be found in Fig. 5.5, which shows an approx. distribution of travel reviews posted by reviewers from the eighteen regions grouped by category and by month. As the season approaches to winter, the proportions of attraction reviews become smaller. Normally, winter's Hokkaido is famous for its illumination events such as the Sapporo Snow Festival, skiing and bathing in hot spring surrounded by snow. But compared to the strong image of snow and ice, attractions in winter (e.g., activities or scenery spots) are relatively limited. For example, in the Hokkaido's official promotion website (see *en.visit-hokkaido.jp*), only three activities (i.e. Ski/Snowboarding, Playing with snow, Drift ice sightseeing) are introduced in winter activities, while ten type of activities (e.g., rafting, golf, trekking) are listed in summer.

Another hypothesis is that the proportion of re-visitors could be larger in winter. Based on Crompton's theory (1979), if the activities such as enjoying skiing in Hokkaido became an annual routine, the pleasure and surprises for vacation could decrease. And that could cause the decrease in motivation to post reviews. To confirm the above hypotheses, further analysis using review contents is necessary in future.

5.4.2.2 Regional posting rate by city

Fig. 5.6 shows the regional posting rate in the sixteen cities recorded with more than one hundred reviewers in total in 2017. Similar as Fig. 5.6, the color is darker when the posting rate is higher in each row. As can be seen, regional posting rate also differs by city. In Hokkaido, Sapporo, or Kutchan had the highest regional number of both overnight travelers and reviewers in 2017; however, regional posting rates were higher in Furano area such as Biei and Nakafurano. In these cities, lavender fields such as Farm Tomita and Shikisai-no-oka, Blue Pond and Shirahige-no-taki Falls occupied most of the reviews. Such nature sceneries could be attractive to international tourists to stimulate their motivation to post reviews.

Regions	Sapporo	Kutchan	Hakodate	Otaru	Noboribetsu	Niseko	Asahikawa	Shimukappu	Toyako	Furano	Chitose	Biei	Kushiro	Abashiri	Nakafurano	Shari
Great Britain	1.325%	2.279%	1.647%	1.587%	1.858%	0.867%	0.768%	1.031%	1.376%	2.113%	0.425%	5.000%	1.272%	2.874%	5.556%	1.575%
Canada	0.955%	2.679%	0.956%	0.884%	1.043%	1.671%	2.278%	0.000%	3.468%	2.344%	0.111%	1.739%	1.418%	1.429%	2.703%	0.699%
Australia	1.207%	1.758%	1.556%	0.621%	1.802%	0.581%	1.504%	2.913%	0.939%	0.000%	0.208%	0.000%	1.016%	1.064%	0.000%	0.742%
France	1.155%	2.083%	1.424%	0.529%	2.128%	0.759%	0.476%	5.000%	1.835%	0.000%	0.287%	4.651%	0.254%	1.316%	0.000%	0.418%
Germany	0.981%	1.018%	1.928%	1.150%	1.171%	0.331%	1.326%	1.589%	1.149%	0.778%	0.688%	3.322%	1.533%	1.937%	3.883%	0.884%
USA	0.696%	1.702%	0.964%	1.582%	0.604%	0.919%	0.631%	0.881%	0.889%	1.264%	0.256%	3.333%	0.817%	0.976%	10.377%	1.219%
Singapore	0.602%	0.851%	0.881%	0.878%	0.735%	0.955%	0.622%	1.280%	0.518%	0.862%	0.530%	2.275%	0.315%	0.895%	3.858%	0.433%
India	0.542%	2.632%	0.180%	3.030%	1.136%	0.000%	0.962%	0.000%		0.000%	0.552%	40.000%	0.000%	4.545%		
Vietnam	0.641%	4.819%	0.763%	0.692%	0.000%	1.190%	0.340%	10.000%	0.000%	0.000%	0.000%	20.000%	0.000%	0.000%		
Philippines	0.497%	0.726%	0.461%	1.412%	0.173%	3.563%	0.683%	0.000%	0.215%	0.529%	0.615%	11.538%	0.000%	0.758%		0.000%
Indonesia	0.327%	0.451%	0.350%	1.405%	0.993%	0.702%	0.498%	0.192%	0.177%	0.244%	0.139%	5.357%	0.000%	0.360%	50.000%	
Russia	0.314%	0.654%	1.172%	0.501%	0.417%	0.383%	0.862%	0.000%	0.000%	0.000%	0.239%	0.000%	0.000%	0.000%		0.000%
Hong Kong	0.245%	0.855%	0.206%	0.640%	0.250%	0.324%	0.347%	0.446%	0.181%	0.684%	0.135%	1.645%	0.212%	0.331%	3.467%	0.281%
Malaysia	0.224%	0.434%	0.536%	0.292%	0.164%	0.238%	0.386%	0.505%	0.264%	0.323%	0.274%	0.613%	0.135%	0.197%	0.814%	0.252%
Thailand	0.132%	0.453%	0.149%	0.207%	0.186%	0.196%	0.074%	0.342%	0.045%	0.108%	0.101%	0.753%	0.093%	0.214%	1.791%	0.091%
Taiwan	0.060%	0.252%	0.147%	0.151%	0.031%	0.060%	0.078%	0.142%	0.015%	0.139%	0.096%	0.183%	0.134%	0.479%	1.848%	0.148%
South Korea	0.059%	0.362%	0.036%	0.227%	0.020%	0.064%	0.170%	0.083%	0.036%	0.218%	0.208%	0.267%	0.021%	0.181%	0.661%	0.049%
MainlandChina	0.016%	0.164%	0.025%	0.030%	0.011%	0.026%	0.018%	0.035%	0.009%	0.027%	0.018%	0.047%	0.003%	0.012%	0.185%	0.042%

Fig. 5.6 Regional posting rate in the sixteen cities in 2017

5.4.3 Limitations and future research directions

Despite of the findings outlined above, this research has many limitations. First of all, in our methodology, we adopted bivariate correlation analysis to examine the relationship between the number of tourist arrivals and the number of reviewers to show it is possible to use the latter to determine the former. However, we didn't provide a method to evaluate the precision of such estimation. Because of the lack of the precise number of tourist arrival, finding the precise precision can be impracticable. However, relative evaluation can be conducted using the latest statistics of tourist arrivals, the predicted results based on the number of reviewers, and common metrics such as MAE or RMSE. By this means, we could find the precision of the estimation compared to traditional surveys.

The same problem also pertains in the evaluation of the estimation at the facility level. In this study, we only examined the relationship between the two variables at the city level. Therefore, to make full use of the quantity of travel reviews, the relationship could be further examined at the facility level with data collected by local governments at each sightseeing spots, festivals and events. Otherwise, estimated values based on travel reviews could only provide general reference in ordinal scales, and should be treated with cautious in ratio or interval scales.

In addition, we only used travel reviews from TripAdvisor and data in Hokkaido. The method itself can be considered transferable. Thus, this method should be applied to other data sources and other areas to further validate our conclusions.

Last but not the least, possible interpretation of the results based on the bivariate analysis alone is limited. Therefore, the analysis of the amount of reviewers should be integrated with analysis of the review contents in search of further insights to questions such as why the data points in Sapporo became outlier values.

5.4.4 Implications

Regarding the investigation of tourist arrivals, both traditional surveys and the analysis of online data suffer from certain limitations outlined above. Thus, to accurately monitoring the trend among tourists, it is important to recognize the limitations of each method, and to understand the nature of the data.

This study contributes in several ways to an understanding of the value and issues of employing online data such as travel reviews into tourism investigations.

First, a method, including the aggregation of the number of reviewers, was devised to examine the relationship between the number of tourist arrivals and the number of reviewers. In our method, we used the information of the username, address, post-date and destination provided with the review. Therefore, this method should be transferrable to other travel websites that provide such information. Also, this method can be applied to other areas that supply the number of monthly sightseeing visitor or overnight travelers. Even more, apart from the number of reviewers, this method can be applied to validate other quantitative indicators from GPS data or those used in content analysis such as the number of travel reviews and the number of reviews containing a certain word if their creators are a distinctive partial of the tourist population.

Next, from the practical point of view, the strong positive correlation between the two examined variables suggests that it is possible to use the number of monthly reviewers in each city to estimate the number of monthly tourist arrivals in each city. In a previous study in this field, Saeki et al. (2015) adopted ranking data and confirmed positive correlation in ordinal scales. In this study, because the number of tourist arrivals and reviewers were adopted directly, liner correlation above ordinal scales was also confirmed.

Also, by defining the posting rate, intriguing research directions were discovered compared to examining the number of tourist arrivals or the number of reviewers independently. Although the interpretation of posting rate requires further analysis with other forms of data, the disagreement suggests that multiple types of data source should be used to extensively understand the behavior of tourists. Furthermore, the differences of regional posting rate in each month or in each city could be resulted by various factors such as the geographical environment in the destinations and the place of residences, the frequency of visit, or psychological factors such as expectation-satisfaction. These possibilities suggest that inbound tourism investigations should not depend solely on the interpretation of results from statistical analysis, but should also incorporate with the geographical or culture differences (e.g., psychology, linguistics) in each country/region.

Additionally, posting rate raises several implications from the managerial point of view. Regarding Hokkaido's inbound tourism, the number of overnight travelers is higher in winter, the number of reviewers is higher in summer or winter, and posting rate is higher in spring or summer. Motivation to post reviews can be influenced by several factors. In the context of tourism, assuming that posting rate is relevant with preferences, higher posting rate in spring or summer may suggest that certain elements during this period are attractive to international visitors. Therefore, the number of international arrivals can be expected to rise during spring and summer provided with appropriate promotions. Meanwhile, promotions of winter should be controlled carefully to avoid excessive expectations. Also, more efforts should be put into the creations of attractions and activities during the winter. Besides, posting rate could be related to the penetration rate of TripAdvisor in a certain consumer group. In that case, spring or summer's Hokkaido may be more attractive to the reviewers group on TripAdvisor. Thus, targeted promotion can be achieved using destination blogs or facility introductions via TripAdvisor.

5.5 Conclusion

This study aimed to show the validity of using the number of reviewers to find the number of tourist arrivals. We presented a method to calculate the number of reviewers on TripAdvisor to be compared to the number of sightseeing visitors and overnight travelers registered in the survey report provided by the Hokkaido government. By examining the correlation and posting rate, the main findings are as followings:

- Pearson's *r* showed strong positive correlation between the number of sightseeing visitors and the number of reviewers in each city in each month. Therefore, it is possible to use the analysis of travel reviews as a low-cost and quicker solution to monitoring tourist' visited places.

- Regarding the inbound tourism, strong positive correlations were found between the number of overnight travelers and the number of reviewers in all eighteen regions of residence in each city in one year. Also, the regional posting rates vary both by month and by city. In addition, for westerners recorded with higher regional posting rate, analyzing travel reviews may compensate the problem of limited samples in traditional survey.

However, our study has a number of limitations, such as limited data sources and limited control of estimating errors. Thus, we should further validate our conclusions with data from other website and other areas. Also, our method can be further incorporated with advanced regression techniques or parameter tuning to reduce the errors.

Chapter 6 Conclusions

6.1 Conclusions

This work aims to show the validity of employing the analysis of online travel reviews for tourist satisfaction and tourist arrivals investigation.

In Chapter 3, to decide whether a certain online data source is suitable for tourism investigation, a content categorization model was created to investigate the percentages of useful information for that purpose. Using travel reviews on TripAdvisor as an example, both manual and automated analysis was applied to find out what do people usually write in travel reviews. By transforming the text in review into categories, information that is useful for finding travel facts, tourists' viewpoints, actions, complaints and compliments can be identified apart from non-needs-related information. Nevertheless, because the percentage of useful information differs by countries/regions, sample size should be selected accordingly. Meanwhile, in the automated analysis based on basic text-mining techniques, some results are consistent with manual analysis, which suggests that it is possible to use automated analysis instead of manual analysis.

In Chapter 4, to determine the validity of identifying tourist needs from online travel reviews through the use of text data mining, a method was presented to manually extract tourist attitudes from reviews to be compared to the satisfaction rates in a traditional survey. The Pearson's r showed (strong) correlations between the attitudes in reviews and the satisfaction rates recorded in the guest survey in six out of seven regions. This finding suggests that it is possible to use the analysis of travel reviews as a low-cost solution to replace traditional surveys. Also, the percentages of positive reviews differ from the satisfaction rates recorded in the guest surveys, whereas the percentages of combined positive and neutral reviews are numerically similar to these satisfaction rates. This finding suggests that survey results can be more positive than they should be, possibly because of the influence of the social desirability bias and the extreme response bias. Besides, tourists with different language background are found to express their satisfaction using different words.

In Chapter 5, to show the validity of using the number of reviewers for finding the number of tourist arrivals, a method was presented to calculate the number of reviewers to be compared to the number of sightseeing visitors and overnight travelers registered in the survey report. As a result, strong positive correlation was found between the number of monthly sightseeing visitors in each city and the number of monthly reviewers in each city in Hokkaido during year 2016 and 2017. This finding suggests that it is possible to use the analysis of travel reviews as a low-cost and quicker solution to monitoring tourist' visited places. Also, regarding the inbound tourism, strong positive correlations were found between the annual number of overnight travelers in each city and the annual number of reviewers in each city in all eighteen regions of residence. Besides, the regional posting rates vary both by month and by city. In addition, westerners are found to be more inclined to provide their information on TripAdvisor than the Asians, and thus analyzing travel reviews may compensate the problem of limited samples in traditional survey.

Therefore, this work would like to believe that the analysis of online travel review is a potential proxy for the current time and money consuming traditional survey methods. Also, it is possible to detect certain biases in results by comparing two distinctive forms of data, in which case, the analysis of online travel review could also serve as an instrument for cross checking the results of traditional survey methods.

6.2 Implications

The results of this research can be useful for both academic and practical field in various ways that concerns the following groups: administrative bodies, researchers, facility owners, and the tourists.

First regarding the methodological contributions, the following three methods were presented, which should be transferrable to other data source, destinations and period of time.

- A content categorization model with three main categories and ten sub categories is created to investigate the percentage of useful information for tourism investigations such as travel facts, tourists' viewpoints, actions, complaints.
- A method was devised to conduct an appropriate comparison between the attitudes expressed in travel reviews and the tourist satisfaction recorded in traditional tourism surveys, based on statistics, social psychology literature, and information technology techniques and was developed through trial-and-error experiments.
- A method, including the aggregation of the number of reviewers, was created to examine the relationship between the number of tourist arrivals and the number of reviewers.

By presenting methods to mine tourist-survey-related information from online travel reviews, this work also contributes to the development of the methodologies for the understanding of the nature of this particular type of data regarding its structure and contents, its embedded attitudes, and its quantities. Researchers could apply these methods to their investigations or extent these methods to other potential tourism statistical data sources. Of course, the utility of travel reviews is not limited to tourism statistics, by identifying what types of data is available and how to retrieve them, this work also helps researchers to make advantage of a particular type of data in their own fields of study.

Regarding the practical contributions, this work showed the possibility of employing the analysis of travel reviews as a low-cost and quicker alternative of the traditional investigation methods in the following the steps: 1) categorizing the content of travel reviews to show that reviews from TripAdvisor contains relevant contents to serve as a potential data source for tourism investigation, 2) tourist satisfaction can be predicted to determine specific areas in which tourists are more (or less) satisfied, and 3) number of reviewers could be utilized to estimate the number of tourist arrivals.

By showing the consistency between those results from travel review analysis and the results from traditional surveys, this work fills the gap between data analysis and tourism surveys, providing a better appreciation of the value of data analysis to administrative bodies, as well as researchers who are dedicated to the utility of data analysis in practical fields. Another group of people who may benefit from this study is the managers and owners of tourism facilities. By recognizing the attitude reflected in travel reviews written by tourists all over the world, facility managers and owners can evaluate their services through the eyes of international tourists, improve the services, and/or foster promotions by geographical segments, which could in turn, improve tourists' travel experience.

Next, by comparing the results of two distinctive methods, the following problem of traditional survey methods were identified.

- Neutrality may be included as satisfaction in survey results. This incorporation may raise concerns about the social desirability bias as survey respondents may shape their answers to please interviewers. It may also elicit the extreme response bias as survey respondents may only select the most extreme options.

From the inconsistency between the results from travel reviews analysis and the results collected from traditional surveys, this work advances the understanding of certain limitations of both two approaches. In this manner, this work helps data analysts to avoid misinterpreting the results acquired by any of these two approaches. And once identified, researchers could develop new instruments to improve or eliminate these methodological faults that could affect the precision of the results.

Besides, as a case study, the following features of the employment of analyzing travel review on TripAdvisor are identified.

- Comments on views are the most, followed by general comments about the whole travel experience (e.g., worth a visit or enjoyable). On the other hand, among the 1,300 reviews, only 33 reviews (2.5%) have clearly identified their travel purposes or motivations. Therefore, sample size should be adjusted accordingly in the investigation of different aspects.
- Some results are consistent with manual analysis in the automated analysis based on basic text-mining techniques, which suggests that it is possible to use automated analysis instead of manual analysis.
- Inevitable bias exists in any sentiment analysis that involves manually annotated data, because even when two individuals fully understand each other's logic and reasoning, they would still not agree with each other because of attitudinal differentials.

- The consistency and/or kappa-value between two individuals could be used to evaluate the results of an automated analysis. That is, if a program can reach similar or higher consistency and kappa-value, we can consider it an individual intelligence and conclude that its performance is as good as that of a non-expert annotator.
- Regarding the estimation of tourist arrivals, the use of travel reviews could yield results that are more sensitive to sudden events.
- Seasonal parameter tuning could also be utilized to improve the precision of estimating tourist arrivals.
- Because the westerners recorded with higher regional posting rate, it is expected that the analysis of travel reviews could compensate the problem of limited samples, especially in countries/regions with less arrivals from western counties.

In addition, the behavior of posting reviews, the content of reviews, and the attitudes embedded in reviews vary by language and/or by countries/regions. To be specific, the following insights were found regarding cross-region and cross-language comparisons, suggesting that inbound tourism investigations should not depend solely on the interpretation of results from statistical analysis, but should also incorporate with the geographical or culture differences (e.g., psychology, linguistics) in each country/region.

- In travel reviews, the Americans wrote more about cultures; Australian and Singaporean wrote more about food; British wrote more about price and services, and little about access; Chinese wrote more about views and less about activities. Since the percentage of useful information differs between countries, sample size should be selected accordingly.
- Although the presence of emotional words or higher ratings may indicate tourist satisfaction within a particular country or language group, we must note that the differences recorded in the statistical distributions may not be equivalent to the disparity in the true feelings of tourists because the expression of satisfaction differs from language to language.
- The proportions of reviewers in TripAdvisor to overnight travelers differ by regions of residence. British recorded the highest among all regions, followed by the Canadian, Australian, France, German, and American. Meanwhile, posting rates in Asian countries/regions are comparatively low.

6.3 Limitations and directions for future research

Despite the findings outlined above, this study suffers from certain limitations.

First, regarding the generality of the conclusions, two methods, one for the comparisons of tourist satisfaction in Chapter 4 and one for the comparisons of tourist arrivals in

Chapter 5 were presented, but these methods were only applied to travel reviews collected from one travel website, TripAdvisor, and posted in one destination, Hokkaido, during a limited period of time. Therefore, further validation is needed to test the generality of the conclusions.

Moreover, there are a plenty of rooms for further improvements of the methods. For example, other statistical metrics for testing, the incorporation with automated analysis or advanced data analysis techniques could be considered. Moreover, as is repeated in this work, the credibility of travel reviews is inevitably limited, especially those collected commercial website wherein some of the negative reviews are reported to be removed to please the owners of the facilities. Therefore, a non-profitable platform is needed to collect tourists' reviews. Techniques that can detect fake reviews are also desired.

Tourist satisfaction and tourist arrivals are only a partial of the tourism investigation. It is possible that the analysis of travel reviews could be utilized in other aspects. For example, in respect of tourist profiling, information about gender, age group, travel companions or nationalities could be presumed from the context of the comment (Fujii et al., 2017) or photos. Comparing the distributions of tourist profiling between the travel reviews and traditional survey results can possibly explain some of the inconsistencies between the two forms of data, which should be an interesting research direction.

In addition, the estimation of tourist arrivals and the investigation of tourist satisfaction based on online travel review analysis could be integrated into one framework to provide insights for problem solving. For example, potential problematic districts/facilities could be identified based on unusual readings of the number reviewers, and content analysis could then be applied to identifying the problematic aspects in services.

As a low-cost approach, the analysis of online travel review has its advantage over the inbound tourism investigation to help understand the needs of tourists who are coming from different countries/regions. Also, on travel websites, managers and owners of tourism facilities actively volunteer introduction information about their properties to reach more costumers. Because those facilities are grouped according to administrative divisions on travel websites, it is possible to evaluate the resources in a destination, which can be considered as another research direction (Song et al., 2018).

Ultimately, as the third direction, by optimizing the needs of tourists and the resources in one destination, the results of this study will not only be useful the governmental bodies and related researchers, but also extendable to the tourists, the destination, and the researchers who are interested in destination recommendation.

Acknowledgements

The author expresses to thank Professor Hidenori Kawamura of Hokkaido University for instruction and guidance from the inception and the completion of the present research.

Deep thanks are also due to Professor Hajime Saito of Hokkaido Information University for years of profitable discussions and suggestions.

She wishes to thank Associate Professor Tomohisa Yamashita and Assistant Professor Soichiro Yokoyama of Hokkaido University and Professor Junichi Uchida of Otaru University of Commerce for valuable discussions and advice.

Many pieces of affectionate advice and discussion given by them were very useful to improve this research.

She also wishes to express sincere gratitude to her friends and students of Hajime Saito Seminar for the participants in the experiments, which enabled this study to be conducted.

She would like to thank all students of the office, Harmonious System Engineering, Complex System Engineering. The weekly presentations often inspired her mind.

Finally, she appreciates all that her family, especially for everything her mother Xiaomei Xu had done to support her life.

Bibliography

Aksu, A., İçigen, E. T., & Ehtiyar, R. (2010). A Comparison of Tourist Expectations and Satisfaction: A Case Study from Antalya Region. Turizam, 14(2), 66-77.

Amaro, S., & Duarte, P. (2017). Social media use for travel purposes: a cross cultural comparison between Portugal and the UK. Information Technology & Tourism, 17(2), 161-181.

Antonio, N., de Almeida, A., Nunes, L. et al. (2018). Hotel online reviews: different languages, different opinions. Inf Technol Tourism. 18, 157–185. DOI: 10.1007/s40558-018-0107-x

Ayeh, J. K., Au, N., & Law, R. (2013). "Do we believe in TripAdvisor?" Examining credibility perceptions and online travelers' attitude toward using user-generated content. Journal of Travel Research, 52(4), 437-452.

Bahtar, A. Z., & Muda, M. (2016). The impact of User–Generated Content (UGC) on product reviews towards online purchasing–A conceptual framework. Procedia Economics and Finance, 37, 337-342.

Baka, V. (2016). The becoming of user-generated reviews: Looking at the past to understand the future of managing reputation in the travel sector. Tourism Management, 53, 148-162.

Bakir, A., Wickens, E., Boluk, K., & Avgeli, V. (2017). Cross-cultural issues in tourism and hospitality. Tourism and Hospitality Research, 17(3), 245-248.

Banerjee, S., & Chua, A. Y. (2016). In search of patterns among travelers' hotel ratings in TripAdvisor. Tourism Management, 53, 125-131.

Beldona, S., Morrison, A. M., & O'Leary, J. (2005). Online shopping motivations and pleasure travel products: a correspondence analysis. Tourism Management, 26(4), 561-570.

Berezina, K., Bilgihan, A., Cobanoglu, C., & Okumus, F. (2016). Understanding satisfied and dissatisfied hotel customers: text mining of online hotel reviews. Journal of Hospitality Marketing & Management, 25(1), 1-24.

Boley, B. B., Magnini, V. P., & Tuten, T. L. (2013). Social media picture posting and souvenir purchasing behavior: Some initial findings. Tourism Management, 37, 27-30.

Bowen, D., & Clarke, J. (2002). Reflections on tourist satisfaction research: Past, present and future. Journal of Vacation Marketing, 8, 297-308. DOI: 10.1177/135676670200800401

Brakel, Jan & Söhler, Emily & Daas, P.J.H. & Buelens, Bart. (2017). Social media as a data source for official statistics; the Dutch Consumer Confidence Index. Survey methodology. 43, 183-210.

Bronner, F., & De Hoog, R. (2010). Consumer-generated versus marketer-generated websites in consumer decision making. International Journal of Market Research, 52(2), 231-248.

Buzova, D., Sanz-Blas, S., & Cervera-Taulet, A. (2019) Does culture affect sentiments expressed in cruise tours' eWOM?, The Service Industries Journal, 39(2), 154-173, DOI: 10.1080/02642069.2018.1476497

Cabral, L., & Hortacsu, A. (2010). The dynamics of seller reputation: Evidence from eBay. The Journal of Industrial Economics, 58(1), 54-78.

Campbell, D. T. & Fiske, D. W. (1967) Convergent and discriminant validation by the multi trait-multi method matrix, in: W. A. MEHRENS & R. L. EBEL (Eds) Principles of Educational and Psychological Measurement, Chicago, Rand McNally.

Chaves, M. S., Gomes, R., & Pedron, C. (2012). Analysing reviews in the Web 2.0: Small and medium hotels in Portugal. Tourism Management, 33(5), 1286-1287.

Cheung, C. M., & Lee, M. K. (2012). What drives consumers to spread electronic word of mouth in online consumer-opinion platforms. Decision support systems, 53(1), 218-225.

Cheung, M. Y. (2006). Do people believe electronic word-of-mouth?: a study on factors affecting readers' perceived credibility of online consumer reviews, Doctoral dissertation, City University of Hong Kong.

Chiba Prefectural Government (2016): Tourists needs and arrivals among foreigners using SNS data (SNS を活用した外国人観光客ニーズ・動向調査結果につ

いて), https://www.pref.chiba.lg.jp/kankou/sns/documents/sns-tyousakekka.pdf (in Japanese) [9/17/2019]

Choi, J., Peters, M., & Mueller, R. O. (2010). Correlational analysis of ordinal data: from Pearson's r to Bayesian polychoric correlation. Asia Pacific Education Review, 11(4), 459-466.

Chung, J. Y., & Buhalis, D. (2008). Information needs in online social networks. Information Technology & Tourism, 10(4), 267-281.

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

Crompton, J. L. (1979). Motivations for pleasure vacation. Annals of tourism research, 6(4), 408-424.

Crowston, K., & Fagnot, I. (2018). Stages of motivation for contributing user-generated content: A theory and empirical test. International Journal of Human-Computer Studies, 109, 89-101.

Cser, K. (2008): Study on survey and conversion methodology of world hotel classification system. Doctoral dissertation. Hokkaido University.

Daas, P. J., Puts, M. J., Buelens, B., & van den Hurk, P. A. (2013). Big data and official statistics. Proceedings of the NTTS (New Techniques and Technologies for Statistics), 5-7.

Daas, P. J., Puts, M. J., Buelens, B., & van den Hurk, P. A. (2015). Big data as a source for official statistics. Journal of Official Statistics, 31(2), 249-262.

Daugherty, T., Eastin, M. S., & Bright, L. (2008). Exploring consumer motivations for creating user-generated content. Journal of interactive advertising, 8(2), 16-25.

Davis, E., Greenberger, E., Charles, S., Chen, C., Zhao, L., & Dong, Q. (2012). Emotion experience and regulation in China and the United States: How do culture and gender shape emotion responding?. International Journal of Psychology, 47(3), 230-239.

Del Chiappa, G., & Dall'Aglio, S. (2012). Factors influencing travelers'e-ratings and e-complaints about hotel services: Insights from an Italian tourism destination. In

Information and Communication Technologies in Tourism 2012. 448-459. Springer, Vienna.

Dickinger, A., & Költringer, C. (2011). Extracting destination representation and competitiveness from online content. Information Technology & Tourism, 13(4), 327-339.

Dickinger, A., & Lalicic, L. (2016). An analysis of destination brand personality and emotions: a comparison study. Information Technology & Tourism, 15(4), 317-340.

Dippelreiter, B., Grün, C., Pöttler, M., Seidel, I., Berger, H., Dittenbach, M., & Pesenhofer, A. (2008). Online tourism communities on the path to Web 2.0: an evaluation. Information technology & tourism, 10(4), 329-353.

Doh, S. J., & Hwang, J. S. (2009). How consumers evaluate eWOM (electronic word-of-mouth) messages. CyberPsychology & Behavior, 12(2), 193-197.

Ehime Prefecture (2017): Basic plan for tourism promotion in Ehime (愛媛県観光振興基本計画 ビッグデータを活用した観光戦略等検討委託業務結果（その 1-3）) https://www.pref.ehime.jp/h30200/documents/bigdata1.pdf, https://www.pref.ehime.jp/h30200/documents/bigdata2.pdf, https://www.pref.ehime.jp/h30200/documents/bigdata3.pdf (in Japanese) [9/17/2019]

Erkan, I., & Evans, C. (2018). Social media or shopping websites? The influence of eWOM on consumers' online purchase intentions. Journal of Marketing Communications, 24(6), 617-632.

European Commission (2016): European Tourism Indicators System for sustainable destination management, https://ec.europa.eu/growth/sectors/tourism/offer/sustainable/indicators_en [access: 9/17/2019]

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.

Fernández, I., Carrera Levillain, P., Sánchez Fernández, F., Paez, D., & Candia, L. (2000). Differences between cultures in emotional verbal and non-verbal reactions. Psicothema. 12, 83-92

Ferrer-Rosell, B., Coenders, G. & Marine-Roig, E. (2017). Is planning through the Internet (un)related to trip satisfaction?. Inf Technol Tourism, 17(2), 229-244. DOI: 10.1007/s40558-017-0082-7

Fesenmaier, D. R., Leppers, A. W., & O'Leary, J. T. (1999). Developing a knowledge-based tourism marketing information system. Information Technology & Tourism, 2(1), 31-44.

Fisher, R. J. (1993). Social desirability bias and the validity of indirect questioning. Journal of consumer research, 20(2), 303-315.

Fowler Jr, F. J., & Mangione, T. W. (1990). Standardized Survey Interviewing: Minimizing Interviewer-Related Error (Applied Social Research Methods Series Vol. 18). Sage.

Fu, J. R., Ju, P. H., & Hsu, C. W. (2015). Understanding why consumers engage in electronic word-of-mouth communication: Perspectives from theory of planned behavior and justice theory. Electronic Commerce Research and Applications, 14(6), 616-630.

Fujii, K., Nanba, H., Takezawa, T., Ishino, A., Okumura, M. & Kurata, Y. (2017) Automatic Identification of Attributes of Travel Blog Entries for Travelers'Behaviour Analysis (in Japanese). Society for Tourism Informatics, 13(1), 83-96.

Fukushima Government (2019): Fukushima Destination Survey (福島県観光地実態調査報告書). https://www.pref.fukushima.lg.jp/uploaded/attachment/343173.pdf [access:9/17/2019] (in Japanese).

García-Pablos, A., Cuadros, M., & Linaza, M. T. (2016). Automatic analysis of textual hotel reviews. Information Technology & Tourism, 16(1), 45-69.

Gayo-Avello, D. (2013). A meta-analysis of state-of-the-art electoral prediction from Twitter data. Social Science Computer Review, 31(6), 649-679.

Goh, K. Y., Heng, C. S., & Lin, Z. (2013). Social media brand community and consumer behavior: Quantifying the relative impact of user-and marketer-generated content. Information Systems Research, 24(1), 88-107.

González, G., López, B., & LLUÍS, J. (2003). Smart user models for tourism: A holistic approach for personaliaed tourism services. Inf. Technology & Tourism, 6(4), 273-286.

Govers, R., & Go, F. M. (2004). Projected destination image online: Website content analysis of pictures and text. Inf. Technology & Tourism, 7(2), 73-89.

Granello, D. H., & Wheaton, J. E. (2004). Online data collection: Strategies for research. Journal of Counseling & Development, 82(4), 387-393.

Gretzel, U., Mitsche, N., Hwang, Y. H., & Fesenmaier, D. R. (2004). Tell me who you are and I will tell you where to go: Use of travel personalities in destination recommendation systems. Information Technology & Tourism, 7(1), 3-12.

Hammersley, M. (1987). Some notes on the terms 'validity'and 'reliability'[1]. British Educational Research Journal, 13(1), 73-82.

Hatoh, Y., Wu, Y., Shimamura, K. & Takeuchi, K. (2013). Comparison of Viewpoints between Japanese and Chinese Review in Tourism (in Japanese). IEICE technical report. AI, Artificial Intelligence and Knowledge-Based Processing, 112(435), 1-4.

He, P., Lai, J., Chou, H. M., & Zhou, L. (2012, January). A cross-cultural comparison of online customer reviews. In Global Conference on Business and Finance Proceedings, 7(1), 122-133.

Hennig-Thurau, T., Gwinner, K. P., Walsh, G., & Gremler, D. D. (2004). Electronic word-of-mouth via consumer-opinion platforms: what motivates consumers to articulate themselves on the internet?. Journal of interactive marketing, 18(1), 38-52.

Hokkaido Government (2016): Survey Concerning Customer Satisfaction. http://www.pref.hokkaido.lg.jp/kz/kkd/H28doutai_home.htm [access: 8/4/2017] (in Japanese)

Hokkaido Government (2019): Report on tourism arrivals in Hokkaido, http://www.pref.hokkaido.lg.jp/kz/kkd/irikomi.htm (in Japanese). (Access: 2019.06.23)

Holt, R. (2002). Faces of the “New China”: a comparison of touristic web sites in the Chinese and English languages. Information Technology & Tourism, 5(2), 105-119.

Hovland, C.I., Janis, I. L., & Kelley, H.H., Communication and persuasion. 1953.

Hu, Y. H., Chen, Y. L., & Chou, H. L. (2017). Opinion mining from online hotel reviews–a text summarization approach. Information Processing & Management, 53(2), 436-449.

Huang, S. S., & Crotts, J. (2019). Relationships between Hofstede's cultural dimensions and tourist satisfaction: A cross-country cross-sample examination. Tourism Management, 72, 232-241.

Huertas, A., & Marine-Roig, E. (2016). User reactions to destination brand contents in social media. Information Technology & Tourism, 15(4), 291-315.

Iinuma, S., Nanba, H., & Takezawa, T.(2017). Automatic summarization of multiple travel blog entries focusing on traveler's behavior(in Japanese). Society for Tourism Informatics, 13(1), 53-68

Inversini, A., & Eynard, D. (2011). Harvesting user-generated picture metadata to understand destination similarity. Information Technology & Tourism, 13(4), 341-351.

Ismagilova, E., Dwivedi, Y. K., & Slade, E. (2019). Perceived helpfulness of eWOM: Emotions, fairness and rationality. Journal of Retailing and Consumer Services.

Jannach, D., Zanker, M., & Fuchs, M. (2014). Leveraging multi-criteria customer feedback for satisfaction analysis and improved recommendations. Information Technology & Tourism, 14(2), 119-149.

Japan Tourism Agency (JTA) (2013): General Information of Tourism Statistics in Japan, http://www.mlit.go.jp/common/000991725.pdf (Access: 2019/7/6)

Japan Tourism Agency (JTA) (2015): Regional Tourism Economic Survey (観光地域経済調査), https://www.mlit.go.jp/kankocho/siryou/toukei/kouzou.html (in Japanese). (Access: 2019.06.23)

Japan Tourism Agency (JTA) (2017): Tourism Promotion based on Tourism Big Data, http://www.mlit.go.jp/kankocho/shisaku/kankochi/gps.html (in Japanese). (Access: 2019.06.23)

Japan Tourism Agency (JTA) (2019a): Program Plans for Tourism Vision 2019 (観光ビジョン実現プログラム 2019), http://www.mlit.go.jp/kankocho/topics02_000170.html (in Japanese). (Access: 2019.06.23)

Japan Tourism Agency (JTA) (2019b): Common Standard for Statistics on Inbound Tourists and Survey Procedure (共通基準による観光入込客統計), http://www.mlit.go.jp/kankocho/siryou/toukei/irikomi.html (in Japanese). (Access: 2019.06.23)

Japan Tourism Agency (JTA) (2019c): Accommodation Survey (宿泊旅行統計調査), https://www.mlit.go.jp/kankocho/siryou/toukei/shukuhakutoukei.html (in Japanese). (Access: 2019.06.23)

Japan Tourism Agency (JTA) (2019d): National Tourism Survey (旅行・観光消費動向調査), https://www.mlit.go.jp/kankocho/siryou/toukei/shouhidoukou.html (in Japanese). (Access: 2019.06.23)

Japan Tourism Agency (JTA) (2019e): International Visitor Survey, https://www.mlit.go.jp/kankocho/en/siryou/toukei/syouhityousa.html. (Access: 2019.06.23)

Jung, H. S. (1999). The analysis of demographic profiles and prospects of Internet users in national tourism organizations. Case study: Korea National Tourism Organization (KNTO). Information Technology & Tourism, 2(2), 131-138.

Jung, T. H., Lee, H., Chung, N., & tom Dieck, M. C. (2018). Cross-cultural differences in adopting mobile augmented reality at cultural heritage tourism sites. International Journal of Contemporary Hospitality Management, 30(3), 1621-1645.

Kawakita, J. (1986). KJ method. Chuokoron-sha, Tokyo.

Kim, H., & Stepchenkova, S. (2017). Understanding destination personality through visitors' experience: A cross-cultural perspective. Journal of Destination Marketing & Management, 6(4), 416-425.

Kim, S. S., Prideaux, B., & Kim, S. H. (2002). A cross-cultural study on casino guests as perceived by casino employees. Tourism management, 23(5), 511-520.

Kim, T. T., Kim, W. G., & Kim, H. B. (2009). The effects of perceived justice on recovery satisfaction, trust, word-of-mouth, and revisit intention in upscale hotels. Tourism management, 30(1), 51-62.

Kliček, B. (2000). Design of a multilevel intelligent decision support system for the improvement of tourist satisfaction. Information technology in Tourism, 5, 55-67.

Koh, N. S. (2011). The valuation of user-generated content: a structural, stylistic and semantic analysis of online reviews. Dissertation.

Krawczyk, M., & Xiang, Z. (2016). Perceptual mapping of hotel brands using online reviews: a text analytics approach. Information Technology & Tourism, 16(1), 23-43.

Kurashima, T., Fujimura, K., & Ukuda, H. (2009). Mining Experience from Large-scale Blog Entries. IEICE Transactions on Fundamentals of Electronics, Communications and Computer Sciences D, 92(3), 301-310. (in Japanese)

Kusumasondjaja, S., Shanka, T., & Marchegiani, C. (2012). Credibility of online reviews and initial trust: The roles of reviewer's identity and review valence. Journal of Vacation Marketing, 18(3), 185-195.

Lee Rodgers, J., & Nicewander, W. A. (1988). Thirteen ways to look at the correlation coefficient. The American Statistician, 42(1), 59-66.

Lee, G., & Tussyadiah, I. P. (2010). Textual and visual information in eWOM: a gap between preferences in information search and diffusion. Information technology & tourism, 12(4), 351-361.

Lee, G., Cai, L. A., & O'Leary, J. T. (2006). WWW. Branding. States. US: an analysis of brand-building elements in the US state tourism websites. Tourism Management, 27(5), 815-828.

Lee, H. A., Law, R., & Murphy, J. (2011). Helpful reviewers in TripAdvisor, an online travel community. Journal of Travel & Tourism Marketing, 28(7), 675-688.

Lefever, S., Dal, M., & Matthiasdottir, A. (2007). Online data collection in academic research: advantages and limitations. British Journal of Educational Technology, 38(4), 574-582.

Levy, S. E., Duan, W., & Boo, S. (2013). An analysis of one-star online reviews and responses in the Washington, DC, lodging market. Cornell Hospitality Quarterly, 54(1), 49-63.

Li, M. (2014). Cross-cultural tourist research: A meta-analysis. Journal of Hospitality & Tourism Research 38.1: 40-77.

Lim, Y., Chung, Y., & Weaver, P. A. (2012). The impact of social media on destination branding: Consumer-generated videos versus destination marketer-generated videos. Journal of Vacation Marketing, 18(3), 197-206.

Litvin, S. W., Goldsmith, R. E., & Pan, B. (2008). Electronic word-of-mouth in hospitality and tourism management. Tourism management, 29(3), 458-468.

Liu, Y., Teichert, T., Hu, F., & Li, H. (2017). How do tourists evaluate Chinese hotels at different cities? Mining online tourist reviewers for new insights. In WHICEB (p. 67).

Liu, Y., Teichert, T., Rossi, M., Li, H., & Hu, F. (2017). Big data for big insights: Investigating language-specific drivers of hotel satisfaction with 412,784 user-generated reviews. Tourism Management, 59, 554-563.

Loh, S., Lorenzi, F., Saldaña, R., & Licthnow, D. (2003). A tourism recommender system based on collaboration and text analysis. Information Technology & Tourism, 6(3), 157-165.

Lu, A. C. C., & Chen, B. T. (2014). Information search behavior of independent travelers: A cross-cultural comparison between Chinese, Japanese, and American travelers. Journal of Hospitality Marketing & Management, 23(8), 865-884.

Lu, B., Ott, M., Cardie, C., & Tsou, B. K. (2011, December). Multi-aspect sentiment analysis with topic models. In Data Mining Workshops (ICDMW), 2011 IEEE 11th International Conference. 81-88. IEEE.

Lu, W., & Stepchenkova, S. (2012). Ecotourism experiences reported online: Classification of satisfaction attributes. Tourism management, 33(3), 702-712.

Mariani, M. M., Di Felice, M., & Mura, M. (2016). Facebook as a destination marketing tool: Evidence from Italian regional Destination Management Organizations. Tourism Management, 54, 321-343.

Mariani, M., Di Fatta, G., & Di Felice, M. (2018). Understanding customer satisfaction with services by leveraging big data: the role of services attributes and consumers' cultural background. IEEE Access, 7, 8195-8208.

Marine-Roig, E., & Clavé, S. A. (2016). A detailed method for destination image analysis using user-generated content. Information Technology & Tourism, 15(4), 341-364.

Marrese-Taylor, E., Velásquez, J. D., Bravo-Marquez, F., & Matsuo, Y. (2013). Identifying customer preferences about tourism products using an aspect-based opinion mining approach. Procedia Computer Science, 22, 182-191.

Martínez-Navarro, J., & Bigné, E. (2017). The value of marketer-generated content on social network sites: media antecedents and behavioral responses. Journal of Electronic Commerce Research, 18(1), 52.

Matsumura, A. Sanseido Books (2006) Daijirin Japanese Dictionary (3rd edition). (in Japanese)

Maurer C., & Schaich S. (2011) Online Customer Reviews Used as Complaint Management Tool. In: Law R., Fuchs M., Ricci F. (eds) Information and Communication Technologies in Tourism 2011. Springer, Vienna

McDonald, J.H. (2014). Handbook of Biological Statistics (3rd ed.). Sparky House Publishing, Baltimore, Maryland. (also see http://www.biostathandbook.com/)

Metzger, M. J. (2007). Making sense of credibility on the Web: Models for evaluating online information and recommendations for future research. Journal of the American Society for Information Science and Technology, 58(13), 2078-2091.

Metzger, M. J., & Flanagin, A. J. (2013). Credibility and trust of information in online environments: The use of cognitive heuristics. Journal of Pragmatics, 59, 210-220.

Morrisonn, A. M., Jing, S., O'Leary, J. T., & Cai, L. A. (2001). Predicting usage of the Internet for travel bookings: An exploratory study. Information Technology & Tourism, 4(1), 15-30.

Murakami, K. & Kawamura, H. (2011). Travel in Japan from Foreigners' Perspectives: Analysis of Destination Image Through English Blogs. Journal of Japanese Society for Artificial Intelligence, 26(3), 286-293. (in Japanese)

Nakajima, H., Niitsuma, H. & Ohta, M.(2013). Travel route recommendation using tweets with location information. IPSJ SIG Techinical Report. 28, 1-6. (in Japanese)

Nakajima, Y., & Ohta, M. (2013). Extraction of Torist Spots and Their Related Information from Travel Blog Entries (in Japanese). DEIM Forum 2013 B8-4

Nakayama, M., & Wan, Y. (2019). Same sushi, different impressions: a cross-cultural analysis of Yelp reviews. Information Technology & Tourism. DOI: 10.1007/s40558-018-0136-5

Nakayama, M., & Wan, Y. (2019). The cultural impact on social commerce: A sentiment analysis on Yelp ethnic restaurant reviews. Information & Management, 56(2), 271-279.

Neidhardt, J., Rümmele, N., & Werthner, H. (2017). Predicting happiness: user interactions and sentiment analysis in an online travel forum. Information Technology & Tourism, 17(1), 101-119.

Nightley Inc. (2019): inbound insight, https://nightley.jp/ (in Japanese). (Access: 2019.06.23)

Niigata Prefectural Government (2019): Attitudes towards Niigata among the residents in Tokyo and Kansai (首都圏・関西圏を対象とした本県観光に対する意識調査), https://www.pref.niigata.lg.jp/sec/kankokikaku/1356880603332.html (in Japanese) [9/17/2019]

Nomori, K. & Kozu, T. (2019): Topic analysis based on Travel Reviews and its Utility for Regional Tourism Promotion. Serviceology: magazine of Society for Serviceology 2(2):1-12 (in Japanese).

Oi, T. (2012): Seasonal Analysis based on Accommodation Survey. https://www.mlit.go.jp/common/000193010.pdf (in Japanese) (Access: 2019.06.23)

Okayama Government (2010): Okayama Tourist Satisfaction Survey (平成 25 年度岡山県観光客満足度調査報告書). http://www.pref.okayama.jp/uploaded/life/92691_2267581_misc.pdf [access:8/4/2017] (in Japanese).

Okinawa Prefectural Government (2019): H30 Foreign Tourism Statistic Report, https://www.pref.okinawa.jp/site/bunka-sports/kankoseisaku/kikaku/report/inbound_survey_report/documents/20190719_h30_foreign_tourism_statistic_report_review.pdf (in Japanese) [9/17/2019]

Oliver, R. L. (1980). A Cognitive Model of the Antecedents and Consequences of Satisfaction Decisions. In Journal of Marketing Research, 17(4), 460-469

Olston, C., & Najork, M. (2010). Web crawling. Foundations and Trends® in Information Retrieval, 4(3), 175-246.

Omar, M. B., Shariff, S. N. F. B. A., Adilin, H., Majid, B. M. A., & Sulong, S. N. B (2017). Tourist Fulfilment and Revisit Intention Antecedent of Culinary Experience. 2nd Business Management and Computing Research Colloquium, 63-66.

Ontario Ministry of Tourism, Culture and Sport. (2006). Travel activities and motivations survey. http://www.mtc.gov.on.ca/en/research/travel_activities/tams.shtml[access: 8/4/2017]

Opdenakker, R. R. (2006). Advantages and disadvantages of four interview techniques in qualitative research. In Forum: Qualitative Social Research, 7, 4, art-11.

Orgad, S. (2009). How can researchers make sense of the issues involved in collecting and interpreting online and offline data?.

Oxford University Press (2019). English Oxford Living Dictionaries. https://en.oxforddictionaries.com

Özdemir, C., & Yolal, M. (2017). Cross-cultural tourist behavior: An examination of tourists' behavior in guided tours. Tourism and Hospitality Research, 17(3), 314-324.

Pandey, S., Dhamdhere, K., & Olston, C. (2004, August). WIC: A general-purpose algorithm for monitoring web information sources. In Proceedings of the Thirtieth international conference on Very large data bases-Volume 30. 360-371. VLDB Endowment.

Pang, B., & Lee, L. (2008). Opinion mining and sentiment analysis. Foundations and Trends® in Information Retrieval, 2(1–2), 1-135.

Pang, B., Lee, L., & Vaithyanathan, S. (2002, July). Thumbs up?: sentiment classification using machine learning techniques. In Proceedings of the ACL-02 conference on Empirical methods in natural language processing. Association for Computational Linguistics. 10, 79-86.

Parrott, W. G. (Ed.). (2001). Emotions in social psychology: Essential readings. Psychology Press.

Petrillo, M., & Baycroft, J. (2010). Introduction to manual annotation. Fairview Research.

Plank, A. (2016). The hidden risk in user-generated content: An investigation of ski tourers' revealed risk-taking behavior on an online outdoor sports platform. Tourism management, 55, 289-296.

Plutchik, R. (1960). The multifactor-analytic theory of emotion. Journal of Psychology, 50(1), 153-171.

Poch, R., & Martin, B. (2015). Effects of intrinsic and extrinsic motivation on user-generated content. Journal of Strategic Marketing, 23(4), 305-317.

Procaccino, J. D., & Miller, F. R. (1999). TOURISM ON THE WORLD WIDE WEB: A COMPARISON OF WEB SITES OF UNITED STATES-AND FRENCH-BASED BUSINESSES. Information Technology & Tourism, 2(3), 173-183.

Qu, H., & Lee, H. (2011). Travelers' social identification and membership behaviors in online travel community. Tourism management, 32(6), 1262-1270.

Rabjohn, N., Cheung, C. M., & Lee, M. K. (2008, January). Examining the perceived credibility of online opinions: information adoption in the online environment. In Proceedings of the 41st Annual Hawaii International Conference on System Sciences (HICSS 2008). 286-286. IEEE.

Racherla, P., & Friske, W. (2012). Perceived 'usefulness' of online consumer reviews: An exploratory investigation across three services categories. Electronic Commerce Research and Applications, 11(6), 548-559.

Ravi, K., & Ravi, V. (2015). A survey on opinion mining and sentiment analysis: tasks, approaches and applications. Knowledge-Based Systems, 89, 14-46.

Researchworks (2017): Response rate of mail survey and improve, http://www.researchworks.co.jp/service/tips/response_rate/(access:2017/10/10) (in Japanese)

Richardson, S. L., & Crompton, J. L. (1988). Latent demand for vacation travel: A cross- cultural analysis of French- and English- speaking residents of Ontario and Quebec. Leisure Sciences, 10(1), 17-26.

Rossetti, M., Stella, F., & Zanker, M. (2016). Analyzing user reviews in tourism with topic models. Information Technology & Tourism, 16(1), 5-21.

Saeki, K., Endo, M., Hirota, M., Kurata, Y., & Ishikawa, H. (2015). Language-Specific Analysis of Domestic Places Visited by Foreign Tourists using Crawled Twitter Data (in Japanese). Society for Tourism Informatics, 11(1), 45-56.

Saito, H. (2011). Analysis of Tourism Informatics on Web. Journal of Japanese Society for Artificial Intelligence, 26(3), 234-239. (in Japanese)

Schmallegger, D., & Carson, D. (2009). Destination image projection on consumer-generated content websites: A case study of the Flinders Ranges. Information Technology & Tourism, 11(2), 111-127.

Schmid, H. (1994): Probabilistic Part-of-Speech Tagging Using Decision Trees. Proceedings of International Conference on New Methods in Language Processing, Manchester, UK.

Schmid, H. (1995): Improvements in Part-of-Speech Tagging with an Application to German. Proceedings of the ACL SIGDAT-Workshop. Dublin, Ireland.

Schouten, K., & Frasincar, F. (2016). Survey on aspect-level sentiment analysis, IEEE Transactions on Knowledge and Data Engineering, 28(3), 813-830

Serna, A., Gerrikagoitia, J. K., & Alzua, A. (2013). Towards a better understanding of the cognitive destination image of Euskadi-Basque Country based on the analysis of UGC. In Information and communication technologies in tourism 2014. 395-407. Springer, Cham.

Serna, A., Gerrikagoitia, J. K., & Bernabé, U. (2016). Discovery and classification of the underlying emotions in the user generated content (UGC). In Information and communication technologies in tourism 2016. 225-237. Springer, Cham.

Shan, Y. (2016). How credible are online product reviews? The effects of self-generated and system-generated cues on source credibility evaluation. Computers in Human Behavior, 55, 633-641.

Shao, J., Chang, X., & Morrison, A. M. (2017). How Can Big Data Support Smart Scenic Area Management? An Analysis of Travel Blogs on Huashan. Sustainability, 9(12),2291.

Sheldon, P. J., & Fox, M. (1988). The Role Of Foodservice I n Vacation Choice And Experience: A Cross-Cultural Analysis. Journal of Travel Research, 27(2), 9-15.

Shibutani, H., Watanabe, S., Yoshimura, H. Kokubo, A., Kashiwaya, I., Sasaki, T., Nakamura, K., & Kihara, H.(2015). A direct comparison between a web survey and a mail survey: comparing characteristics of respondents and self-efficacy scores based on the same sample. Journal of Aomori University Multidisciplinary Research Institute. 17(1), 1-22. (in Japanese)

Shimane Prefectural Government (2019): Survey on tourism awareness towards Shimane (しまねの観光認知度調査), https://www.pref.shimane.lg.jp/tourism/tourist/kankou/chosa/ninchido.html (in Japanese) [9/17/2019]

Shuang Song, Tomohisa Yamashita, Hidenori Kawamura, Junichi Uchida, Hajime Saito : Study on the survey and matching of destination and tourist models: Introducing intrinsic travel motives to recommender system, Proceedings of the18th Workshop of the Society for Tourism Informatics, pp.35-38, Hiroshima(2018)

Simmons, L. L., Mukhopadhyay, S., Conlon, S., & Yang, J. (2011). A computer aided content analysis of online reviews. Journal of Computer Information Systems, 52(1), 43-55.

Soldatenko, D., & Backer, E. (2019). A content analysis of cross-cultural motivational studies in tourism relating to nationalities. Journal of Hospitality and Tourism Management, 38, 122-139.

Sparks, B. A., & Browning, V. (2011). The impact of online reviews on hotel booking intentions and perception of trust. Tourism management, 32(6), 1310-1323.

Sparks, B. A., Perkins, H. E., & Buckley, R. (2013). Online travel reviews as persuasive communication: The effects of content type, source, and certification logos on consumer behavior. Tourism Management, 39, 1-9.

Statistics Canada (1995): Tourism Attitude and Motivation Study, http://www23.statcan.gc.ca/imdb/p2SV.pl?Function=getSurvey&SDDS=3822 [access: 9/17/2019]

Statistics Canada (2018): Travel Survey of Residents of Canada, https://www.statcan.gc.ca/eng/survey/household/3810 [access: 9/17/2019]

Statistics Canada (2019): Results and documentation of surveys and statistical programs, https://www150.statcan.gc.ca/n1/en/type/surveys?subject_levels=24 [access: 9/17/2019]

Stepchenkova, S. (2012). 23 Content analysis. Handbook of research methods in tourism: Quantitative and qualitative approaches, 443.

Stepchenkova, S., Kim, H., & Kirilenko, A. (2015). Cultural differences in pictorial destination images: Russia through the camera lenses of American and Korean tourists. Journal of Travel Research, 54(6), 758-773.

Strapparava, C., & Valitutti, A. (2004). WordNet Affect: an Affective Extension of WordNet. In LREC, 4,1083-1086.

Struijs, P., Braaksma, B., & Daas, P. J. (2014). Official statistics and big data. Big Data & Society, 1(1), 2053951714538417.

Sun, Q., Niu, J., Yao, Z., & Yan, H. (2019). Exploring eWOM in online customer reviews: Sentiment analysis at a fine-grained level. Engineering Applications of Artificial Intelligence, 81, 68-78.

Suzuki, S. & Kurata, Y.(2017). An Analysis of Characteristc of Tourism Destinations using User Profile of Twitter (in Japanese). Society for Tourism Informatics, 13(1), 39-52.

Suzuki, S., & Kurata, Y. (2017). An Analysis of Characteristic of Tourism Destinations using User Profile of Twitter. Society for Tourism Informatics, 13(1), 39-52. (in Japanese)

Taylor, D. C., Barber, N. A., & Deale, C. (2015). To tweet or not to tweet: that is the question for hoteliers: a preliminary study. Information Technology & Tourism, 15(1), 71-99.

The British Tourist Authority (BTA) (2018): Visitor Attraction Trends in England 2018 Full Report, https://www.visitbritain.org/sites/default/files/vbcorporate/Documents-Library/documents/England-documents/annual_attractions_survey_2018_trends_report.pdf [9/17/2019]

The British Tourist Authority (BTA) (2019a): GB Tourism Survey (domestic overnight tourism): Latest results, https://www.visitbritain.org/great-britain-tourism-survey-latest-monthly-overnight-data [9/17/2019]

The British Tourist Authority (BTA) (2019b): GB Day Visits Survey: Latest results, https://www.visitbritain.org/gb-day-visits-survey-latest-results [9/17/2019]

The British Tourist Authority (BTA) (2019c): Accommodation Occupancy: Latest results, https://www.visitbritain.org/accommodation-occupancy-latest-results [9/17/2019]

Tokuhisa, N., Takahara, A., Murata, M. & Murakami, J. (2015). Sentence Extraction for Idea Support on Tourism Development. Society for Tourism Informatics, 11(1), 35-44. (in Japanese)

Tokyo Metropolitan Government (2018). Foreign Travelers Behavioral Trend Survey. Available: http://www.metro.tokyo.jp/tosei/hodohappyo/press/2018/06/06/documents/13_01.pdf [9/17/2019] (in Japanese)

Topp, N. W., & Pawloski, B. (2002). Online data collection. Journal of Science Education and Technology, 11(2), 173-178.

Tourangeau, R., Rips, L. J., & Rasinski, K. (2000). The psychology of survey response. Cambridge University Press.

Tourism Research Australia (TRA) (2019a): National Visitor Survey (NVS) Methodology 2019, https://www.tra.gov.au/Domestic/national-visitor-survey-methodology [9/17/2019]

Tourism Research Australia (TRA) (2019b): International Visitor Survey Methodology, https://www.tra.gov.au/International/ivs-methodology [9/17/2019]

Tseng, C., Wu, B., Morrison, A. M., Zhang, J., & Chen, Y. C. (2015). Travel blogs on China as a destination image formation agent: A qualitative analysis using Leximancer. Tourism Management, 46, 347-358.

U.S. Department of Transportation (2019): U.S. International Air Passenger and Freight Statistics Report, https://www.transportation.gov/policy/aviation-policy/us-international-air-passenger-and-freight-statistics-report [access:9/17/2019]

Ukpabi, D., Olaleye, S., Mogaji, E. and Karjuluoto, H. (2018). Insights into Online Reviews of Hotel Service Attributes: A Cross-National Study of Selected Countries in Africa, Information and Communication Technologies in Tourism 2018, 243-256, Springer

Van der Heijden, H., Verhagen, T., & Creemers, M. (2003). Understanding online purchase intentions: contributions from technology and trust perspectives. European journal of information systems, 12(1), 41-48.

Vermeulen, I. E., & Seegers, D. (2009). Tried and tested: The impact of online hotel reviews on consumer consideration. Tourism management, 30(1), 123-127.

Vu, H. Q., Li, G., Law, R., & Ye, B. H. (2015). Exploring the travel behaviors of inbound tourists to Hong Kong using geotagged photos. Tourism Management, 46, 222-232.

Wan, C. S. (2002). The web sites of international tourist hotels and tour wholesalers in Taiwan. Tourism Management, 23(2), 155-160.

Wang, X., & Li, Y. (2014). Trust, psychological need, and motivation to produce user-generated content: A self-determination perspective. Journal of Electronic Commerce Research, 15(3), 241-253.

Wang, Y., & Morais, D. B. (2014). An examination of tourists' identity in tourist weblogs. Information Technology & Tourism, 14(3), 239-260.

Wang, Y., So, K. K. F., & Sparks, B. A. (2017). Technology readiness and customer satisfaction with travel technologies: A cross-country investigation. Journal of Travel Research, 56(5), 563-577.

Wathen, C. N., & Burkell, J. (2002). Believe it or not: Factors influencing credibility on the Web. Journal of the American society for information science and technology, 53(2), 134-144.

Weeks, P., & Crouch, I. (1999). Sites for sore eyes: an analysis of Australian tourism and hospitality web sites. Information Technology & Tourism, 2(3), 153-172.

Wenger, A. (2008). Analysis of travel bloggers' characteristics and their communication about Austria as a tourism destination. Journal of Vacation Marketing, 14(2), 169-176.

Willemsen, L. M., Neijens, P. C., Bronner, F., & De Ridder, J. A. (2011). “Highly recommended!” The content characteristics and perceived usefulness of online consumer reviews. Journal of Computer-Mediated Communication, 17(1), 19-38.

Wissler, L., Almashraee, M., Díaz, D. M., & Paschke, A. (2014, June). The Gold Standard in Corpus Annotation. In IEEE GSC. DOI: 10.13140/2.1.4316.3523

World Tourism Organization (UNWTO) (2018): UNWTO Tourism Highlights, 2018 Edition, https://www.e-unwto.org/doi/pdf/10.18111/9789284419876

World Tourism Organization (UNWTO) (2019), Compendium of Tourism Statistics, Data 2013 – 2017, 2019 Edition, UNWTO, Madrid, DOI: https://doi.org/10.18111/9789284420421

Xiang, Z., & Gretzel, U. (2010). Role of social media in online travel information search. Tourism management, 31(2), 179-188.

Xiang, Z., Du, Q., Ma, Y., & Fan, W. (2017). A comparative analysis of major online review platforms: Implications for social media analytics in hospitality and tourism. Tourism Management, 58, 51-65.

Xiang, Z., Schwartz, Z., Gerdes Jr, J. H., & Uysal, M. (2015). What can big data and text analytics tell us about hotel guest experience and satisfaction?. International Journal of Hospitality Management, 44, 120-130.

Yagasaki, N.. (2015). Tourism Policy challenges -towards the internationally competitive tourism industry. Society for Serviceology, 1(4), 28-35. (in Japanese).

Yan, Q., Wu, S., Wang, L., Wu, P., Chen, H., & Wei, G. (2016). E-WOM from e-commerce websites and social media: Which will consumers adopt?. Electronic Commerce Research and Applications, 17, 62-73.

Yang, F. X. (2017). Effects of restaurant satisfaction and knowledge sharing motivation on eWOM intentions: the moderating role of technology acceptance factors. Journal of Hospitality & Tourism Research, 41(1), 93-127.

Ye, Q., Law, R., Gu, B., & Chen, W. (2011). The influence of user-generated content on traveler behavior: An empirical investigation on the effects of e-word-of-mouth to hotel online bookings. Computers in Human behavior, 27(2), 634-639.

Yoo, K. H., & Gretzel, U. (2008). What motivates consumers to write online travel reviews?. Information Technology & Tourism, 10(4), 283-295.

Yüksel, A. (2017). A critique of “Response Bias” in the tourism, travel and hospitality research. Tourism Management, 59, 376-384.

Zanasi, A. (2000). Web mining through the Online Analyst. WIT Transactions on Information and Communication Technologies, 25.

Zhang, L., & Li, L. (2018). The Differences of Online Review Textual Content: A Cross-Cultural Empirical Study. WHICEB 2018 Proceedings. 66.

Zhang, L., Pan, B., Smith, W., & Li, X. R. (2009). An exploratory study of travelers' use of online reviews and recommendations. Information Technology & Tourism, 11(2), 157-167.

Zhang, Y., & Cole, S. T. (2016). Dimensions of lodging guest satisfaction among guests with mobility challenges: A mixed-method analysis of web-based texts. Tourism Management, 53, 13-27.

Zhou, L., Ye, S., Pearce, P. L., & Wu, M. Y. (2014). Refreshing hotel satisfaction studies by reconfiguring customer review data. International Journal of Hospitality Management, 38, 1-10.

Appendix A

List of local tourism surveys and touristic projects in Japan (Up to Aug. 2018)

Pref.	Translated Survey Name / Original Survey Name	Survey Period
Hokkaido / 北海道	Survey on accommodation of domestic school trip / 来道修学旅行宿泊実態調査	2004
	Survey Concerning Customer Satisfaction among foreigners / 訪日外国人来道者動態・満足度調査	2005,2006,2007
	Survey on usage of rental bus among foreigners / 訪日外国人貸切バス利用状況調査	2008,2009
	Survey on tourism arrivals / 北海道観光入込客数調査	2010-2018
	Survey Concerning Customer Satisfaction / 観光客動態・満足度調査	2016
	Survey on the tourism economic impact in Hokkaido / 北海道観光産業経済効果調査	1988-2017 (every 5 years)
Aomori / 青森県	Tourism arrivals in Aomori / 青森県観光入込客統計	2006-2016
	Open recruitment of foreigner attitude survey / 外国人意識調査公募	
	Tourism arrivals in Aomori after the opening of the Tohoku Shinkansen Line / 東北新幹線全線開業後における本県観光の動向について	
Iwate / 岩手県	Overview of tourism statistics in Iwate / 岩手県観光統計概要	2001,2007-27,2017
	Inbound tourism arrivals of education trips in Iwate / いわての観光統計教育旅行客外国人観光客の入込動向	2017
	Arrivals of ski tourists in 2017 by season / 平成 29 年シーズンのスキー客入込状況について	2017
Miyagi / 宮城県	Survey on tourism trends / 観光動態調査	2003,2006,2009,2012
	Overview of tourism statistics / 観光統計概要	2003-2016
Akita / 秋田県	Tourism statistics in Akita in 2010 (Tourism arrivals and trends in Akita) /平成 22 年秋田県観光統計(秋田県観光客入込・動態調査)	2010

Yamagata / 山形県	Tourism arrivals in Yamagata / 山形県観光者数調査	2006-2015
Fukushima / 福島県	Tourism arrivals / 観光客入込状況調査	2007-2016
	fact-finding investigation in destinations in Fukushima (2016) / 2016 年福島県観光地実態調査	2016
Ibaraki / 茨城県	Survey on tourism trends / 観光客動態調査	2004-2016
	Tourism arrivals during the golden week / ゴールデンウィークの入込客数について	2015-2018
	Basic Plans of tourism promotion in Ibaraki / 茨城県観光振興基本計画	2016
	Analysis of tourism consumption and its economic impact in Ibaraki (2013 and 2014) / 観光消費が本県にもたらす経済波及効果分析（平成２３年及び平成２４年）	2014
Tochigi / 栃木県	Survey on tourism trends in Tochigi / 栃木県観光動態調査	2015-2017
	Estimation on tourism arrivals and accommodations in Tochigi /栃木県観光客入込数・宿泊数推定調査	2003-2017
Gunma / 群馬県	Report on the usage of hot springs / 温泉利用状況報告	1990~2000
	Estimation of Tourism Arrivals / 観光客入込数推計	1994~2016
	Tourism consumptions / 観光消費額推計	1994~2016
	Tourism Arrivals by inbound/domestic and by one-day/overnight / 県内外別・日帰宿泊別観光客入込数	1994~2016
Saitama / 埼玉県	Survey on tourism arrivals / 観光入込客統計調査	2010-2016
Chiba / 千葉県	Survey on tourism arrivals /観光入込調査	2000-2016
	Tourists needs and arrivals among foreigners using SNS data / SNS を活用した外国人観光客ニーズ・動向調査	2015
	Pamphlet creation for Muslins regarding foreigners' food culture / 「訪日観光客の食文化等に関する調査・推進事業」に係るムスリム観光客向けパンフレットの作成	
	Investigation of innovative local tourism development / 新たな観光地域づくりに係る調査	2018
Tokyo / 東京都	Fact-finding investigation of tourism arrivals and etc. in Tokyo / 東京都観光客数等実態調査	2008-2016

	Foreign Travelers Behavioral Trend Survey / 国別外国人旅行者行動特性調査	2013- 2017
Kanagawa / 神奈川県	Basic survey concerning tourism industry / 観光産業に関する基礎調査	2009
	Survey on tourism consumptions an etc. / 観光客消費動向等調査	2009- 2017
	Tourism Arrivals / 入込観光客調査	2007- 2017
	Fact-finding investigation of foreign tourists / 外国人観光客実態調査	2009- 2017
Niigata / 新潟県	Tourist arrivals during golden week and etc. in Niigata / ゴールデンウィーク等における県内観光動向	2010- 2018
	Tourist arrivals in ski resorts / スキー場利用客入込状況	2007- 2017
	Attitudes towards Niigata among the residents in Kansai / 関西方面を対象とした本県観光に対する意識調査	2012
	Tourist Arrivals of sea bathing facilities / 海水浴客入込状況	2010- 2017
	Attitudes towards Niigata among the residents in Tokyo and Kansai / 首都圏・関西圏を対象とした本県観光に対する意識調査	2017
	Foreigner accommodations / 外国人宿泊客数調査	2010- 2016
	Investigation of economic impact in destinations in Niigata / 県内観光地の経済波及効果調査	2004
	Tourism arrivals in Niigata / 県内観光動向	2014,2015
	Tourism trends in Niigata / 新潟県観光動態	2004,2008,2009
	Tourism arrivals in Niigata / 新潟県観光入込客統計調査	2010- 2016
	Tourism arrivals in Sado / 佐渡観光客入込状況	2007- 2011
Toyama / 富山県	Estimation of tourism arrivals / 観光客入込数推計	2007- 2016
Ishikawa / 石川県	Tourism statistics in Ishikawa (tourism statistics and survey on tourism trends) / 統計からみる石川県の観光（観光統計＋観光動態調査）	2010- 2016
Fukui / 福井県	Tourism arrivals in Fukui (Estimation) / 福井県観光客入込数（推計）	2004- 2016
	Local brands for inbound promotion / 外国人誘客に向けたブランド	

	Multi-language guideline sheet / 観光地等の多言語表記ガイドライン、飲食店等での外国人接客用指さし会話シート	
Yamanashi / 山梨県	Survey on tourism arrivals in Yamanashi / 山梨県観光入込客統計調査	2010-2017
	Tourism trends in Yamanashi / 山梨県観光客動態調査	2003-2009
	Survey on accommodations / 宿泊旅行統計調査	2014-2018
	Tourism arrivals during the golden week / ゴールデンウィークの観光客	2005-2018
	Investigation of the influence of the Sasago tunnel incident / 笹子トンネル天井板落下事故による観光への影響調査	2012,2013
Nagano / 長野県	Tourism arrivals at ski/skate borating facilities / スキー・スケート場利用者統計調査	2007-2018 (every 2 years)
	Current status of ski and etc. facilities / スキー場等現況調査	2009-2015
	Survey on tourism arrivals / 観光地利用者統計調査	2003-2016
	Survey on inbound accommodations / 外国人延宿泊者数調査	2005-2016
	Fact-finding investigation about educational trips / 学習旅行実態調査	2009-2016
Gifu / 岐阜県	Survey on tourism trends / 観光動態調査	2016
	Survey on tourism arrivals / 観光入込客統計調査	2000-2016
Shizuoka / 静岡県	[cannot access to the website]	
Aichi / 愛知県	Survey on tourism arrivals based on the Common Standards /「観光入込客統計に関する共通基準」に基づく観光入込客統計	2010-2016
	Fact-finding investigation concerning MICE(meeting, incentive travel, convention, event/exhibition) / MICE 実態調査	2012
	Investigation of travel routes in Aichi using big data analysis /ビッグデータを活用した愛知県の観光拠点及び回遊に関する調査	2016
	Survey on inbound tourism arrivals in Aichi / 愛知県訪日外客動向調査	2015-2017
	Usage of tourism recreation facilities / 観光レクリエーション利用者統計	2006-2017

Mie / 三重県	Tourism arrivals (actual number) / 入込客数(実数)	2005-2017
Shiga / 滋賀県	Survey on Tourism arrivals / 観光入込客統計調査	2002-2015
Kyoto / 京都府	Tourism arrivals and tourism consumptions /観光入込客数及び観光消費額について	2008-2017
Oosaka / 大阪府	Survey on tourism statistics in Oosaka / 大阪府観光統計調査	2004-2012
	Report on fact-finding investigations regarding the preparation of receiving tourism accommodation /大阪府観光客受入環境整備の推進に関する宿泊実態調査報告書	2015
Hyogo / 兵庫県	[cannot access to the website]	
Nara / 奈良県	Report on tourism trends / 観光客動態調査報告書	2006-2015
	Report on accommodations in Nara / 奈良県宿泊統計調査報告書	2009-2016
Wakayama / 和歌山県	Tourism trends / 観光客動態	2009-2017
	Tourism arrivals during the new year season / 年末年始の観光客入込状況	2011-2013,2015-2017
	Tourism arrivals during summer (Jul. 1st to Aug. 31st) / 夏季(7 月 1 日から 8 月 31 日)の観光客入込状況	2009-2017
	Tourism arrivals during the golden week / ゴールデンウィークの観光客	2009-2018
	Survey on tourism statistics / 観光統計調査	2008,2014,2017
Tottori / 鳥取県	Tourism satisfaction survey / 観光客満足度調査	2016~2017
	Inbound tourism arrivals / 外国人観光入込客数	2014-2018
	Tourism arrivals / 観光入込動態調査	2000-2016
	Tourism arrivals in major tourism facilities in Tottori / 鳥取県内の主要観光施設における外国人観光入込客数について	2018
Shimane / 島根県	Survey on tourism awareness towards Shimane / しまねの観光認知度調査	2013-2016
	Reports on tourism trends in Shimane / 島根県観光動態調査結果	2002-2017
	Attitude towards matchmaking among female tourists / 女性観光客動向調査・「縁結び」に関する女性観光客意識調査	2014,2015

	Monthly tourism arrivals in 2017 / 平成２９年島根県月別主要観光動向	2017
	Tourism arrivals and travel routes in San-in region (Tottori and Shimane) using GPS and GAP data analysis / 山陰（鳥取・島根）観光動態調査（GPS・GAP 調査）の調査結果について	2013
	About inbound tourism / 外国人観光客	2013-2017
Okayama / 岡山県	Tourism satisfaction survey / 観光客満足度調査結果	2007-2010
	Inbound tourism accommodation / 外国人旅行者宿泊者数調査結果	2015-2017
	Tourism trends in Okayama / 岡山県観光客動態調査結果	2003-2017
Hiroshima / 広島県	Tourist arrivals in Hiroshima / 広島県観光客数の動向	2004-2017
	On-spot parameter survey / 観光地点パラメータ調査	2016,2017
Yamaguchi / 山口県	Accommodation and tourism arrivals in Yamaguchi / 山口県の宿泊者及び観光客の動向	2009-2016
Tokushima / 徳島県	Tourism arrivals during the golden week / ゴールデンウィーク期間中の入込客状況	2017,2018
Kagawa / 香川県	[official website not found]	
Ehime / 愛媛県	Basic plan for tourism promotion in Ehime / 愛媛県観光振興基本計画	2012-2016
	Tourism arrivals in major destinations during the golden week /ゴールデンウィーク期間中の主要観光地観光客数	2016-2018
	Tourism arrivals and consumptions / 観光客数とその消費額	2013-2015
Kochi / 高知県	Report on tourism arrivals and trends / 県外観光客入込・動態調査報告書	2008-2015
Fukuoka / 福岡県	Estimation of tourism arrivals in Fukuoka / 福岡県観光入込客推計調査	2011-2016
Saga / 佐賀県	Survey on Tourism trends / 観光客動態調査	2014-2016
Nagasaki / 長崎県	Tourism arrivals in Nagasaki / 長崎県観光動向調査	2012-2018
	Tourism statistics in Nagasaki	2012-

	/ 長崎県観光統計	2017
	Usages in major tourism facilities and etc. / 主要観光施設等の利用者数	2009- 2018
Kumamoto / 熊本県	Tourism statistics in Kumamoto / 熊本県観光統計	2007- 2016
Oita / 大分県	Survey on tourism statistics in Oita / 大分県観光統計調査	2008- 2017
	Tourism arrivals during the golden week / ゴールデンウィーク観光動向調査	2012- 2018
	Fact-finding investigation / 観光実態調査報告書	2005,2010-2017
	Report of the questionnaire regarding accommodations / 宿泊客アンケート調査報告書	2006- 2009
Miyazaki / 宮崎県	Tourism arrivals in Mayazaki / 宮崎県観光入込客統計調査	2012- 2016
Kagoshima / 鹿児島県	Tourism trends in Kagoshima – tourism statistics- / 鹿児島県の観光の動向～鹿児島県観光統計～	2005- 2016
	Survey on tourism arrivals in Kagoshima / 鹿児島県観光動向調査	2007- 2018
	Tourism accommodations / 宿泊旅行統計	No data
Okinawa / 沖縄県	Report on tourism satisfaction survey in Okinawa 2003 / 平成 15 年度沖縄観光客満足度調査報告書	2003
	Report on current status of tourism industry in Okinawa / 沖縄県観光産業実態調査報告書	2014,2015
	Reginal tourism arrivals / 入域観光客数	2010- 2017
	Statistics concerning accommodation facilities / 宿泊施設に関する統計データ	2004,2011-2016

Appendix B

Numbers of positive (P), neutral (E) and negative (N) reviews

#	Items	Taiwan (n=109)				Mainland China (n=75)			
		P	E	N	Total	P	E	N	Total
1	for the entire trip and sightseeing	0	0	0	0	0	1	0	1
2	meals at each tourist destination	39	18	2	59	18	12	6	36
3	souvenirs	4	8	2	14	2	5	0	7
4	accommodations	28	9	1	38	10	5	0	15
5	tourist attractions	9	19	3	31	10	11	2	23
6	Wi-Fi accessibility	0	2	2	4	0	0	0	0
7	multilingual informational signs	0	0	1	1	0	0	0	0
8	local staff's linguistic abilities	1	1	1	3	0	1	2	3
9	transportation system	15	36	1	52	10	17	2	29
10	customer service	17	1	1	19	12	9	3	24
11	scenery	33	10	0	43	17	10	0	27

#	Items	Hong Kong (Chinese) (n=127)				Hong Kong (English) (n=122)			
		P	E	N	Total	P	E	N	Total
1	for the entire trip and sightseeing	0	0	0	0	0	0	0	0
2	meals at each tourist destination	42	29	3	74	47	20	3	70
3	souvenirs	6	10	0	16	3	5	0	8
4	accommodations	34	9	1	44	28	14	2	44
5	tourist attractions	8	15	0	23	6	16	2	24
6	Wi-Fi accessibility	0	0	1	1	2	0	0	2
7	multilingual informational signs	2	0	0	2	1	0	2	3
8	local staff's linguistic abilities	2	3	3	8	4	3	2	9
9	transportation system	28	16	2	46	22	31	1	54
10	customer service	22	3	3	28	37	5	8	50
11	scenery	16	8	0	24	26	11	1	38

Numbers of positive (P), neutral (E) and negative (N) reviews (Continued)

#	Items	Singapore (n=332)				Australia (n=190)			
		P	E	N	Total	P	E	N	Total
1	for the entire trip and sightseeing	0	0	0	0	0	0	0	0
2	meals at each tourist destination	139	63	9	211	94	22	5	121
3	souvenirs	5	29	0	34	5	9	0	14
4	accommodations	75	24	7	106	32	9	3	44
5	tourist attractions	36	38	3	77	30	25	5	60
6	Wi-Fi accessibility	2	5	2	9	3	1	0	4
7	multilingual informational signs	4	6	2	12	1	6	1	8
8	local staff's linguistic abilities	8	19	9	36	3	5	3	11
9	transportation system	57	67	9	133	26	26	4	56
10	customer service	77	31	7	115	44	14	5	63
11	scenery	72	16	2	90	43	5	2	50

#	Items	America (n=168)				Britain (n=35)			
		P	E	N	Total	P	E	N	Total
1	for the entire trip and sightseeing	2	1	0	3	0	0	0	0
2	meals at each tourist destination	74	24	7	105	19	3	1	23
3	souvenirs	2	12	1	15	0	0	0	0
4	accommodations	29	9	8	46	5	2	2	9
5	tourist attractions	27	22	3	52	5	3	1	9
6	Wi-Fi accessibility	1	0	1	2	0	0	1	1
7	multilingual informational signs	2	9	0	11	1	0	1	2
8	local staff's linguistic abilities	4	8	2	14	0	1	1	2
9	transportation system	20	42	2	64	4	4	1	9
10	customer service	42	17	4	63	13	3	2	18
11	scenery	40	11	0	51	7	0	0	7

Appendix C

Correlations by city in 2016 and 2017 (*n* = 12)

(*r* is Pearson's *r*, *a* is slope, *b* is intercept, n_1 is the number of sightseeing visitors (unit: thousand), n_2 is the number of reviewers)

179 cities in Hokkaido	2016					2017				
	Correlations			Numbers		Correlations			Numbers	
	r	*a*	*b*	n_1	n_2	*r*	*a*	*b*	n_1	n_2
Sapporo	0.204	0.031	397.4	13879.5	5200	0.080	0.007	410.4	15270.9	5032
Otaru	0.708	0.142	44.0	7907.7	1653	0.170	0.040	107.8	8061.6	1615
Asahikawa	0.937	0.138	61.3	5310.0	1469	0.959	0.152	55.8	5357.0	1484
Hakodate	0.777	0.213	130.7	5606.9	2765	0.758	0.244	94.6	5246.8	2416
Chitose	0.690	0.108	28.2	5187.4	901	0.787	0.099	26.6	5240.5	838
Kushiro	0.673	0.127	21.9	4599.5	847	0.714	0.128	13.3	5239.4	832
Noboribetsu	0.421	0.088	60.6	3851.9	1067	0.410	0.094	42.7	4048.9	892
Toyako Town	0.947	0.184	6.6	3067.6	643	0.840	0.103	23.4	2931.7	583
Obihiro	0.736	0.066	32.0	2481.9	547	0.920	0.096	26.9	2704.2	582
Kimobetsu Town	0.858	0.009	0.4	2552.9	28	0.829	0.007	1.1	2583.6	31
Sobetsu	0.461	0.078	1.6	2332.0	202	0.899	0.152	-13.3	2187.3	173
Ishikari	0.480	0.009	2.3	2106.9	46	0.861	0.017	-0.1	2048.5	33
Tomakomai	0.555	0.053	12.4	1932.9	251	0.920	0.112	-2.9	1994.8	188
Furano	0.892	0.271	18.5	1859.8	726	0.916	0.222	15.6	1894.3	608

Kamikawa Town	0.785	0.106	8.2	1874.3	297	0.784	0.098	12.2	1853.9	328
Nanae Town	0.572	0.069	2.2	1998.1	164	0.648	0.068	4.4	1838.3	178
Date	0.288	0.024	5.0	1811.3	103	0.460	0.055	-1.3	1810.7	84
Shimukappu Village	0.823	0.159	7.4	1487.3	325	0.436	0.186	18.5	1756.7	549
Shiraoi	0.573	0.058	0.2	1766.7	104	0.695	0.072	-3.5	1735.5	83
Biei Town	0.964	0.419	-4.5	1659.5	641	0.967	0.407	-3.5	1679.5	641
Niseko town	0.728	0.765	-10.7	1671.3	1150	0.699	0.525	4.0	1672.4	925
Abashiri	0.788	0.167	16.5	1530.2	454	0.861	0.159	18.8	1624.1	483
Kutchan	0.783	2.230	-115.4	1566.7	2109	0.781	1.698	-77.4	1614.1	1813
Otofuke	0.267	0.029	9.3	1370.6	152	0.439	0.063	5.5	1543.4	163
Rusutsu Village	0.922	0.169	1.5	1507.1	273	0.964	0.137	2.6	1510.1	239
Kitami	0.763	0.131	1.5	1461.6	209	0.912	0.160	-0.2	1493.3	236
Higashikawa Town	0.197	0.029	13.5	1450.9	205	0.418	0.050	7.6	1488.8	166
Sunagawa	0.533	0.017	-0.2	1209.3	19	0.672	0.026	-0.6	1412.7	29
Akaigawa Village	-0.036	-0.018	20.0	1140.1	219	0.355	0.093	5.1	1351.7	187
Eniwa	0.585	0.024	2.3	1267.8	58	0.600	0.031	1.2	1351.1	57
Mikasa	0.758	0.041	0.4	997.5	46	0.383	0.014	1.6	1283.4	36
Shari Town	0.951	0.278	16.8	1188.3	532	0.890	0.227	19.4	1217.7	510
Muroran	0.519	0.034	9.4	1286.8	156	0.888	0.076	4.4	1201.8	144
Ozora Town	0.607	0.044	1.9	935.2	64	0.628	0.041	1.7	1171.4	69
Yoichi Town	0.862	0.149	7.7	1282.7	283	0.864	0.145	4.8	1163.6	226
Iwamizawa	-0.144	-0.005	4.4	1194.9	47	0.288	0.009	2.3	1145.4	37

Nakafurano Town	0.947	0.182	11.3	1068.4	330	0.940	0.165	7.4	1110.9	272
Shakotan	0.827	0.094	6.5	1220.2	193	0.872	0.099	4.0	1058.6	153
Kitahiroshima City	0.587	0.045	6.0	1056.6	119	0.531	0.052	5.6	1050.6	121
Hokuto	0.719	0.021	1.2	1237.7	40	-0.142	-0.008	5.8	997.1	61
Fukagawa	-0.026	-0.001	2.0	912.4	23	0.199	0.007	0.6	938.5	13
Teshikaga	0.957	0.407	-0.6	914.5	365	0.960	0.252	5.2	936.7	299
Ebetsu	0.061	0.002	3.3	1046.1	41	-0.368	-0.013	3.9	915.3	35
Ashibetsu	0.249	0.015	1.3	909.6	29	0.764	0.048	-1.1	904.3	30
Nakasatsunai Village	0.769	0.059	1.0	757.1	56	0.832	0.049	1.0	887.0	55
Mori Town	0.128	0.005	2.3	907.1	32	0.560	0.035	0.1	882.1	32
Shintoku Town	0.279	0.019	6.7	933.4	98	-0.220	-0.014	7.4	859.2	77
Tobetsu Town	-0.251	-0.025	2.4	419.8	19	0.351	0.013	0.8	834.2	21
Rankoshi Town	0.640	0.053	1.1	829.0	57	0.445	0.049	1.6	814.2	59
Kyogoku Town	0.609	0.046	-0.4	846.9	35	0.825	0.079	-1.9	812.5	41
Naganuma Town	0.530	0.035	-0.7	690.2	16	0.720	0.029	-0.5	737.7	16
Bihoro Town	0.627	0.036	2.1	713.0	51	0.292	0.011	4.3	726.6	59
Shikaoi Town	0.415	0.026	2.6	744.9	51	0.077	0.005	3.4	716.2	44
Makubetsu Town	0.552	0.063	1.0	628.4	52	0.685	0.067	-0.3	668.9	41
Takigawa	0.238	0.005	1.3	736.4	19	0.668	0.027	0.3	661.2	22
Kamifurano Town	0.868	0.171	1.9	610.1	127	0.917	0.143	2.6	631.4	121
Shiranuka Town	0.410	0.009	0.2	600.8	8	0.610	0.026	-0.2	621.2	14
Kenbuchi Town	0.273	0.013	0.3	628.6	12	0.259	0.009	0.3	601.0	9

Kikonai Town	0.831	0.048	-0.6	626.1	23	0.126	0.003	0.8	570.7	12
Yakumo Town	0.340	0.027	0.9	628.0	28	0.800	0.100	-1.1	560.7	43
Rausu	0.829	0.076	4.4	539.0	94	0.932	0.096	2.8	546.2	86
Hokuryu Town	0.882	0.060	-0.4	503.8	26	0.969	0.041	0.1	539.2	23
Chichibu Betsucho	-0.254	-0.014	1.1	437.2	7	0.455	0.012	0.0	538.8	7
Yubari	0.675	0.156	0.1	489.1	77	0.617	0.094	0.8	534.1	60
Koshimizu Town	0.829	0.086	0.9	410.2	46	0.777	0.043	0.0	531.2	23
Yubetsu-cho	0.575	0.041	0.6	519.1	29	0.600	0.037	0.4	521.7	24
Wakkanai	0.838	0.279	14.8	507.6	319	0.865	0.430	4.8	520.8	282
Ashoro Town	0.817	0.073	3.0	462.2	70	0.802	0.079	1.4	516.2	57
Oshamanbe Town	0.673	0.036	-0.4	507.0	13	0.415	0.030	-0.3	499.6	11
Monbetsu	0.338	0.037	6.6	470.2	97	0.544	0.055	5.8	491.4	97
Honbetsu Town	-0.242	-0.009	1.0	505.2	7	0.106	0.007	0.8	478.5	13
Shikabe Town	0.510	0.045	1.1	479.5	34	0.106	0.007	1.7	459.7	24
Kuriyama Town	0.769	0.019	0.4	448.5	13	0.518	0.015	1.4	456.1	24
Akkeshi	0.684	0.061	1.6	414.6	45	0.891	0.093	0.1	449.1	43
Matsumae Town	0.893	0.059	1.0	434.9	38	0.579	0.026	1.8	448.4	33
Nayoro	0.702	0.062	-0.5	485.0	24	0.179	0.015	1.4	446.2	23
Toma town	0.535	0.020	-0.2	431.6	7	0.424	0.028	-0.2	444.3	10
Makkari Village	0.416	0.117	1.7	409.9	69	0.718	0.293	-4.9	441.2	70
Kamishihoro Town	0.265	0.032	4.5	359.1	66	0.806	0.142	2.8	439.0	96
Hamanaka	0.755	0.129	-0.4	380.8	44	0.482	0.053	1.0	431.4	35

Minamifurano Town	0.749	0.087	1.2	379.0	47	0.455	0.038	2.9	429.4	51
Bifuka	0.671	0.054	-0.6	400.8	15	0.670	0.031	-0.5	414.3	7
Iwanai Town	0.075	0.003	2.5	431.9	31	0.609	0.033	2.3	410.8	41
Toyoura	0.297	0.022	0.7	420.3	18	0.823	0.046	-0.6	403.4	11
Shihoro Town	0.353	0.074	0.5	99.5	13	0.745	0.081	-0.3	402.0	29
Nemuro	0.636	0.170	5.3	377.0	128	0.583	0.083	5.9	397.1	104
Hidaka	0.108	0.020	2.0	364.2	31	0.315	0.065	-0.8	367.6	14
Niikappu-cho	0.925	0.081	-0.5	351.8	22	0.709	0.082	0.8	366.0	40
Shibetsu Town	0.773	0.060	0.4	360.0	26	0.719	0.038	0.9	361.4	25
Shibetsu	0.091	0.006	0.4	323.9	7	0.697	0.119	-2.6	360.8	12
Anping Town	0.332	0.026	0.7	331.4	17	0.506	0.032	0.3	358.6	15
Utashinai	0.107	0.013	0.1	358.0	6	0.561	0.090	-2.0	349.3	7
Kuromatsunai Town	0.596	0.078	-0.1	152.5	11	0.248	0.020	0.1	346.0	8
Esashi	0.504	0.021	2.6	345.8	39	0.481	0.016	1.6	345.4	25
Yuni Town	0.715	0.040	0.2	350.3	16	0.552	0.082	-0.2	337.8	25
Shinhidaka Town	0.892	0.096	0.0	313.4	30	0.929	0.089	0.5	319.6	34
Betsukai	0.768	0.162	2.1	282.3	71	0.508	0.069	2.8	309.8	55
Mashike Town	0.280	0.034	3.2	328.0	49	0.927	0.092	0.1	306.3	29
Bibai	0.400	0.056	0.7	300.8	25	0.797	0.185	-2.0	302.3	32
Engaru Town	0.417	0.055	2.3	264.4	42	0.539	0.054	1.3	289.6	31
Nakashibetsu Town	0.704	0.219	2.9	289.2	98	0.658	0.105	4.4	286.2	83
Higashi Kagura Town	0.193	0.049	1.9	268.3	36	0.026	0.005	1.9	283.7	24

Rich town	0.161	0.029	2.2	275.1	34	0.418	0.044	1.6	280.0	32
Rainy town	-0.321	-0.016	0.9	253.5	7	0.131	0.006	0.3	262.0	5
Tsubetsu Town	0.718	0.118	0.1	244.6	30	0.474	0.084	-0.1	257.6	20
Ikeda Town	0.208	0.040	2.0	256.6	34	0.501	0.100	-0.1	257.4	25
Uraporo Town	0.192	0.009	0.4	263.2	7	0.373	0.032	0.3	252.0	12
Rumoi	0.531	0.046	1.8	258.2	34	0.726	0.038	1.5	245.3	27
Nanporo Town	0.019	0.002	0.3	270.7	4	-0.277	-0.023	0.8	241.6	4
Niki Town	-0.103	-0.003	0.3	220.4	3	0.841	0.039	-0.2	240.7	7
Akabira	-0.271	-0.008	0.3	229.4	2	0.698	0.032	-0.3	240.2	4
Esashi Town	0.035	0.002	0.7	227.1	9	0.451	0.034	0.0	239.2	8
Suttsu-cho	0.460	0.017	0.2	203.8	6	0.724	0.041	0.0	238.5	10
Teshio	-0.186	-0.020	1.2	226.2	10	0.575	0.025	-0.2	231.8	4
Setana town	0.361	0.050	0.8	229.4	21	0.186	0.013	0.5	231.3	9
Urausu	0.550	0.043	0.5	222.2	15	-0.032	-0.001	0.4	226.4	5
Sister back cow town	0.543	0.060	-0.9	225.5	3	-0.225	-0.029	0.8	221.8	3
Biratori town	0.413	0.058	0.3	206.8	16	0.633	0.161	-0.8	220.7	26
Kyowa Town	0.874	0.081	0.9	191.2	26	0.621	0.042	1.3	213.0	24
Kamieuchi Village	0.029	0.002	1.1	195.9	13	0.445	0.025	0.2	208.4	7
Horokanai Town	0.052	0.003	1.0	183.7	13	0.086	0.004	0.5	205.1	7
Tsurui Village	0.067	0.018	2.5	155.4	33	-0.245	-0.190	6.3	204.8	37
Memuro-cho	0.521	0.058	-0.3	200.7	8	-0.365	-0.032	0.9	198.9	5
Hifu Town	0.265	0.033	0.2	205.5	9	0.074	0.004	0.4	190.5	5

Kodaira Town	0.390	0.025	0.7	187.7	13	0.875	0.065	-0.3	189.3	9
Saroma Town	0.657	0.157	0.8	166.0	35	0.333	0.048	0.8	188.3	19
Naie Town	0.069	0.007	0.3	117.6	4	0.338	0.067	-0.4	184.9	8
Asasawabe Town	-0.062	-0.008	0.5	165.8	5	0.680	0.102	-0.9	179.1	8
Rikubetsu Town	-0.183	-0.016	0.7	169.4	6	0.508	0.040	-0.1	178.6	6
Shinshinotsu Village	-0.164	-0.046	2.0	160.1	17	0.545	0.117	-0.6	178.5	14
Erimo Town	0.858	0.195	0.2	169.8	36	0.623	0.130	1.5	172.4	40
Sarufutsu Village	-0.024	-0.003	1.2	181.7	14	0.760	0.103	-0.6	172.3	11
Kiyosato	0.763	0.358	1.1	152.3	68	0.740	0.541	0.3	168.2	94
Atsuma	-0.226	-0.006	0.2	156.6	1	-0.255	-0.007	0.2	159.3	1
Numata Town	0.645	0.038	0.3	163.6	10	-0.109	-0.010	1.5	154.9	16
Mukawa town	0.173	0.024	0.5	138.9	9	0.640	0.104	-0.1	152.9	15
Rishiri Town	0.769	0.292	1.6	139.9	60	0.811	0.243	0.9	146.3	46
Rishiri Fuji Town	0.867	0.388	0.8	139.9	64	0.818	0.401	0.4	146.3	63
Shintotsugawa Town	-0.039	-0.006	0.6	152.2	6	0.255	0.069	-0.4	145.2	5
Shiriuchi town	0.482	0.059	-0.2	169.8	7	0.254	0.058	0.0	145.1	9
Hamatonbetsu Town	0.265	0.039	0.8	141.8	15	0.463	0.055	0.3	140.8	11
Urakawa	0.150	0.024	1.4	134.6	20	0.808	0.137	0.3	136.8	22
Tomamae Town	0.367	0.047	0.4	137.0	11	0.621	0.117	-0.4	133.1	11
Rebun Town	0.887	0.475	1.4	117.2	73	0.884	0.552	0.3	126.0	73
Shimokawa Town	0.593	0.081	-0.4	118.3	5	-0.193	-0.047	1.2	123.3	9
Tsukigata Town	0.291	0.037	0.5	109.8	10	0.073	0.007	0.2	122.0	3

Shibecha Town	0.819	0.387	1.8	102.7	61	0.926	0.263	0.7	121.9	41
Hiroo Town	-0.031	-0.003	1.0	123.0	11	0.630	0.102	-0.1	115.3	10
Takasu Town	-0.510	-0.038	0.8	106.1	6	-0.397	-0.030	0.7	111.4	5
Enbetsu Town	-0.248	-0.019	0.3	109.5	2	0.649	0.042	0.0	109.7	4
Horonobe	0.582	0.068	0.5	95.6	12	0.743	0.075	0.0	109.6	8
Tomari Village	0.128	0.014	0.5	110.4	7	0.127	0.012	0.3	105.1	5
Samani Town	0.905	0.109	0.0	111.7	12	0.236	0.026	0.4	104.6	8
Okido Town	0.132	0.020	0.1	29.0	2	-0.188	-0.024	0.4	101.0	2
Otake Town	0.020	0.013	1.6	82.1	20	-0.417	-0.325	3.6	99.7	11
Otobe Town	-0.151	-0.009	0.3	105.1	3	0.062	0.008	0.5	98.4	7
Kamisunagawa Town	-0.101	-0.035	0.5	101.3	3	-0.003	0.000	0.1	97.4	1
Kushiro Town	0.918	0.199	0.3	109.8	26	0.915	0.288	0.4	94.3	32
Kaminokuni Town	-0.030	-0.004	0.6	115.9	7	0.372	0.046	0.2	89.9	7
Nakagawa Town	-0.224	-0.028	0.4	84.0	3	0.575	0.058	-0.3	89.6	2
Kodaira	0.316	0.054	0.4	86.5	9	0.052	0.003	0.1	86.4	1
Haboro Town	0.767	0.196	1.0	85.5	29	0.776	0.138	2.4	85.6	40
Shimamaki Village	-0.127	-0.019	0.9	79.0	9	0.444	0.105	0.0	68.4	7
Hatsuyama Betsumura	0.033	0.008	0.7	64.8	9	0.411	0.068	0.2	67.1	7
Takigami Town	0.556	0.209	1.2	64.8	28	0.960	0.215	0.1	66.5	15
Fukushima Town	0.240	0.038	0.5	75.0	9	0.728	0.142	-0.1	65.3	8
Toyokoro	-0.379	-0.116	1.3	53.5	9	0.045	0.011	0.7	61.0	9
Aibetsu	0.170	0.020	0.1	59.6	2	#DIV/0!	0.000	0.0	58.8	0

Nakatonbetsu Town	-0.476	-0.104	0.7	47.8	4	0.309	0.039	0.1	56.3	3
Taiki-cho	-0.317	-0.152	1.9	54.5	14	0.700	0.361	-0.4	54.3	15
Otoiko Fumura	-0.274	-0.283	2.7	54.5	17	0.450	0.240	-0.4	49.9	7
Imanamachi	-0.010	-0.002	0.4	66.9	5	-0.074	-0.027	0.6	49.6	6
Kunikofu Town	0.136	0.008	0.0	53.0	1	0.007	0.000	0.1	49.2	1
Sarabetsu Village	-0.289	-0.049	0.8	43.2	7	0.319	0.064	0.2	47.4	6
Wassamu Town	-0.112	-0.019	0.7	46.6	8	-0.055	-0.006	0.2	47.1	2
Shimizu Town	0.908	0.265	0.5	50.9	20	0.805	0.392	0.3	40.6	19
Okushiri town	0.524	0.236	0.4	27.1	11	0.749	0.582	-0.2	35.0	18
Okbe-cho	-0.201	-0.038	0.7	38.0	7	0.128	0.028	0.3	34.0	4
Saikobe Village	0.128	0.083	0.7	34.3	11	0.241	0.195	0.1	31.3	7