DOCTORAL DISSERTATION

Analysis on the Changes in Preferences of Chinese Tourists:

For the Recovery of Inbound Tourism in Japanese Local

Regions after COVID-19 Pandemic

(中国人観光客の選好性の変化に関する分析: コロナ禍以降の日本の

地方部におけるインバウンド観光回復を目的として)

HUO MINGMING

Graduate School of Sciences and Technology for Innovation,

Yamaguchi University

Sep 2024

Table of contents

Acknowledgments:

CHAPTER I
PRODROME

1. 1 Introduction

Due to declining birth rates and an aging population, Japan faces serious issues of population decrease and economic stagnation, especially in local regions where the problems of population decline and aging are particularly pronounced. Furthermore, the industrial structure lacking diversity in some local regions has led to regional depopulation. Local governments and enterprises urgently need to find new paths for development. In the early 21st century, the global tourism industry was rapidly developed and became a key driver of economic growth for many countries. Japan also recognized this trend and, in 2003, then-Prime Minister Junichiro Koizumi proposed the slogan “Tourism Nation,” aiming to vigorously develop tourism to address the issues of population decline and economic stagnation, especially to revitalize the local economy. Inbound tourism not only brings direct revenue to Japan but also stimulates related industries such as transportation, accommodation, and catering, creating more job opportunities and increasing the income of employees, thereby expanding the consumer market and revitalizing Japan's economy. In recent years, with the continuous increase in the number of foreign visitors to Japan, the number of tourists traveling to local regions has also been rising, effectively boosting the local economy and positively impacting the mitigation of labor shortages and economic decline in these areas. Therefore, the Japanese government considers the tourism industry as a crucial sector for enhancing the national economy and revitalizing local economies.

To attract more foreign tourists visiting Japan, the Japanese government launched the “Tourism Nation Promotion Basic Plan” in June 2013, including strategically relaxing visa requirements for foreign visitors to Japan, expanding the tax-free system, strengthening the entry and exit management system, and expanding the airline network [1]. Based on the effects of the inbound tourism policies mentioned above, in 2019, the number of foreign visitors who visited Japan was 31.88 million, and the travel consumption was 4,813.5 billion yen, both of which were record highs [2]. Among them, the number of Chinese tourists who visited Japan was 8.57 million and their travel consumption was 1,770.4 billion yen, which were also record highs. Moreover, China has been Japan's largest inbound tourism market since 2015. According to the statistical data of Japan Tourism Agency “International Visitor Survey”, the per capita travel consumption of Chinese tourists who visited Japan increased sharply from 2014 to 2015, and a phenomenon called shopping spree emerged. Although the phenomenon has gradually flattened since 2016, Japan's tourism market remains highly attractive to Chinese visitors, with its food, culture, scenery, and other attractions drawing more and more Chinese tourists to visit, besides shopping. Additionally, with the increase of Chinese tourists who visited Japan for personal tourism and repeated visits to Japan, the number of Chinese visitors visiting local regions, in addition to the capital region (Tokyo, Kanagawa, Saitama, Chiba), Keihanshin region (Osaka, Kyoto

and Hyogo) and Aichi prefecture, has also been on an upward trend [3]. It can be speculated that the needs and tourism behaviors of Chinese tourists traveling to Japan are becoming more and more diverse. In addition, Yagasaki et al.'s research also indicates that with the increase of repeated visitors, the number of visitors to regions other than the “Golden Route” including Tokyo, Kyoto, and Osaka is increasing. Looking at the changes in the total number of foreign visitors staying in the three major metropolitan regions and other local regions, the annual growth rate and the overall share of local regions has exceeded that of the three major metropolitan regions since 2015 [4]. Therefore, From the perspective of revitalizing the local economy, inbound tourism is expected to develop. However, local regions are considered to be at a relative disadvantage in terms of tourists’ recognition and mobility convenience. Consequently, strategic tourism promotion is necessary particularly for local regions.

On the other hand, since the outbreak of COVID-19 in early 2020, the global tourism industry has suffered a huge impact [5]. As the pandemic has spread globally, almost all countries and regions have implemented border closures and restrictive measures [6, 7], and industries such as aviation, hotels, restaurants, and tourism operations have been devastated [8, 9]. According to the data from the UNWTO, the number of international tourists decreased by 72%, and international tourism export earnings declined by 63% in 2020 compared to figures from the previous year, equivalent to a loss of $1.1 trillion in annual revenue [10]. As one of the world's important economic pillar industries, the tourism industry's impact is not only short-term direct economic losses but may also have long-term effects on the psychology, demand and tourism preferences of tourism consumers. This may subsequently affect their behavior and decision-making when it comes to participating in outdoor activities. For example, some research findings indicate that influenced by the COVID-19 pandemic, tourists may pay more attention to the hygiene and safety conditions of travel destinations and tourism service providers [11 - 15]. This has led some tourism consumers to be more willing to choose destinations with smaller COVID-19 impact or that can provide space from others, such as natural scenic areas, outdoor activities, and other low-contact tourism destinations, instead of densely populated urban areas with high human traffic [16 - 26]. Due to the advantages of local tourist spots, such as relatively sparse crowds, good air quality, and easier maintenance of social distancing, the local tourism market is expected to experience a period of growth. This is especially true for areas that can offer unique natural landscapes and a wealth of outdoor activities. Furthermore, considering the differences in cultural backgrounds, economic levels, tourism development stages, as well as variations in the infection situation and prevention and control measures of the COVID-19 pandemic in different countries and regions, the demands, preferences, and consumption habits of the tourism market may also differ [27, 28]. This may lead to differences in the behavior and decision- making of tourist consumers from different countries. Therefore, timely understanding of the impact of the COVID-19 pandemic on the tourism behavior and preferences of potential outbound tourism consumers from different countries plays a positive role in the recovery of inbound tourism in various

countries after the pandemic.

Based on the above background, this paper analyzes the yearly changes in the tourism behavior of Chinese tourists to Japan before the pandemic and explores the impact of the COVID-19 pandemic on the attitudes and preferences of potential outbound Chinese tourists. By comprehensively studying the behavior of Chinese tourists before and after the pandemic, it aims to provide effective policy recommendations for Japanese local governments to help them restore the tourism industry post-pandemic.

1. 2 LITERATURE REVIEW

1. 2. 1 The study on the tourism behavior of foreign tourists visiting Japan

Hishida [29] using the data on foreign visitors to Japan from the Japan Tourism Agency, analyzed the differences in the behavior of tourists visiting Japan from various regions of China to show changes in destination choice. Matsui [30] used the data of the "International Visitor Survey" of the Japan Tourism Agency to analyze changing trends in visiting locations and tourism behavior of foreign tourists to Japan. Furuya and Liu [31] also used the data of the "International Visitor Survey" of the Japan Tourism Agency, using the latent class model, to extract the visit patterns of foreign tourists to prefectures and to explain the relationship between the composition ratio of each visit pattern and factor such as nationality/region, departure time, and visit frequency. In addition, Furuya [32] used the data of the "International Visitor Survey" of the Japan Tourism Agency to classify the combined patterns of visiting places by the hierarchical Pachinko allocation model (hPAM) and indicated the relationship between nationality, region, visit frequency, and visit patterns. In particular, the study focused on the visit patterns that include municipalities and attractions, which are more detailed than prefecture classifications. Moreover, Tatsumi and Tsukai [33] analyzed the flow of foreigners-data (FF-data) from the "Ministry of Land, Infrastructure, Transport and Tourism" with a topic model. They extracted the visit patterns of foreign tourists who visited Japan and analyzed its changes over time. However, the data used in the abovementioned studies were generally collected from the questionnaire surveys and the interviews with foreign tourists. Such surveys usually only capture the tourists' behaviors by prefectures, and it is difficult to understand the specific visited places (by city), attractions, and travel experiences of tourists.

1. 2. 2 The study on the application of online travel text data

With the massive growth of the Internet, text data has become one of the main formats of tourism big data. As an effective means of expression of tourists' opinions, text-based tourism data mining has the potential to transform the tourism industry [34]. Currently, web text mining, machine learning, and other technologies that analyze online travel text data are mainly used to create tourist profiles [35, 36]

to develop tourism value analysis models [37, 38], to build tourism recommendation systems [39, 40, 41], to provide tourism personalized recommendation [42] and to generate tourism destination image perception [43]. In addition, some studies apply sentiment analysis for examining online travel review texts [44, 45, 46] and for hotel reviews [47], as well as conduct a satisfaction analysis of the hotel [48, 49, 50]. Mou [51] used the online travel diaries data as the digital footprint, and analyzed the spatial patterns of tourism flows in Qingdao by using the gravity center model, association rules mining, and social network analysis. Then an uneven distribution of tourist visits across destinations and the structural hole phenomenon were found. Song and Furuya 52) used the LDA model to categorize the visit patterns and travel contents and display the travel behaviors and characteristics of Chinese tourists who visited Japan by linking them with the personal and travel information of travel notes. However, this study focused on primary tourism sites such as Tokyo, Kyoto and Osaka rather than local regions, and they did not discuss the changes of topics over the time.

1. 2. 3 The study on the impact of covid-19 pandemic on tourists' behavior and preferences

Some scholars have investigated the impact of the COVID-19 pandemic on changes in the travel intentions, restricting factors, and preferences of tourism consumers in certain countries. Liu [53] found that the willingness of Chinese residents to travel was generally low in June and July 2020. The perception of COVID-19 had a negative impact on Chinese residents' outbound travel, while risk tolerance had a negative moderating effect on the association between perception of COVID-19 after the pandemic and travel intentions. Wen [54] summarized the constantly changing consumption patterns of Chinese tourists based on news related to "COVID-19" and "tourism" reported by multiple media outlets in 2020, supplemented by an overview of literature on tourism marketing, management, and behavior. The authors concluded that the consumption patterns of Chinese tourists will include travel choices based on nature and a trend towards "slow travel," which emphasizes longer stays and more fulfilling travel experiences. However, a survey conducted on Costa Rican residents from March 15 to April 15, 2020, showed that, in addition to concerns about healthcare issues related to COVID-19 when deciding to visit a travel destination, potential tourism consumers were more inclined to travel with family or friends for short trips of 2-3 days [55]. Furthermore, the studies of Santos [56] and Marques [57] indicate that, after the lifting of domestic mobility restrictions in Portugal, tourists have shown a significantly increased preference for tourism in low-density areas (rural areas), and regions that are best suited to provide rural accommodation have experienced a stronger and more sustained recovery in domestic demand. However, Orden-Mejía [58] conducted a survey on the preferences, attitudes, and expectations of Guayaquil residents in Ecuador when planning travel after the pandemic (April 2020 to May 2020). The results indicate that, after the outbreak, residents of the area showed a greater preference for urban tourism, followed by cultural tourism and family tourism. In addition, Park [59] recruited 197 adult participants from the United States to investigate how the COVID-19

pandemic affected their tourists' preferences for crowded and non-crowded options. The results indicated that, when the pandemic is prominent, tourists tend to reduce their preferences for crowded options and choose non-crowded ones instead. However, for participants with high sensory-seeking and uniqueness needs, the tendency to avoid crowded options is significantly reduced. NAKAMURA [60] conducted a survey on the intention of Japanese people to travel abroad in February 2021, which showed that before the COVID-19 pandemic stabilizes, most Japanese people try to avoid traveling abroad. Even for those who choose to travel abroad, they tend to choose countries/regions that are closer to Japan or English-speaking countries/regions. Additionally, they expect to have travel experiences that are close to normal with fewer restrictions, and they are less willing to accept various restrictions and conditions when traveling overseas. Based on the aforementioned literature, it can be found that the COVID-19 pandemic has had varying impacts on the tourism preferences of different groups in different countries.

1. 3 The purpose of the study

This study aims to provide reference for the recovery strategy of inbound tourism in Japan's local regions after the COVID-19 pandemic. In Chapter 2, we used the latent Dirichlet allocation model and Word2vec model of natural language processing to analyze the web travel notes posted on the Chinese visitor website Mafengwo written by Chinese tourists who visited Japan. Using the analytical results, we identified the tourism themes and the visit patterns of the Chinese tourists travelling to Japan and how they have changed. Furthermore, we also elucidated impressive and attractive local and regional tourism resources in some regions in Japan. Based on the analytical results, the policy to promote inbound tourism industry in local regions was also discussed. In Chapter 3, we conducted an online survey in August 2022 with Chinese respondents who had previous outbound travel experience. The purpose was to timely understand the attitude and preference changes of Chinese travel consumers towards outbound travel after the COVID-19 pandemic. However, considering the different infection situations and prevention measures of COVID-19 in different countries and regions, this may lead to differences in the attitudes and preferences of tourist consumers towards outbound travel in different countries. Therefore, in September of the same year, we conducted the same online survey with Japanese respondents who had previous outbound travel experience and had significant differences in their COVID-19 prevention policies from China, aiming to understand the differences in the impact of the COVID-19 pandemic on outbound travel among different national tourist consumer groups.

1. 4 The composition of the study

This study consists of four parts in total.

Chapter 1, summarizing research background and previous studies, clarifying the positioning and purposes of this study.

Chapter 2, analyzes the tourism behaviors and changes of Chinese tourists visiting Japan before the pandemic, with a particular emphasis on their behavior changes in local regions.

Chapter 3, The attitudes, perception risks, and preference changes of potential tourists from China and Japan towards outbound travel were investigated after the COVID-19 pandemic.

Chapter 4, as an overview, the conclusions of each chapter were summarized.

References

1) Japan Tourism Agency homepage. Available at: https://www.mlit.go.jp/kankocho/seisaku_seido/suishintaisei/kakuryokaigi/juzen.html
2) White Paper on Tourism, Japan Tourism Agency. Available at: https://www.mlit.go.jp/kankocho/en/siryou/whitepaper.html
3) International Visitor Survey (2012-2020), Japan Tourism Agency. Available at: https://www.mlit.go.jp/kankocho/en/siryou/toukei/syouhityousa.html
4) Yagasaki, N.: Significance, trends and issues of tourism by foreign tourists visiting Japan. IATSS Review, Vol. 45(1), pp. 6-17, 2020
5) Uğur, N. G., Akbıyık, A.: Impacts of COVID-19 on global tourism industry: A cross-regional comparison. Tourism management perspectives, Vol. 36, 100744, 2020
6) CONNOR, P.: More than nine-in-ten people worldwide live in countries with travel restrictions amid COVID-19. 2020
7) Luković, S., Stojković, D.: Covid-19 pandemic and global tourism. Hotel and Tourism Management, Vol. 8(2), pp. 70-88, 2020
8) Gössling, S., Scott, D., Hall, C.M.: Pandemics, tourism and global change: A rapid assessment of COVID-19. Journal of Sustainable Tourism, Vol. 29(1), pp. 1–20, 2021
9) Sharma, A., Nicolau, J. L.: An open market valuation of the effects of COVID-19 on the travel and tourism industry. Annals of Tourism Research, Vol. 83, 102990, 2020
10) UNWTO.: IMPACT ASSESSMENT OF THE COVID-19 OUTBREAK ON INTERNATIONAL TOURISM, https://www.unwto.org/impact-assessment-of-the-covid-19-outbreak-on-international-tourism
11) Kupi, M., Szemerédi, E.: Impact of the COVID-19 on the Destination Choices of Hungarian Tourists: A Comparative Analysis. Sustainability, Vol. 13(24), 13785, 2021
12) Ivanova, M., Ivanov, I. K., Ivanov, S.: Travel behaviour after the pandemic: The case of Bulgaria. Anatolia, Vol. 32(1), pp. 1-11, 2021
13) Mirzaei, R., Sadin, M., Pedram, M.: Tourism and COVID-19: changes in travel patterns and tourists' behavior in Iran. Journal of Tourism Futures, Vol. 9(1), ISSN: 2055-5911, 2021
14) Riestyaningrum, F., Ferdaos, E., Bayramov, B.: Customer behavior impact on international tourist's travel intention due to Covid-19. Journal of Sustainable Tourism and Entrepreneurship, Vol. 1, No. 3, pp. 231-243, 2020
15) Orîndaru, A., Popescu, M.-F., Alexoaei, A.P., Căescu, Ş.-C., Florescu, M.S., Orzan, A.-O.: Tourism in a Post-COVID-19 Era: Sustainable Strategies for Industry's Recovery. Sustainability, Vol. 13(12), 6781, 2021
16) Huang, S., Shao, Y., Zeng, Y., Liu, X., Li, Z.: Impacts of COVID-19 on Chinese nationals' tourism preferences. Tourism Management Perspectives, Vol 40, ISSN 2211-9736, 2021

17) Li, Z., Zhang, X., Yang, K., Singer, R., Cui, R.: Urban and rural tourism under COVID-19 in China: research on the recovery measures and tourism development. Tourism Review, Vol. 76(4), pp. 718-736, 2021
18) Kinczel, A., Müller, A.: STUDY ON TRAVEL HABITS AND LEISURE ACTIVITIES IN THE LIGHT OF COVID-19 TRIGGERED CHANGES IN ROMANIA AND HUNGARY. GeoJournal of Tourism and Geosites, Vol. 41, No. 2, pp. 440-447, 2022
19) Umukroro, G. M., Odey, V.E., Yta, E.M.: The effect of pandemic on homebased tourism: post Covid-19. International Journal of Humanities and Innovation (IJHI), Vol. 3, No. 3, pp. 115-120, 2020
20) Fan, X., Lu, J., Qiu, M., Xiao X.: Changes in travel behaviors and intentions during the COVID-19 pandemic and recovery period: A case study of China. Journal of Outdoor Recreation and Tourism, Vol. 41, 100522, 2022
21) Marques, C.P., Guedes, A., Bento, R.: Rural tourism recovery between two COVID-19 waves: the case of Portugal. Current Issues in Tourism, Vol. 25(6), pp. 857-863, 2022
22) Wachyuni, S.S., Kusumaningrum, D.A.: The Effect of COVID-19 Pandemic: How are the Future Tourist Behavior?. Journal of Education, Society and Behavioural Science, Vol. 33(4), pp.67-76, 2020
23) Jafari, K., Saydam, M. B., Erkanlı, E., Olorunsola, V. O.: The impacts of the COVID-19 pandemic on the consumer behavior of Turkish tourists. Revista Turismo Estudos e Práticas-RTEP/UERN, No. 5, pp. 1-17, 2020
24) Popescu, A.: The impact of COVID-19 pandemic on Romania's tourist flows in the year 2020. Scientific Papers Series Management, Economic Engineering in Agriculture and Rural Development, Vol. 21(1), pp. 655-666, 2021
25) Wen, J., Kozak, M., Yang, S., Liu, F.: COVID-19: potential effects on Chinese citizens' lifestyle and travel. Tourism Review, Vol. 76(1), pp. 74-87, 2021
26) Afanasiev, O.E., Afanasieva, A.V.: The Effect of COVID-19 Pandemic on Russian Tourists' Demand and Preferences. Tourism Through Troubled Times, pp. 203-230, 2022
27) Aronica, M., Pizzuto, P., Sciortino, C.: COVID-19 and tourism: What can we learn from the past?. The World Economy, Vol. 45(2), pp. 430-444, 2022
28) Chemli, S., Toanoglou, M., Valeri, M.: The impact of Covid-19 media coverage on tourist's awareness for future travelling. Current Issues in Tourism, Vol. 25(2), pp. 179-186, 2022
29) Hishida, N., Hibino, N., Morichi, S.: Chinese Tourists in Japan: Time-Series Analysis in Their Destination Choices and its Multiplicities. Journal of Japan Society of Civil Engineers D3, Vol. 68(5), pp. I_667-I_677, 2012
30) Matsui, Y., Hibino, N., Morichi, S., Ieda, H.: Behavior analysis on inbound tourists to Japan focus on their destinations and activities based on individual data. Journal of Japan Society of Civil Engineers D3, Vol. 72(5), pp. I_533-I_546, 2016
31) Furuya, H., Liu, Y.: Analysis of combination of visiting places by foreign visitors in Japan by Latent

Class Model. Journal of Japan Society of Civil Engineers D3, Vol. 72(5), pp. I_571-I_583, 2016

32) Furuya, H.: Study on identification of destinations combination pattern of foreign visitors to japan based on similarity by hPAM. Journal of Japan Society of Civil Engineers D3, Vol. 75(5), pp. I_507-I_517, 2019

33) Tatsumi, Y., Tsukai, M.: A tour of analysis of foreign visitors to Japan by applying topic model. Transport Policy Studies' Review, Vol. 23, pp. 20-34, 2021. Available at: https://www.jstage.jst.go.jp/article/tpsr/advpub/0/advpub_TPSR_23R_04/_pdf/-char/ja

34) Li, Q., Li, S., Zhang, S., Hu, J., Hu, J.: A review of text corpus-based tourism big data mining. Applied Sciences, Vol. 9(16), 3300, 2019

35) Hao, Q., Cai, R., Wang, C., Xiao, R., Yang, J. M., Pang, Y., Zhang, L.: Generating location overviews with images and tags by mining user-generated travelogues. Proceedings of the 17th ACM international conference on Multimedia, pp. 801–804, 2009

36) Vu, H. Q., Li, G., Law, R.: Discovering implicit activity preferences in travel itineraries by topic modeling. Tourism Management, Vol. 75, pp. 435-446, 2019

37) Abeysinghe, S., Manchanayake, I., Samarajeewa, C., Rathnayaka, P., Walpola, M. J., Nawaratne, R., Bandaragoda, T., Alahakoon, D.: Enhancing decision making capacity in tourism domain using social media analytics. International Conference on Advances in ICT for Emerging Regions, 2018

38) Haris, E., Gan, K.H., Tan, T.-P.: Spatial information extraction from travel narratives: Analysing the notion of co-occurrence indicating closeness of tourist places. Journal of Information Science, Vol. 46(5), pp. 581–599, 2020

39) Yuan, H., Xu, H., Qian, Y., Li, Y.: Make your travel smarter: summarizing urban tourism information from massive blog data. International Journal of Information Management, Vol. 36(6), pp. 1306-1319, 2016

40) Hayashi, T.: Applying the document vector model to tour recommendation. Journal of Global Tourism Research, Vol. 4, No. 1, pp. 15-20, 2019

41) Shafqat, W., Byun, Y.-C.: A recommendation mechanism for under-emphasized tourist spots using topic modeling and sentiment analysis. Sustainability, Vol. 12(1), 320, 2020

42) Du, S., Zhang, H., Xu, H., Yang, J., Tu, O.: To make the travel healthier: a new tourism personalized route recommendation algorithm. Journal of Ambient Intelligence and Humanized Computing. Vol. 10(5), pp. 3551–3562, 2019

43) Chen X., Li J, Han W, Liu S.: Urban tourism destination image perception based on LDA integrating social network and emotion analysis: the example of Wuhan. Sustainability, Vol. 14(1), 12, 2021

44) He, W., Tian, X., Tao, R., Zhang, W., Yan, G. and Akula, V.: Application of social media analytics: a case of analyzing online hotel reviews. Online Information Review, Vol. 41, pp. 921-935, 2017

45) Ren, G.; Hong, T.: Investigating online destination images using a topic-based sentiment analysis approach. Sustainability, Vol. 9(10), 1765, 2017

46) Zheng, X., Luo, Y., Sun, L., Zhang, J., Chen, F.: A tourism destination recommender system using users' sentiment and temporal dynamics. Journal of Intelligent Information Systems, Vol. 51(3), pp. 557-578, 2018

47) Nakamura, Y., Oomiya, N.: Applying the topic model to hotel reviews of internet sites and analyzing their results. Journal of Global Tourism Research, Vol. 6, No. 1, pp. 15-20, 2021

48) Guo, Y., Barnes, S.J., Jia, Q.: Mining meaning from online ratings and reviews: Tourist satisfaction analysis using latent dirichlet allocation. Tourism Management, Vol. 59, pp. 467–483, 2017

49) Hu, N., Zhang, T., Gao, B., Bose, I.: What do hotel customers complain about? Text analysis using structural topic model. Tourism Management, Vol. 72, pp. 417-426, 2019

50) Zhai, Y., Chen, P.: Sentiment analysis on tourist satisfaction with rural homestay inns based on reviews from the website of online travel agency. International Journal of Sustainable Development and Planning, Vol. 15, No. 5, pp. 705-712, 2020

51) Mou, N., Zheng, Y., Makkonen, T., Yang, T., Tang, J., Song, Y.: Tourists' digital footprint: the spatial patterns of tourist flows in Qingdao China. Tourism Management, Vol. 81, 104151, 2020

52) Song, Z., Furuya, H.: Analysis of Chinese tourists' travel literatures by Latent Dirichlet Allocation Model. Proceedings of Research Meeting on Civil Engineering Planning, 2018. Available at: http://lbrary.jsce.or.jp/jsce/open/00039/201806_no57/57-01-01.pdf

53) Liu, Y., Shi, H., Li, Y. and Amin, A.: Factors influencing Chinese residents' post-pandemic outbound travel intentions: an extended theory of planned behavior model based on the perception of COVID-19. Tourism Review, Vol. 76(4), pp. 871-891, 2021

54) Wen, J., Kozak, M., Yang, S., Liu, F.: COVID-19: potential effects on Chinese citizens' lifestyle and travel. Tourism Review, Vol. 76(1), pp. 74-87, 2021

55) Moya Calderón, M., Chavarría Esquivel, K., Arrieta García, M. M., Lozano, C. B.: Tourist behaviour and dynamics of domestic tourism in times of COVID-19. Current Issues in Tourism, Vol. 25(14), pp. 2207-2211, 2022

56) Santos, N., Moreira, C. O.: Uncertainty and expectations in Portugal's tourism activities. Impacts of COVID-19. Research in Globalization, Vol. 3, 100071, 2021

57) Marques, C.P., Guedes, A., Bento, R.: Rural tourism recovery between two COVID-19 waves: the case of Portugal. Current Issues in Tourism, Vol. 25(6), pp. 857-863, 2022

58) Orden-Mejía, M., Carvache-Franco, M., Huertas, A., Carvache-Franco, W., Landeta-Bejarano, N., Carvache-Franco, O.: Post-COVID-19 tourists' preferences, attitudes and travel expectations: A Study in Guayaquil, Ecuador. International Journal of Environmental Research and Public Health, Vol. 19(8), 4822, 2022

59) Park, I. J., Kim, J., Kim, S. S., Lee, J. C., Giroux, M.: Impact of the COVID-19 pandemic on travelers' preference for crowded versus non-crowded options. Tourism Management, Vol. 87, 104398, 2021

60) NAKAMURA, T.: Intention of Japanese to Travel Abroad During and After the COVID-19 Pandemic. The journal of Tamagawa University College of Tourism and Hospitality, No. 9, pp. 43-61, 2022 (Japanese)

CHAPTER 2
Analysis of the tourism experiences of Chinese tourists to Japan using natural language processing techniques: focusing on local regions

2.1 Introduction

In 2019, the number of foreign visitors who visited Japan was 31.88 million, the tourism expenditure was 4,813.5 billion yen. Among them, Chinese tourists accounted for 8.57 million visitors and spent 1.77 trillion yen on travel, both of which were record highs. Since 2015, Chinese tourists have held the top spot in terms of the number of inbound tourists to Japan and have played an important role in the country's tourism industry [1]. According to the statistical data of Japan Tourism Agency "International Visitor Survey," the per capita travel consumption of Chinese tourists who visited Japan increased sharply from 2014 to 2015, and a phenomenon called shopping spree emerged. However, since 2016, with the easing down of the Chinese tourist shopping spree, the per capita travel consumption of Chinese tourists who visited Japan has decreased. Nonetheless, Japan's tourism market remains highly attractive to Chinese visitors, with its food, culture, and scenery drawing more and more Chinese tourists to visit, besides shopping. Furthermore, with the increase of Chinese tourists who visited Japan for personal tourism and repeated visits to Japan, the number of Chinese visitors visiting local regions, in addition to the capital region (Tokyo, Kanagawa, Saitama, Chiba), Keihanshin region (Osaka, Kyoto, and Hyogo) and Aichi prefecture, has also been on an upward trend [2]. It can be speculated that the needs and tourism behaviors of Chinese tourists traveling to Japan are becoming more and more diverse. Therefore, in order to maintain the continued appeal of Japan's tourism industry to Chinese tourists, it is necessary to understand the changing needs and tourism behaviors of Chinese visitors, and in particular, significant attention should be paid to the changes in their tourism behaviors in local regions. The development of inbound tourism is expected to contribute to the revitalization of local economy. On the other hand, local regions are considered to be at a relative disadvantage in terms of tourists' recognition and mobility convenience. Consequently, strategic tourism promotion is necessary particularly for local regions.

In the local regions, traditional methods of questionnaire surveys or interviews are limited by time and location due to the relatively low number of visitors, making it difficult to obtain large-scale data of tourist experience data for analysis. In contrast, tourists are increasingly willing to record and share their travel experiences on weblogs, forums, or travel communities using textual travelogues and photos taken during the trips [3]. We refer to the textual travelogues as "travel notes" in this study. As this is spontaneously generated by consumers, it reflects the perceptions that are not easily obtained by other survey methods. The benefit of this unsolicited content is that people may be more open and honest online than in face-to-face situations [4, 5], and the travel notes contain a lot of relevant

information such as place names, attractions, tourism resources, and tourism experiences. By analyzing such data, valuable information such as visit patterns and tourism themes of tourists can be obtained. Although the information in a single travel note is possibly noisy or biased, numerous travelogues as a whole could reflect people's overall preference and understanding of travel resources, and thus, can serve as a reliable knowledge source. As most travel note data collected from social media are unstructured, natural language processing techniques are required to assist with data analysis.

This study uses the Latent Dirichlet Allocation (LDA) and word to vector (Word2vec) models of natural language processing to analyze the web travel notes written by Chinese tourists who visited Japan, which were posted on the Mafengwo, with the aim of revealing the tourism themes and visit patterns and their changes for Chinese tourists who visited Japan over the years. Furthermore, we also tried to limelight the impressive and attractive regional resources of local regions for Chinese tourists who visited Japan. Finally, based on the analytical results, the policy to promote inbound tourism in local regions is also discussed.

2.2 literature review

With the massive growth of the Internet, text data have become one of the main formats of tourism big data. As an effective means of expression of tourists' opinions, text based tourism data mining has the potential to transform the tourism industry [6]. Currently, web text mining, machine learning, and other technologies that analyze online travel text data are mainly used to create tourist profiles [7, 8] to develop tourism value analysis models [9, 10], to generate tourism destination image perception [11]. to build tourism recommendation systems [12-14], and to provide tourism personalized recommendation [15], In addition, some studies apply sentiment analysis for examining online travel review texts [16-18] and hotel reviews [19], as well as analyzing satisfaction levels during the hotel stays [20-22]. For example, Quiaoit and Furuya [23] conducted an analysis of online reviews of the Philippine tourism experiences, employing methods such as VADER sentiment analysis, Word2vec, and clustering analysis. This analysis identified aspects of tourism that tourists disliked and which needed improvement. Furthermore, the study utilized Word2vec models to analyze word vectors associated with destination names and activity names, providing alternative travel destination suggestions and recommending other travel activities that matched tourists' interests. In contrast, this study utilizes the Word2vec model to analyze word vectors related to destination names with the aim of identifying attractive and impressive visiting locations, attractions, and local tourism resources for Chinese tourists in Japan, in order to gain a more detailed understanding of the tourism behavior of Chinese tourists in the local region of Japan. Additionally, this study also employed topic modeling to identify the tourism themes, visit patterns, and their variations for Chinese tourists traveling to Japan.

Previous studies on foreign tourists who visited Japan mainly used the statistical data from Japan Tourism Agency. Hishida et al. [24] using the data on foreign visitors to Japan from the Japan Tourism Agency, analyzed the differences in the behavior of tourists visiting Japan from various regions of China to show changes in destination choice. Matsui et al. [25] used the data of the "International Visitor Survey" of the Japan Tourism Agency to analyze changing trends in visiting locations and tourism behavior of foreign tourists to Japan. Furuya and Liu [26] also used the data of the "International Visitor Survey" of the Japan Tourism Agency, using the latent class model, to extract the visit patterns of foreign tourists to prefectures and to explain the relationship between the composition ratio of each visit pattern and factors such as nationality/region, departure time, and visit frequency. In addition, Furuya [27] used the data of the "International Visitor Survey" of the Japan Tourism Agency to classify the combined patterns of visiting places by the hierarchical Pachinko allocation model (hPAM) and indicated the relationship between nationality, region, visit frequency, and visit patterns. In particular, the study focused on the visit patterns that include municipalities and attractions, which are more detailed than prefecture classifications. Tatsumi and Tsukai [28] analyzed the flow of foreigners-data (FF-data) from the "Ministry of Land, Infrastructure, Transport and Tourism" with a topic model. They extracted the visit patterns of foreign tourists who visited Japan and analyzed their changes over time. However, the data used in the abovementioned studies were generally collected from the questionnaire surveys and the interviews with foreign tourists. Such surveys usually only capture the tourists' behaviors by prefectures, and it is difficult to understand the specific visited places (by city), attractions, and travel experiences of tourists. In the study reported in 2018, Song and Furuya [29] analyzed the travel notes that described the travel progress and impressions of tourists. They used the LDA model to categorize the visit patterns and travel contents and display the travel behaviors and characteristics of Chinese tourists who visited Japan by linking them with the personal and travel information of travel notes. However, their study focused on primary tourism sites such as Tokyo, Kyoto, and Osaka rather than local regions, and they did not discuss the changes in topics over time.

The purpose of this study is to analyze the tourism themes and the visit patterns and their changes over time by using the online travel notes. In the analysis, we try to detect the visit patterns, including relatively "minor" tourism sites that are remote from large cities.

2.3 Analytical method

2.3.1 Outline of analytical framework

Based on data from the Japan Tourism Agency, the number of foreign visitors to Japan has started to rebound from October 2022. However, the number of Chinese tourists visiting Japan did not

significantly increase until March 2023. Due to limited data available for analyzing the tourism behavior of Chinese tourists, this study used travel notes data from before the pandemic.

This study used the web crawler technology built in the programming language (Python) to collect the travel notes of Chinese tourists who visited Japan and posted travel notes on the Chinese visitor website Mafengwo, and the personal information of travel notes' writers. From April 14, 2021, to May 30, 2021, a total of 37,259 travel notes contributed by 31,012 authors were collected. These travel notes excluded pre-travel descriptions. As shown by the non-bold values (the statistics of Mafengwo) in **Table 2.1**, the departure time to Japan from 2015 to 2019 accounted for a relatively large proportion. The percentage of females is greater than that of males. The distribution of the departure season is relatively even, with most stays between 7 and 13 days. Regarding accompanying persons, the proportions of lovers/spouses, family members, and friends are about the same.

Additionally, based on statistics from the Japan Tourism Agency (The bold values in **Table 2.1**), from 2015 to 2019, more female than male Chinese tourists visited Japan each year. The proportion of tourists traveling to Japan in the summer is slightly higher than in other seasons. The length of stay was mainly concentrated in the range of four to seven days. In contrast to Mafengwo's data collection method, the Japan Tourism Agency only records the duration of visitors' stays within Japan, while Mafengwo's data record the overall time from their departure to their return to China. In terms of companions, the proportion of family members was relatively high. Although the sources and collection methods for the two groups of data are different, the statistical results in **Table 2.1** indicate that the two sets of data exhibit basic consistency in various statistical aspects. Therefore, we can preliminarily utilize data from Mafengwo to analyze the trends in Chinese tourists traveling to Japan.

In addition, Mafengwo jointly established the "Joint Laboratory of Big Data on Independent Tourism" with the China Tourism Academy in 2018 [30]. It is based on the travel information of users, travel notes, comments, and transaction data that delve deeply into the tourism behavior of Chinese tourists; it regularly prepares and publishes reports. Therefore, the travel note data from the Mafengwo are considered reliable and persuasive. However, it should be noted that Mafengwo did not collect information on tourists' age, income, frequency of visits to Japan, and spending habits. Additionally, some users have not provided their personal information, which renders this data unsuitable for in-depth demographic analysis.

Table 2.1 The summary of travel notes' data.

Name	Options	Quantity（articles）	Proportion
Number of travel notes recorded		37259	100%
Departure times of travel note writers	~ 2013	561	1.50%
	2014	825	2.20%
	2015	1963	5.30%
	2016	3660	9.80%
	2017	5646	15.20%
	2018	7890	21.20%
	2019	6692	18.00%
	2020 ~	623	1.70%
	Null	9399	25.20%
Gender	Male	8129	21.80% / **37.3%**
	Female	23026	61.80% / **62.7%**
	Null	6104	16.40%
Departure time (Month)	Spring (3 ~ 5)	7439	20.00% / **22.7%**
	Summer (6 ~ 8)	6939	18.60% / **30.8%**
	Autumn (9 ~ 11)	7201	19.30% / **24.5%**
	Winter (12 ~ 2)	6281	16.90% / **22%**
	Null	9399	25.20%
The length of stay (days)	~ 3	968	2.60% / **1.1%**
	4 ~ 6	6976	18.70% / **56.1%**
	7 ~ 13	18027	48.40% / **40.3%**
	14 ~ 20	1394	3.70% / **1.8%**
	21 ~	277	0.70% / **0.7%**
	Null	9617	25.80%
Accompanying person	Personage	2862	7.70% / **9.3%**
	Lovers / Spouse	7059	18.70% / **17%**
	Family	7073	19.00% / **42%**
	Classmate	419	1.10% / **5.8%**
	Friend	8973	24.10% / **25.2%**
	Others	1161	3.10% / **0.7%**
	Null	9712	26.10%

Note: Null is empty value. The bold values represent the mean percentage for each project from 2015 to 2019. These data are from the statistics of the Japan Tourism Agency.

This study uses the LDA and Word2vec models to analyze the contents of the travel notes posted on tourism websites in China by Chinese tourists who visited Japan to explore the changes in tourists' travel behavior over time thereby. The LDA is a topic modeling technique used to identify latent topics within a collection of documents. Its assumption is that each document is composed of multiple topics, each represented by a set of words, and it attempts to find these topics and their probability distributions within the documents [31, 32]. Therefore, LDA primarily focuses on the topic structure and does not directly consider the proximity of words in sentences. In contrast, Word2vec is a shallow neural network model that learns vector representations for each word in the vocabulary, extracting semantic relationships between adjacent words in actual sentences. This means that if two words have similar representations in the Word2vec vector space, they tend to appear in similar contexts (the context consists of several surrounding words), making Word2vec more effective at capturing the meanings and contextual relationships between words [33].

In addition, the LDA model allows a word to have multiple meanings based on cooccurrence relationships, which helps assign words to multiple topics, thus better capturing the diversity and complexity of travel note content. In contrast, the Word2vec model may suffer from the problem of polysemy, as it provides only one vector representation for each word. This can lead to unclear expressions of multiple meanings or low correlation between vectors. Given that the output of the Word2vec model in this study was related to place names, we introduced a custom dictionary containing information on 5841 Japanese place names and tourist attractions when preprocessing the travel note content. The purpose of this step is to enhance the accuracy of the outputs from both models and reduce ambiguity in the Word2vec output results.

In summary, LDA is mainly used for topic modeling and analysis, aiding in understanding the thematic structure of documents, while Word2vec is more suitable for word embeddings and semantic analysis, helping to understand the semantic associations between words and phrases in documents. As shown in **Fig. 2.1**, in this study, the LDA model and the word2vec model share a corpus (the word segmentation results of travel notes data). The feature words of each topic extracted by the LDA model can be marked as the tourism theme and visit pattern. Thereafter, by using the place names in the visit pattern as the input of the word2vec model, the words with more high similarity to the names of the places are obtained as the output to understand the detailed travel behavior in various regions of tourists.

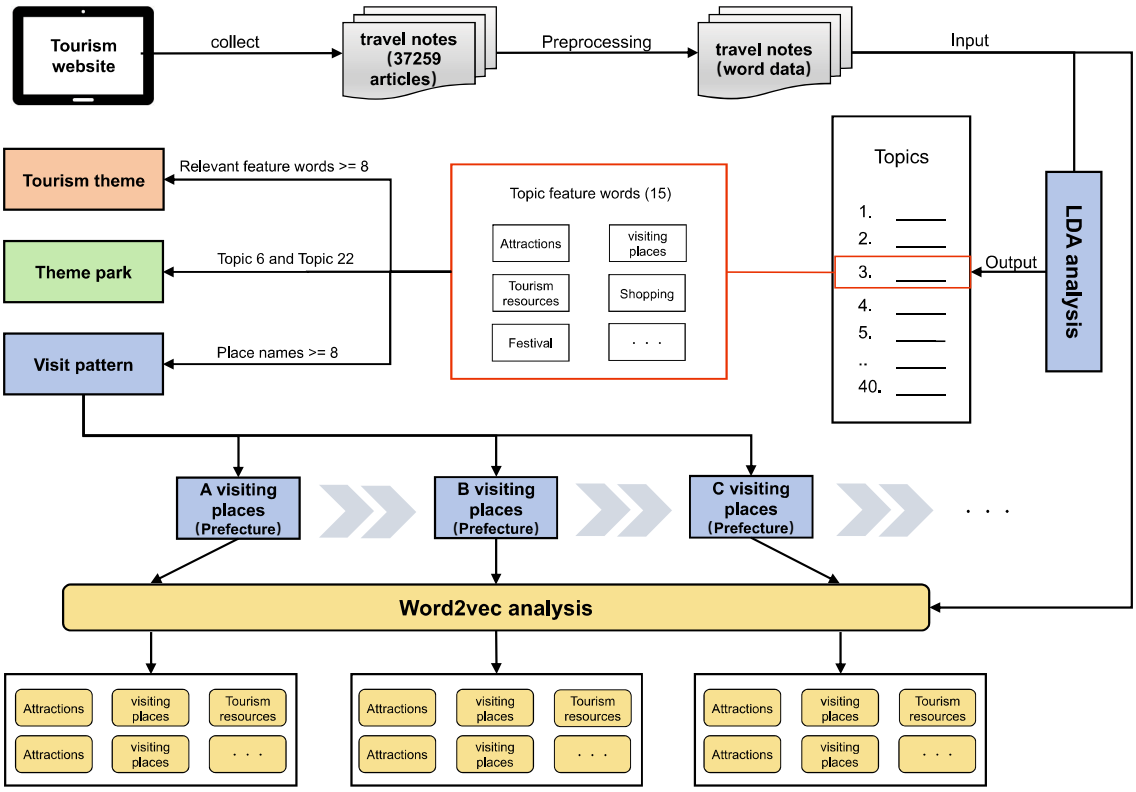


Fig. 2.1 The combined analysis process of the LDA and Word2vec models in this study

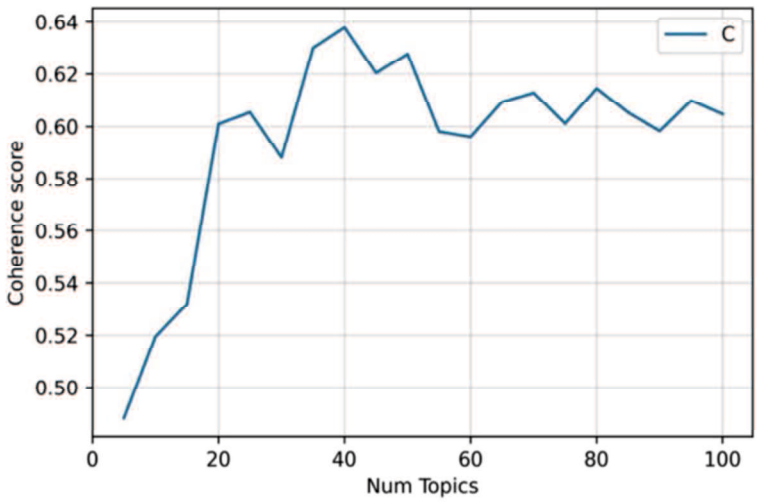
2.3.2 Data preprocessing

This study used the Jieba package in Python to segment travel note data and the sentences were decomposed into individual words. Then, conjunctions, adverbs, interjections/modal particles, idioms, onomatopoeia, prepositions, quantifiers, pronouns, auxiliary words and some adjectives and verbs were combined with the Chinese stop words list on Gitee homepage to generate the stop words list (61991 words). The stop words were not used in the following analysis. In addition, a custom word table was generated from the thesaurus of Sogou input method, Baidu input method, and QQ input method, which are commonly used for Chinese input, utilizing words related to Japan place names, sightseeing spots, etc. A total of 5841 custom words specific to the context of the theme of this article were added to Jieba's dictionary. Finally, the stop word list and custom words list were imported into Python's Jieba package, and word segmentation on the travel data was performed again. After excluding morphemes with a frequency of 10 or less, we obtained 27,814,431 morphemes of 80,886 types. These morphemes were applied to the LDA and Word2vec model.

2.3.3 The LDA model

In this study, LDA model analysis was performed using Python's Gensim package. In LDA model analysis, it was necessary to determine the number of topics. I examined the coherence of the LDA model to determine the number of topics.

Figure 2.2 shows the relationship between the number of topics and the coherence score. Since the model giving the highest coherence value before flattening out or a major drop should be selected [34], 40 topics were set for the LDA model. In addition, the number of passes through the corpus during training was set to 10, and the minimum probability of topics included in each travel note was set to more than 0.01. In contrast, the random state was set to 80 to reproduce the result of executing the LDA model multiple times [35]. All the above settings were adopted to perform the LDA model analysis.



Note: Num Topics is the number of topics

Fig. 2.2 The evaluation index of the LDA model (Coherence).

2.3.4 The Word2vec model

Word2vec model includes two types of models, a continuous bag-of-words (CBOW) model and a skip-gram model [36, 37]. The skip-gram model has been found to work well with small amounts of data and represent rare words well, while the CBOW model is better suited for representing words that are highly similar to input words [33]. In order to understand tourist' behaviors in the local regions in detail, I expect the Word2vec model to output the words that are more similar to the input word (place name) rather than rare words. Therefore, this study employed the CBOW model for analysis. The input layer of the CBOW model consists of a one-hot encoded input context $\{X_1, ..., X_C\}$, where the window size is C, and the vocabulary size is V. The hidden layer is an N-dimensional vector. The output layer is also one-hot encoded and represents the output word y. One-hot encoding is a vector representation

method where each word corresponds to a unique one-hot vector. In this vector, only one position contains the value 1, and all other positions contain 0. The one-hot encoded input vectors are connected to the hidden layer through a weight matrix W of size $V \times N$, and the hidden layer is connected to the output layer through the weight matrix W' of size $N \times V$ [33]. The specific calculation process is as follows:

- The first step is to calculate the output of the hidden layer h.

$$h = (1/C)\, W \cdot \sum_{i=1}^{C} X_i \tag{1}$$

- The second step is to calculate the input at each node in the output layer.

$$u_j = v'^{T}_{wj} \cdot h \tag{2}$$

v'^{T}_{wj} is the j-th column of the output matrix W'.

- Finally, we calculate the output of the output layer, denoted as y_j.

$$Y_{c,j} = p(w_{y,j} \mid w_1, \ldots, w_c) = \exp(u_j) / \sum_{j'=1}^{V} \exp(u'j) \tag{3}$$

Based on the above formula, the Word2vec model sets the vector size (dimensionality of word vectors) to 300, the window C (maximum distance between the current and predicted word within a sentence) to 5, the number of epochs (number of iterations over the corpus) to 10, and the seed (seed for the random number generator) to 20. However, the model should be limited to a single worker thread (workers = 1) to reproduce the results of executing the Word2vec model multiple times [38]. All the above settings were adopted to perform the Word2vec model analysis.

2.4 Results and discussion

This section presents the results of the analysis of the LDA and Word2vec models using travel note data collected from Mafengwo. In the following, “metropolitan regions” refer to the eight prefectures of Tokyo, Kanagawa, Saitama, Chiba, Aichi, Osaka, Kyoto, and Hyogo, and “local regions” refer to the prefectures other than “metropolitan regions.”

2.4.1 The results of the LDA model

In this study, 40 topics were extracted by the LDA model, and the top 15 feature words and the

probability distribution of each feature word were output. As shown in **Table 2.2**, if there are eight or more feature words associated with a certain visit place in the top 15 feature words, the topic is also categorized as "visit pattern." The topics including eight or more feature words related to theme park are identified as "theme park." For example, topic 6 and topic 22 are called "Osaka (USJ)" and "Tokyo (Disneyland)," respectively. In the remaining topics, if there are eight or more feature words associated with a certain theme in the top 15 feature words, the topic is categorized as "tourism theme." A total of 20 visit patterns, 2 theme park, and 9 tourism themes are identified. Although most of the probability values for the topic feature words in **Table 2.2** are low, the low probability values may be due to the large amount of data used for analysis, and the variety of words that appear in the text. Most of the topic feature words point to their respective topics and rarely contain noisy words. Therefore, we consider the results of the LDA model to be effective.

Table 2.2 Example of extracted topics from the LDA model's results. (1)

Topic ID	The probability distribution of topic feature word (top 15 words)			Topic name
5	0.172*"Sakura"	0.013*"bloom"	0.008*"Sakura viewing"	Sakura (Tourism Theme)
	0.047*"park"	0.012*"Shinjuku Gyoen"	0.008*"Ueno Park"	
	0.024*"Sakura tree"	0.011*"Daigoji"	0.008*"petal"	
	0.021*"Philosophical way"	0.009*"shrine"	0.007*"Coinage"	
	0.016*"Sakura season"	0.009*"Meguro River"	0.007*"Hirano Shrine"	
6	0.145*"USJ"	0.016*"roller coaster"	0.010*"Movie"	Osaka (USJ) (Theme Park)
	0.038*"Harry Potter"	0.012*"Minions"	0.009*"castle"	
	0.028*"queue"	0.012*"Spiderman"	0.008*"experience"	
	0.026*"Osaka"	0.011*"tickets"	0.008*"magic"	
	0.017*"USJ"	0.010*"theme"	0.007*"Carcharodon carcharias"	
7	0.056*"Fireworks display"	0.023*"fireworks"	0.016*"fireworks"	Fireworks Display (Tourism Theme)
	0.040*"yukata"	0.022*"convention"	0.016*"Tagawa"	
	0.034*"summer"	0.019*"concert"	0.014*"festival"	
	0.034*"fireworks"	0.017*"summertime"	0.014*"convention place"	
	0.026*"summer day"	0.017*"court"	0.011*"stage"	
14	0.066*"Hiroshima"	0.023*"Kurashiki"	0.015*"Itsukushima Shrine"	Hiroshima_ Okayama_ Tottori (Visit Pattern)
	0.040*"Conan"	0.021*"Tottori"	0.012*"beautiful"	
	0.037*"Miyajima"	0.019*"Korakuen"	0.012*"shrine"	
	0.028*"Okayama"	0.018*"Hokuei Town"	0.011*"Matsuyama"	
	0.025*"small town"	0.015*"JR"	0.010*"Nojima"	
23	0.048*"Sendai"	0.014*"Tokyo"	0.012*"Aomori Prefecture"	Tohoku region (Skiing) (Visit Pattern)
	0.030*"Aomori"	0.013*"JR"	0.011*"Shinkansen"	
	0.027*"skiing"	0.013*"Akita"	0.011*"Hirosaki Park"	
	0.021*"ski resort"	0.012*"northeast region"	0.009*"Snow pack"	
	0.015*"Matsushima"	0.012*"Kakunodate"	0.009*"beef tongue "	

Table 2.2 Example of extracted topics from the LDA model's results. (2)

Topic ID	The probability distribution of topic feature word (top 15 words)			Topic name
29	0.048*"Fukuoka"	0.023*"Kagoshima"	0.015*"Beppu"	
	0.038*"Kumamoto"	0.020*"Jigoku"	0.014*"JR"	Kyushu
	0.033*"Hakata"	0.017*"Nagasaki"	0.012*"Dazaifu"	region
	0.026*"Kyushu"	0.016*"Kumamoto Castle"	0.010*"Saga"	(Visit Pattern)
	0.024*"Yufuin"	0.016*"hot spring"	0.010*"Aso"	
36	0.140*"Okinawa"	0.016*"dolphin"	0.012*"Churaumi"	Okinawa
	0.057*"aquarium"	0.015*"beach"	0.011*"Shurijo Castle Park"	(Visit Pattern)
	0.033*"Naha"	0.013*"Manzamo"	0.011*"diving"	
	0.024*"American Village"	0.012*"park"	0.010*"Shuri Castle"	
	0.022*"Kokusaidori"	0.012*"hotel"	0.009*"Churaumi Aquarium"	
37	0.074*"art gallery"	0.017*"works"	0.012*"Ritsurin park"	Seto Inland
	0.040*"Naoshima"	0.014*"Seto Inland Sea"	0.011*"olives"	Sea and
	0.033*"Takamatsu"	0.014*"art"	0.011*"pumpkin"	Takamatsu
	0.027*"park"	0.013*"Takamatsu City"	0.008*"angel"	(Art)
	0.019*"Toshima"	0.012*"on the island"	0.008*"architecture"	(Visit Pattern)
…	……	……	……	……
40	0.060*"child"	0.017*"zoo"	0.013*"father"	
	0.035*"kid"	0.015*"children"	0.012*"adult"	Family trip
	0.026*"mother"	0.015*"Museum"	0.011*"Child"	(Tourism
	0.022*"baby"	0.014*"husband"	0.011*"Kaiyukan"	Theme)
	0.020*"son"	0.014*"daughter"	0.010*"Mom"	

Note: The repeated words of some topics in **Table 2.2** indicates that the word is extracted twice as Chinese characters and Japanese characters in the LDA model results. Moreover, multiple Chinese words may appear to refer to the same English word pair, so the same English word may appear more than once in the feature words of the topics.

LDA can generate a topic distribution for each travel note to represent the weight of each topic within the travel note. In this study, we define the association between a travel note and a specific topic as follows: if the probability value of a topic in one travel note is greater than 0.01, the travel note is considered as "corresponding" to the topic. The proportion of the travel notes corresponding to the topic is called the "corresponding ratio (R)." Since a single travel note may correspond to multiple topics, the sum of the R-values for all topics in **Fig. 2.3** and **Fig. 2.4** exceeds 1. The definition of the corresponding ratio R can be shown as follows:

$$R = N/S \tag{4}$$

R : Corresponding ratio of a topic.

S : Total number of travel notes.

N : The number of the travel notes corresponding to the topic.

Figure 2.3 shows the corresponding ratio (*R*) of visit patterns and theme parks extracted by the LDA model, including some visit patterns and theme parks not shown in **Table 2.2**. The top 10 in terms of corresponding ratio are mostly related to the metropolitan regions (Osaka, Kyoto, Tokyo, Kanagawa, Chiba) and surrounding regions (Nara, Shizuoka). On the other hand, in the local regions, the corresponding ratios (*R*) of the visit patterns of Izu Peninsula and Hokkaido are relatively higher. Additionally, the visit patterns in local regions, such as Seto Inland Sea /Takamatsu, Kyushu, Tohoku, Takayama/Kanazawa, Okinawa, North Kanto, and Hiroshima_Okayama_Tottori, are identified. Although the corresponding ratios of these visit patterns in the local regions are not high compared with those in metropolitan regions, the abovementioned regions can be regarded as the relatively popular visited places.

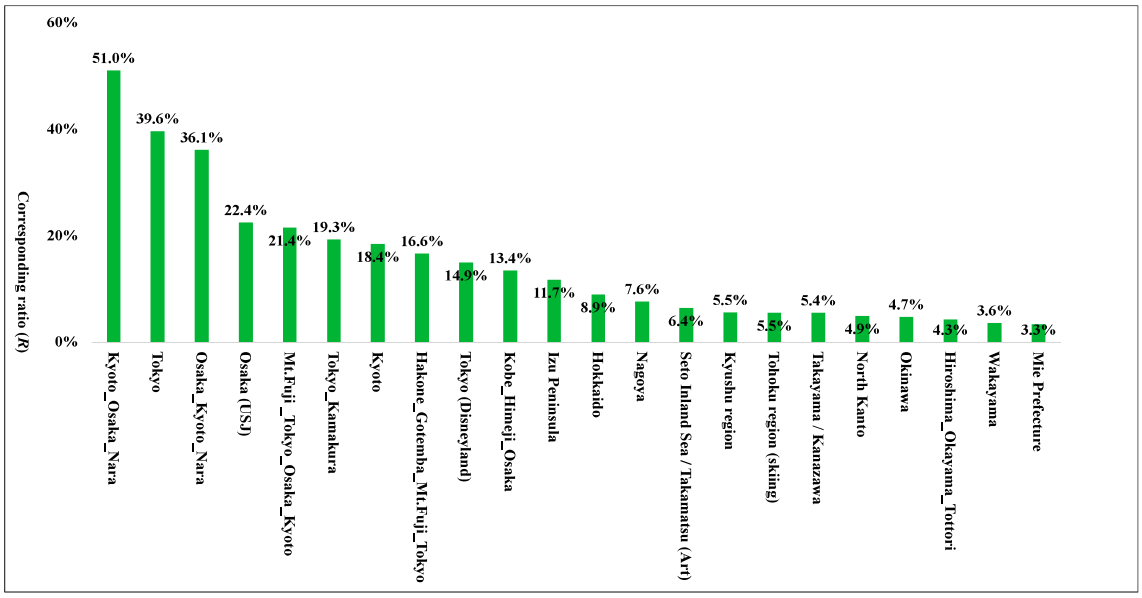


Fig. 2.3 The corresponding ratio (*R*) of each visit pattern and theme park

Figure 2.4 shows the corresponding ratio (*R*) of the tourism themes extracted by the LDA model, including some tourism themes not shown in **Table 2.2**. The result implies that the tourism themes of Chinese tourists who visited Japan are diverse. Among them, the corresponding ratios (*R*) of the gastronomy, hot-spring hotel, shrine/temple, and shopping themes are relatively higher, reaching more than 30%. In particular, almost half of the tourists mentioned the "gastronomy" topic in their travel notes. In this topic, the feature word with the highest probability is "Delicious," which shows that Chinese tourists like the Japanese gastronomy (gastronomy topics include "ramen," "sushi," "beef," "sashimi," "puffer fish," etc.) so much. Additionally, tourism themes, such as family trip, history/culture, sakura, cruise, and firework, are identified.

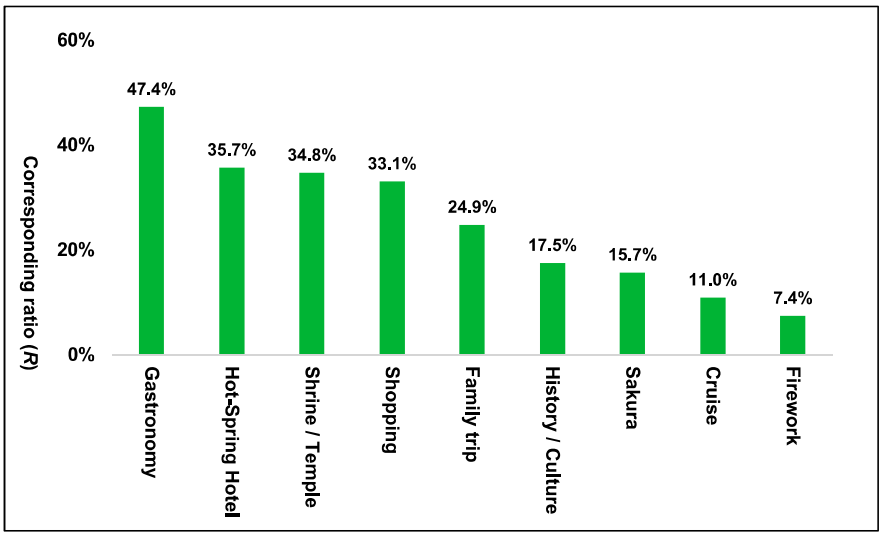


Fig. 2.4 The corresponding ratio (R) of each tourism theme

As shown in **Table 2.1**, we examined the temporal transition of the corresponding ratios of the tourism themes, theme parks, and the visit patterns between 2015 and 2019. The annual corresponding ratio R'_y of each topic (visit pattern, theme park, and tourism theme) in the specific year y (value range: 2015-2019) is defined as follows:

$$R'_y = N_y / S_y \tag{5}$$

R'_y : The annual corresponding ratio of a topic in year y.

S_y : The total number of the travel notes in year y.

N_y : The number of the travel notes corresponding to the topic in year y.

Using R'_y values as the vertical axis values, **Fig. 2.5**, **Fig. 2.6,** and **Fig. 2.7** were created to show the trend of change in each topic from 2015 to 2019. In addition, the test for the difference of the population proportions is conducted to examine the difference in R'_y values of visit patterns, theme parks, and tourism themes between 2015 and 2019. The P-values for the test are shown in **Fig. 2.5**, **Fig. 2.6,** and **Fig. 2.7** to illustrate the difference in R'_y values. If the P-value is less than 0.05, it is considered that there is a 5% significant difference in R'_y values of each visit pattern or theme park or tourism theme between 2015 and 2019.

Figure 2.5 shows the temporal transition of R'_y values of the visit patterns and the theme parks related to the metropolitan regions from 2015 to 2019. The red P-value means that the difference in R'_y values of the visit pattern between 2015 and 2019 is 5% significant (P-value is less than 0.05), while the black P-value means that the difference is not 5% significant (P-value is more than 0.05). In eight out of nine visit patterns and two theme parks, the differences in R'_y values between 2015 and 2019 are 5% significant. Except for "Osaka (USJ)," R'_y values for visit patterns that showed a 5% significant difference decreased over the five-year period. It suggests that the relative weight of the

metropolitan regions such as Tokyo, Kyoto, and Osaka as travel destination had declined between 2015 and 2019.

According to the annual changes in the visit rate of Chinese tourists visiting each prefecture [39], the visit rate of Osaka was on an upward trend. Combined with the results of the Osaka (USJ) visit pattern in **Fig. 2.5**, which was increasing annually, it can be considered that more Chinese tourists visiting Osaka were choosing USJ as the destination. Moreover, according to the annual changes in the visit rate of Chinese tourists visiting each prefecture [39], the visit rates of Tokyo, Chiba Prefecture, Kanagawa Prefecture, Yamanashi Prefecture, and Shizuoka Prefecture were decreasing. This is consistent with our analytical results of the visit patterns in the metropolitan regions in **Fig. 2.5**. Therefore, it can be inferred that Chinese tourists visiting Japan were shifting their destinations from the metropolitan regions to other regions.

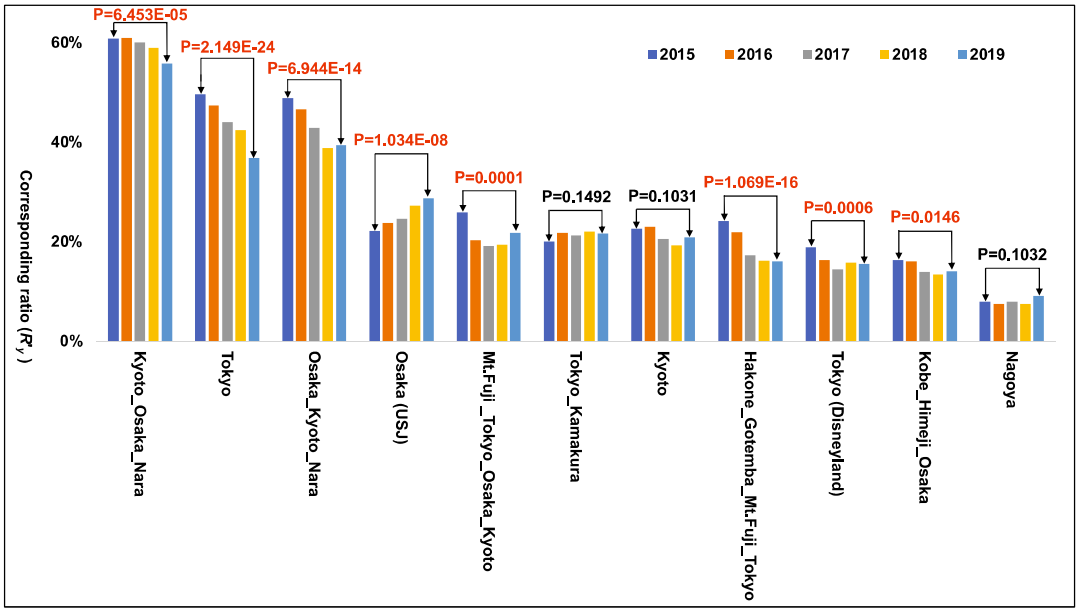


Fig. 2.5 The annual change in the visit pattern and the theme park in the metropolitan regions (The. red P-values are less than 0.05, and the black P-values are more than 0.05)

Figure 2.6 shows the R'_y values of the visit patterns on the local regions from 2015 to 2019. In 6 out of 11visit patterns, the differences in R'_y values between 2015 and 2019 are 5% significant. Except for "Hokkaido," R'_y values for visit patterns that show a 5% significant difference increase over the five-year period. The visit patterns with increasing trend are "Seto Inland Sea / Takamatsu (Art)," "Tohoku region," "Takayama / Kanazawa," "North Kanto," and "Hiroshima_Okayama_Tottori." Based on these results, it can be inferred that the increase in the R'_y values of the visit patterns to some local regions may be due to several reasons: tourists who visited metropolitan regions also visited these local regions; repeat visitors to Japan started exploring local regions from their second visit

onwards; and these local regions offer unique tourism resources that cannot be experienced in metropolitan regions.

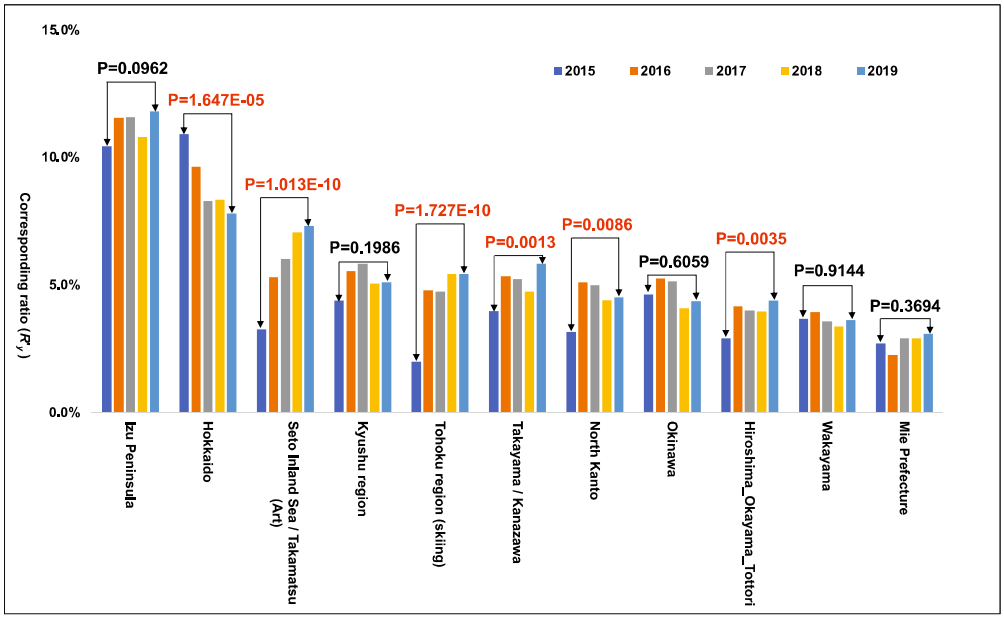


Fig. 2.6 The annual change in the visit pattern in the local regions by years

(The red P-values are less than 0.05, and the black P-values are more than 0.05)

Figure 2.7 shows the R'_y values of tourism themes from 2015 to 2019. The R'_y values of the major tourism themes such as "Gastronomy," "Hot-Spring Hotel," "Shrine/temple," and "Shopping" showed 5% significant decrease, while the R'_y values of "Family trip," "Cruise," and "Fireworks" has increased significantly over the five-year period.

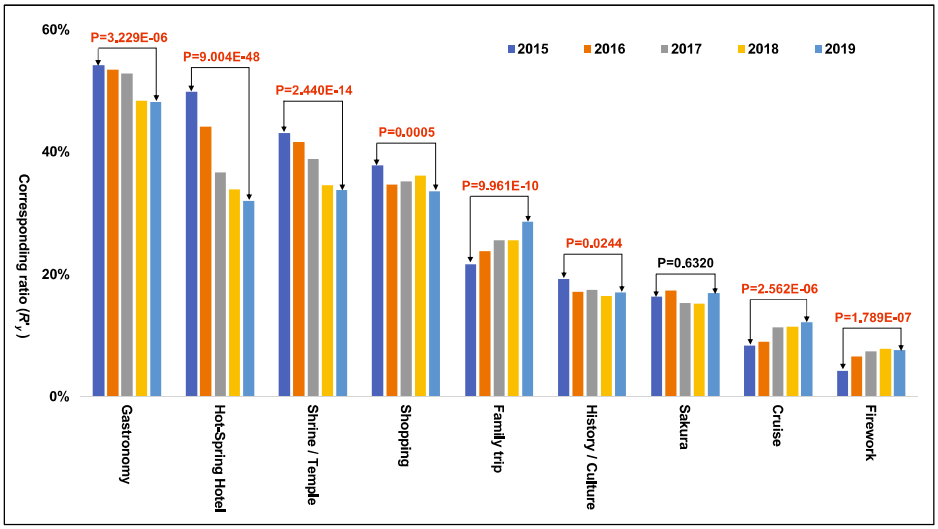


Fig. 2.7 The annually change in each tourism theme

(The red P-values are less than 0.05, and the black P-values are more than 0.05)

These results suggest that the Chinese tourists who visit Japan are changing their travel purposes from the traditional purposes such as “Gastronomy,” “Hot-Spring Hotel,” “Shrine/temple,” and “Shopping” to the diversified experiential travels such as “Family trip,” “Cruise,” and “Fireworks.” As for “Sakura” tourism theme, the P-value is more than 0.05 and it implies that the proportion of visitors who mention “Sakura (Cherry blossom)” is relatively constant over the five-year period.

Table 2.3 shows the percentage of the travel notes corresponding to a tourism theme in the travel notes corresponding to the specific visit pattern. Proportion (P_{ij}) is defined as follows:

$$P_{ij} = t_{ij}/v_i \quad (6)$$

P_{ij} : The percentage of the travel notes corresponding to the tourism theme *j* in the travel notes corresponding to the visit pattern *i*.

v_i : The number of the travel notes corresponding to the visit pattern *i*.

t_{ij} : The number of the travel notes corresponding to both the visit pattern *i* and the tourism theme *j* simultaneously.

Using P_{ij}, the importance of each tourism theme in each visit pattern can be discussed. The values in pink cells in **Table 2.3** show the top three percentages of visit patterns for each tourism theme. The proportions of travel notes corresponding to “Firework” tourism theme in the North Kanto, Tohoku region, and Mie Prefecture visit pattern were relatively high, and the top 15 feature words of these visit patterns included the venues of firework festivals. The tourism theme of cruise tourism is relatively prominent in the visit patterns of the Kyushu region and Okinawa, indicating that most Chinese tourists opting for cruise trips to Japan are inclined towards visiting the Kyushu and Okinawa regions. This result is associated with the fact that cruise routes to Japan primarily concentrate on the Kyushu and Okinawa regions. As for the tourism theme “Shopping,” the visiting patterns closely related to the tourism theme are mainly included in metropolitan regions, such as Osaka (USJ), Osaka_Kyoto_Nara, and Tokyo. The percentages of the “Hot-Spring Hotel” tourism theme in the visit patterns such as Takayama / Kanazawa, Wakayama, and Hakone_Gotemba_Mt.Fuji_Tokyo were relatively high, and the top 15 feature words of “Takayama / Kanazawa,” “Wakayama,” and “Hakone_Gotemba_ Mt.Fuji_Tokyo” visit patterns included “Gero Onsen,” “Shirahama Onsen,” and “Hakone Onsen,” which are the names of popular hot springs included in the abovementioned visit patterns. The percentages of the “Shrine/temple” tourism theme in the visit patterns such as Kobe_Himeji_Osaka, Kyoto, and Wakayama were high. These results show the possibility that a higher proportion of Chinese tourists visiting these regions chose to visit the local shrine/temple. As for the tourism theme “Family travel,” the percentages among the visit patterns such as Tokyo (Disneyland) and Okinawa were high. These results suggest that the Chinese tourists choosing the

Disneyland and Okinawa tended to travel with their families. On the other hand, the percentages in the top three visit patterns on the tourism themes “Sakura” and “Gastronomy” were similar to those of other visit patterns. P_{ij} values were similar for each visit pattern due to the possibility of enjoying cherry blossoms and gourmet food throughout Japan.

Table 2.3 The proportion of travel note corresponding to each tourism theme in the travel notes corresponding to each visit pattern. (1)

Topic ID		5	7	13	25	31	35	38	39	40
	Themes / Visit Patterns	Sakura	Fire-work	Gruise	Shop-ping	Hot-Spring Hotel	Shrine Temple	Gastro-nomy	History Culture	Family trip
1	Kobe_Himeji_Osaka	20.6%	10.7%	10.3%	30.5%	46.8%	52.0%	56.8%	25.1%	24.3%
2	Mt.Fuji _Tokyo_ Osaka_Kyoto	18.8%	6.9%	10.1%	38.0%	37.3%	38.7%	45.8%	25.2%	26.2%
3	Takayama / Kanazawa	23.1%	14.2%	12.1%	25.0%	55.9%	40.4%	51.3%	30.8%	25.7%
6	Osaka (USJ)	13.4%	8.3%	7.5%	41.0%	34.5%	40.3%	55.9%	13.8%	27.5%
10	Nagoya	18.9%	10.2%	11.6%	34.7%	44.6%	39.1%	50.5%	26.9%	29.9%
11	North Kanto	29.4%	20.7%	14.2%	24.0%	50.6%	38.6%	45.0%	30.1%	24.4%
12	Kyoto	24.9%	8.5%	6.3%	32.6%	38.9%	62.5%	56.1%	19.8%	23.0%
14	Hiroshima_Okayama_Tottori	24.1%	16.2%	19.0%	21.9%	42.3%	41.6%	46.9%	37.3%	24.9%
17	Wakayama	24.4%	16.3%	14.6%	30.7%	62.7%	54.8%	53.1%	36.4%	30.8%
18	Tokyo_Kamakura	19.1%	12.3%	8.3%	34.2%	37.7%	35.3%	52.7%	18.1%	22.7%
19	Kyoto_Osaka_Nara	17.8%	7.1%	5.7%	37.4%	36.7%	47.7%	52.9%	18.6%	24.0%
22	Tokyo (Disneyland)	14.2%	10.2%	12.9%	38.0%	38.6%	26.8%	48.5%	14.9%	40.7%
23	Tohoku region (skiing)	26.4%	17.6%	15.7%	23.8%	53.8%	29.0%	47.3%	25.5%	28.1%
24	Mie Prefecture	23.5%	20.6%	20.8%	29.1%	41.4%	44.4%	51.4%	38.0%	28.8%
26	Hakone_Gotemba_ Mt.Fuji_Tokyo	20.9%	8.4%	7.9%	35.6%	54.3%	35.2%	51.7%	17.4%	25.4%
28	Izu Peninsula	23.0%	14.7%	16.7%	23.5%	48.5%	36.2%	45.5%	26.2%	25.7%
29	Kyushu region	20.2%	14.4%	34.4%	27.9%	42.7%	31.1%	48.8%	27.2%	30.1%
30	Hokkaido	17.6%	12.0%	12.9%	29.7%	51.2%	25.1%	50.1%	14.6%	27.8%

Table 2.3 The proportion of travel note corresponding to each tourism theme in the travel notes corresponding to each visit pattern. (2)

Topic ID		5	7	13	25	31	35	38	39	40
	Themes / Visit Patterns	Sakura	Fire-work	Gruise	Shop-ping	Hot-Spring Hotel	Shrine Temple	Gastro-nomy	History Culture	Family trip
32	Osaka_Kyoto_Nara	18.3%	6.7%	6.6%	39.8%	36.5%	51.2%	58.2%	19.0%	25.0%
34	Tokyo	17.6%	8.2%	7.2%	39.1%	38.8%	35.0%	53.7%	17.5%	23.4%
36	Okinawa	18.2%	16.7%	33.4%	32.9%	36.4%	27.5%	47.4%	21.6%	42.2%
37	Seto Inland Sea / Takamatsu (Art)	19.3%	14.0%	16.3%	28.2%	43.3%	40.5%	49.4%	26.6%	26.6%

Note: The values in pink show the top three percentages of visit patterns for each tourism theme. Topics 4, 8, 9, 15, 16, 20, 21, 27, and 33 are not included in **Table 2.3** because they could not be classified under any of the categories.

2.4.2 The results of the Word2vec model

The LDA model focuses on the co-occurrence of words in documents, and it does not make any assumption about the structure of text or the syntactical or grammatical properties of the language. Consequently, the context within the article is not considered well in the LDA model. On the other hand, the Word2vec model focuses on the co-occurrence relationships among neighboring words in a text. The word vector obtained by the Word2vec model represents more features that incorporate co-occurrence of words considering context. Therefore, the Word2vec model can be used in text mining in a complementary manner to the LDA model.

In this section, the Word2vec model is used to understand the behaviors of Chinese tourists in more detail, especially to explore tourists' activities in the local regions, where the numbers of Chinese tourists are limited.

In **Fig. 2.8**, topic 14 (Hiroshima_Okayama_Tottori) is used as an example of the application of the Word2vec analysis. Hiroshima, Okayama, and Tottori are the names of prefectures in a local region, and topic 14 is one of the visit patterns in the local region. As shown in **Fig. 2.8**, we expanded the number of feature words of topic 14 obtained from the LDA model from 15 in **Table 2.2** to 30. Among them, there are six feature words related to Hiroshima prefecture, four feature words related to Okayama prefecture, and five feature words related to Tottori prefecture.

Then, the words "Hiroshima," "Okayama," and "Tottori" were input into the Word2vec model for analysis. The lower half of **Fig. 2.8** shows the 15 feature words with high association with "Hiroshima," "Okayama," and "Tottori," obtained by the Word2vec model. The red-letter words appear in the results of both the LDA and Word2Vec models. Although the feature words of the visit patterns obtained by

the LDA model include the names of large cities (Hiroshima city, Okayama city, etc.) and major tourism spots (Miyajima, Korakuen, Sand dunes, etc.), the Word2vec model can extract more detailed information, such as small municipalities (Takehara city, Kurayoshi city, Misasa town, etc.), tourism spots (peace memorial park, shukkeien garden, etc.), and local tourism resources (Sanfrecce [local football team,] Momotaro [hero in Japanese folklore,] etc.) In addition, as shown in Figure 2.8, both the topic 14 extracted from the LDA model and the Word2vec analysis of this topic revealed the word "Conan." Upon examining all the results from the LDA model and the Word2vec model related to local regions, we found that the only anime-related term identified was "Conan." Therefore, we speculate that visiting the town related to Conan in Tottori Prefecture is a unique preference among Chinese tourists traveling to Japan.

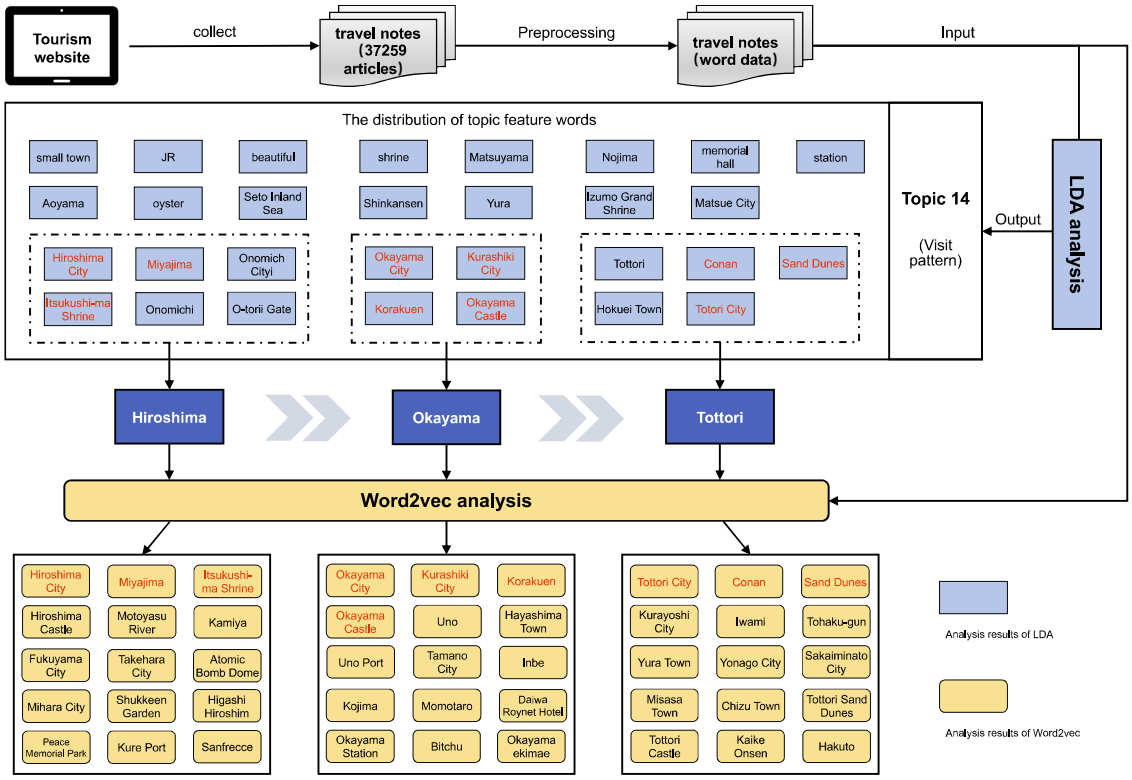


Fig. 2.8 The Word2vec model is used to complement the results of topic14
(The red part indicates that the two models output the same results.)

Table 2.4 summarizes the results of the same analysis shown in **Fig. 2.8** for the visit patterns in other local regions. The red words appear in the results of both the LDA and Word2vec models. It can be found that when both models output the same number of analysis results (30 words), the Word2vec model outputted more information related to the input place name (visited places, attractions, and local tourism resources). Additionally, the cosine similarity of the results produced by the Word2vec model

and the input words is primarily distributed between 0.4 and 0.6, indicating that there is a certain degree of semantic similarity between them. On the other hand, when using only the LDA model to obtain detailed travel information, it is necessary to increase the number of feature words within each topic. However, in this case, it is expected that many noisy words with low relevance to the visit pattern will also be extracted. Therefore, it is expected that useful information can be obtained more efficiently by using the LDA model and the Word2vec model in a complementary manner, as proposed in this study.

2.4.3 Discussion on the policy to promote inbound tourism in local regions

In this section, based on the analytical results from LDA and Word2vec models, the policy to promote inbound tourism in local regions in Japan is discussed.

The local regions far from metropolitan regions and major tourism sites are considered to be at a relative disadvantage in terms of inbound tourism. However, as shown in **Fig. 2.3**, the Chinese tourists mentioned not only the topics on the metropolitan regions, but also those on the local regions such as Seto Inland Sea /Takamatsu and Hiroshima_Okayama_Tottori. **Figure 2.6** also shows that the percentage of mentions of the visit patterns in many local regions is increasing. These results suggest that other local regions also have the opportunity to promote inbound tourism from China.

From **Table 2.2**, Seto Inland Sea /Takamatsu visit pattern included “art gallery,” and Hiroshima_Okayama_Tottori included “Conan,” together with the major tourism sites in these regions. These are unique local resources that cannot be experienced elsewhere. Leveraging these unique tourism resources will help in developing effective promotional strategies for tourism.

Table 2.4 shows that Chinese tourists mentioned small municipalities and local foods, and festivals, etc. Additionally, we can focus not only on the mentioned words, but also on the unmentioned words. In other words, we can determine which travel information within specific region has been extracted and which has not, allowing us to predict the preferences of Chinese tourists traveling to the local regions in Japan. The information on the preferences of Chinese tourists can be useful to propose an effective tourism promotion. In the local regions where the number of foreign tourists is relatively small, it is difficult to discuss the tourism promotion policy for the visitors from the specific country quantitatively. The analytical methodology used in this study can be helpful especially for such local regions to develop an effective tourism policy.

After the COVID-19 pandemic, Chinese tourists are expected to be more inclined to travel independently. It is important for the recovery of inbound tourism for each region to develop highly unique tourism promotion measures, based on the analysis of tourists’ preferences, such as this study.

Table 2.4 The comparison of LDA model and word2vec model analysis results (1)

Topic ID	LDA model results		The words as input to Word2vec model	Word2vec model results			
	Visit place	Others		Visit place	Attractions	Regional resources	Others
23 Tohoku Region	Tokyo, Kakunodate, Hirosaki Park, Oirase, Zao	ski, ski facility*2, jr, Shinkansen, Station, apple, Hot spring, Dingding, cable car	Sendai	Miyagi Prefecture, Yamagata, Aomori, Sendai City, Utsunomiya, Ishinomaki, Ishinomaki City, Aoba, Aobaku, Tohoku University, Ogawara, Morioka, Koriyama, Niigata	Matsushima, Sendai Castle, Akiho, Akiho Onsen, Zuihoden Temple, Naruko Onsen	Beef tongue	Sendai Station, looplebus, Tohoku Region, Date,Suisui, Rikyu, Kisuke, Inosuke, Uzen
			Aomori	Aomori Prefeture*2, Iwate, Hirosaki, Hirosaki City, Aomori City, Hachinohe City, Goshogawara, Goshogawara City, Misawa, Towada, Towada City, Morioka	KAI Tsugaru, Nebuta House, Lake Towada, Mt.Hakkodasan, Hoshino Resort, Apple Orchard, Asamushi	Nebuta*2 (festival)	Shirakami, Hakkoda, Hoshino, Yasumiya, Ukiyu, Gono Line, Tohoku Region, Stream
23 Tohoku Region			Akita	Akita Prefecture*2, Yamagata, Aomori Prefecture, Odate, Odate City, Akita City, Ugo, Noshiro, Oga, Kakunodate, Kakunodate Town, Semboku City, Yokote, Yokote City, Omagari, Takanosu, Hirosaki, Shinjo, Goshogawara	Senshu Park, Lake Tazawa, Nyutou (Onsen)	Shiba Inu, puppy, dog	Want to raise, Gono Line, Inaniwa (Udon Shop), land of beauty

Table 2.4 The comparison of LDA model and word2vec model analysis results (2)

Topic ID	LDA model results		The words as input to Word2vec model	Word2vec model results			
	Visit place	Others		Visit place	Attractions	Regional resources	Others
29 Kyushu Region	Tenjin	Jigoku, jr, small town, train, shrine, Tenmangu, volcano, bus	Fukuoka	Fukuoka*2, Kitakyushu, Saga, Nagasaki, Kagoshima, Hakata, Hakata Ward, Fukuoka City, Nakasu*2, Dazaifu City, Kurume City, Kitakyushu City, Kokura, Saga City, Tosu, Sasebo, Beppu, Mojiko	Dazaifu, Ohori, Ohori Park, Hakata Bay	Beef Intestine Pot	kyushu, Reflector, not climbing tower, zaifu, maizuru
			Nagasaki	Nagasaki Prefecture, Fukuoka, Sasebo, Sasebo City, Nagasaki City, Isahaya, Nishihama, Urakami, Unzen, Omura City	Inasayama, Glover Garden*2, Oura Church, Oranda-zaka, Huis Ten Bosch, Shiambashi, Peace Park, Atomic Bomb Museum, Kujukushima, Fukusaya	Kakuni (food)	Oura, Church*2, Dejima, Return Ship, Pull in to Shore, Nether-lands, go ashore
			Kumamoto	Kumamoto Prefecture*2, Kagoshima, Kitakyushu, Nagasaki, Kumamoto City, Shinshigai, Shimotori, Aso, Takamori, Tosu, Kurume	Bear Shop, Bear Pavilion, Kumamoto Castle, Johsaien, Jojuen Karashima, Takefue, Honmyoji, Kurokawa, Aso Volcano, Suizenji Jojuen,	Horseme at sashimi	Minister, Kyushu, Suganoya, Lovely Bear, Kokutei, Tsuruya,

Table 2.4 The comparison of LDA model and word2vec model analysis results (3)

Topic ID	LDA model results		The words as input to Word2vec model	Word2vec model results			
	Visit place	Others		Visit place	Attractions	Regional resources	Others
29 Kyushu Region	Tenjin	Jigoku, jr, small town, train, shrine, Tenmangu, volcano, bus	Oita	Oita Prefecture, Saga, Saga Prefecture, Miyazaki Prefecture, Kumamoto*2, Nagasaki Prefecture, Fukuoka Prefecture, Fukuoka, Yufu, Yufu City, Yufuin, Beppu, Hita City, Bungo, Yufuin Town, Hitoyoshi, Kurume, Chikuzen, Kitakyushu City, Unzen,Kirishima	Beppu Onsen, Yabakei*2	Beppu Hatto (hot spring)	Oita Station, Hisatsu Sen, Kyudai Sen*2
			Kagoshima	Kagoshima, Sasebo, Kumamoto, Saga, Miyazaki, Fukuoka, Nagasaki, Takamori, Kagoshima City, Kinsei, Yatai village, Ibusuki, Ibusuki City, Kirishima City, Yoshimatsu, Kumamoto City	Yakushima, Senganen Garden*2, Kirishima, Sakurajima, Shiroyama, Kinkowan, Planetarium	Black pig, Black Pork, Hairy pig	Return ship, amuplaza, Kyushu

Table 2.4 The comparison of LDA model and word2vec model analysis results (4)

Topic ID	LDA model results		The words as input to Word2vec model	Word2vec model results			
	Visit place	Others		Visit place	Attractions	Regional resources	Others
37 Seto Inland Sea / Takamatsu	Teshima, Kotohira Shrine, Shikoku, marina, pumpkin (Art work)	Museum, work,olive, Building, bus,port, Small island, wharf,Tadao Ando, angel, soy sauce, bicycle, Artist, plan, Photography	Seto Inland Sea	Seto, Shikoku*2, Kagawa, Takamatsu, Shimonada, Imabari, Takehara City, Onomichi, Tamano City	Shodoshima, Seto Ohashi, Inujima, Naoshima, Innoshima, Ikuchijima Island	Arts Festival	Island Island, hopping*3, Uchigashima, Each Island*2, Sea Wave, Island, Utazu Remote Kirousan, Hayama, doshima,
37 Seto Inland Sea / Takamatsu	Teshima, Kotohira Shrine, Shikoku, marina, pumpkin (Art work)	Museum, work, olive, Building, bus, port, small island, wharf, Tadao Ando, angel, soy sauce, bicycle, Artist, plan, Photography	Takamatsu	Kagawa, Okayama, Tokushima, Seto Inland Sea, Takamatsu City, Marukame, Kotoden, Kotohira, Tonosho, Kawaramachi Station, Tamano, Katahara, Uno, Tamano City, Uno Port	Shodoshima, Ritsurin, Tamamo Park, Shimayama	Honetsukidori Ikkaku (Udon Restaurant)	Takamatsu Station, JR Hotel clement takamatsu, Ikkaku, Island hopping*2, Utazu, Torimi, catch boat, each Island, paicei

Note: Input words are the words that have been inputted into the Word2vec model. The red part is the duplicate content of the analysis results of the LDA model and the Word2vec model. The contents in parentheses are explanatory notes. The number after the asterisk in **Table 4** indicates that the word is extracted twice as Chinese characters and Japanese characters in the LDA or Word2vec model results, or multiple Chinese words may appear to refer to the same English word pair, so the same English word may appear more than once in the feature words. The number in front of the asterisk represents the cosine distance between the Word2vec model's output word and the input word.

2.5 Conclusion

In chapter 2 used the LDA model and the Word2vec model, which are natural language processing techniques to analyze the web travel notes written by Chinese tourists who visited Japan that are posted on the Chinese visitor website Mafengwo. Based on the results, this study analyzed the tourism themes and visit patterns and the changes in the choices of these Chinese tourists. Furthermore, we clarified the tourism resources specific to some regions. Based on the analytical results, the policy to promote inbound tourism in local regions in Japan was also discussed.

The results of the LDA model analysis showed that the Chinese tourists who visited Japan were changing their purposes from the traditional purposes such as “Gastronomy,” “Hot-Spring Hotel,” “Shrine/temple,” and “Shopping” to the diversified experiential travels such as “Family trip,” “Cruise,” and “Fireworks.” In terms of the visit patterns of Chinese tourists to Japan, more Chinese tourists visiting Osaka were choosing to visit the USJ. According to the annual changes in the visit rate of Chinese tourists visiting each prefecture [39], the visit rates of Tokyo, Chiba Prefecture, Kanagawa Prefecture, Yamanashi Prefecture, and Shizuoka Prefecture, which are included in the metropolitan regions or neighboring regions, are decreasing yearly. Additionally, it is clear that the visit rate of Chinese tourists to Hokkaido is decreasing, while the visit rate to Okayama and Kagawa prefectures is increasing. These findings are consistent with the results obtained from the LDA model, as shown in Figs. 5 and 6. Therefore, it can be inferred that Chinese tourists visiting Japan are shifting from the metropolitan regions to other local regions.

“International Visitor Survey” data collected by the Japan Tourism Agency are aggregated by prefecture. As a result, it is difficult to analyze the detailed tourism behaviors in specific prefectures. On the other hand, the LDA model specifies the topics directly from the travel notes, which include names of the places that tourists visited actually. Consequently, as shown in **Table 2.2**, the topics categorized as visit pattern include the names of cities and tourism spots. Such information is useful to estimate the excursion pattern within prefectures. Additionally, this study also utilizes a Word2Vec model to analyze word vectors related to specific place names, with the aim of extracting travel information pertaining to the specific regions. This includes visited locations, attractions, tourism resources, and shop names. Based on the extracted geographic information, we can develop the tour routes for the local region and offer recommendations to the relevant departments regarding sightseeing transportation routes. Furthermore, from the extracted travel information, we can also analyze the tourist attractions of particular interest to Chinese tourists, such as famous shrines, Sakura spots, and natural scenic areas. This analysis will help us to formulate corresponding development and preservation policies to ensure that these resources receive proper maintenance and protection, thereby promoting the sustainable development of the local tourism industry. Furthermore, the Word2vec model can also help us discover tourism resources in different regions, as shown in **Table 2.4**, such as

Sendai's gastronomic beef tongue, Aomori's Nebuta Festival, Okayama's Momotaro, Tottori's Conan, Seto Inland Sea/Takamatsu's art festival, Kumamoto's basashi (horse meat sashimi), and so on. With this information, we can not only provide more precise travel recommendations for visitors but also predict the preferences of Chinese tourists in various regions.

In summary, using the LDA model and the Word2vec model in a complementary manner proposed in this study, it is possible to understand Chinese tourists' behaviors and tourism resources in detail in the regions that they visit. Based on these results, it is possible to discuss the tourism promotion policy in local regions, including discovering unique tourism resources. Furthermore, the local tourism resources of destination could be strategically developed to create new tourism content for foreign tourists who visit Japan.

References

1) White Paper on Tourism (2015-2020), Japan Tourism Agency. Available at: https://www.mlit.go.jp/kankocho/en/siryou/whitepaper.html
2) International Visitor Survey (2012-2020), Japan Tourism Agency. Available at: https://www.mlit.go.jp/kankocho/en/siryou/toukei/syouhityousa.html
3) Hao, Q., Cai, R., Wang, C., Xiao, R., Yang, J. M., Pang, Y., Zhang, L.: Equip tourists with knowledge mined from travelogues. Proceedings of the 19th international conference on World wide web, pp. 401–410, 2010
4) Kozinets, R.: The field behind the screen: using netnography for marketing research in online communities. Journal of Marketing Research, Vol. 39(1), pp. 61-72, 2002
5) Reid, E.: Informed consent in the study of on-line communities: a reflection on the effects of computer-mediated social research. The Information Society, Vol. 12(2), pp. 169-174, 1996
6) Li, Q., Li, S., Zhang, S., Hu, J., Hu, J.: A review of text corpus-based tourism big data mining. Applied Sciences, Vol. 9(16), 3300, 2019
7) Hao, Q., Cai, R., Wang, C., Xiao, R., Yang, J. M., Pang, Y., Zhang, L.: Generating location overviews with images and tags by mining user-generated travelogues. Proceedings of the 17th ACM international conference on Multimedia, pp. 801–804, 2009
8) Vu, H. Q., Li, G., Law, R.: Discovering implicit activity preferences in travel itineraries by topic modeling. Tourism Management, Vol. 75, pp. 435-446, 2019
9) Abeysinghe, S., Manchanayake, I., Samarajeewa, C., Rathnayaka, P., Walpola, M. J., Nawaratne, R., Bandaragoda, T., Alahakoon, D.: Enhancing decision making capacity in tourism domain using social media analytics. International Conference on Advances in ICT for Emerging Regions, 2018
10) Haris, E., Gan, K.H., Tan, T.-P.: Spatial information extraction from travel narratives: Analysing the notion of co-occurrence indicating closeness of tourist places. Journal of Information Science, Vol. 46(5), pp. 581–599, 2020
11) Yuan, H., Xu, H., Qian, Y., Li, Y.: Make your travel smarter: summarizing urban tourism information from massive blog data. International Journal of Information Management, 36(6), pp. 1306-1319, 2016
12) Hayashi, T.: Applying the document vector model to tour recommendation. Journal of Global Tourism Research, Vol. 4, No. 1, pp. 15-20, 2019
13) Shafqat, W., Byun, Y.-C.: A recommendation mechanism for under-emphasized tourist spots using topic modeling and sentiment analysis. Sustainability, Vol. 12(1), 320, 2020
14) Du, S., Zhang, H., Xu, H., Yang, J., Tu, O.: To make the travel healthier: a new tourism personalized route recommendation algorithm. Journal of Ambient Intelligence and Humanized Computing 10, pp. 3551–3562, 2019
15) Chen X., Li J, Han W, Liu S.: Urban tourism destination image perception based on LDA integrating social network and emotion analysis: the example of Wuhan. Sustainability, Vol. 14(1), 12, 2021

16) He, W., Tian, X., Tao, R., Zhang, W., Yan, G. and Akula, V.: Application of social media analytics: a case of analyzing online hotel reviews. Online Information Review, 41, pp. 921-935, 2017
17) Ren, G.; Hong, T.: Investigating online destination images using a topic-based sentiment analysis approach. Sustainability, Vol. 9(10), 1765, 2017
18) Zheng, X., Luo, Y., Sun, L., Zhang, J., Chen, F.: A tourism destination recommender system using users' sentiment and temporal dynamics. Journal of Intelligent Information Systems, Vol. 51(3), pp. 557-578, 2018
19) Nakamura, Y., Oomiya, N.: Applying the topic model to hotel reviews of internet sites and analyzing their results. Journal of Global Tourism Research, Vol. 6, No. 1, pp. 15-20, 2021
20) Guo, Y., Barnes, S.J., Jia, Q.: Mining meaning from online ratings and reviews: Tourist satisfaction analysis using latent dirichlet allocation. Tourism Management, Vol. 59, pp. 467–483, 2017
21) Hu, N., Zhang, T., Gao, B., Bose, I.: What do hotel customers complain about? Text analysis using structural topic model. Tourism Management, Vol. 72, pp. 417-426, 2019
22) Zhai, Y., Chen, P.: Sentiment analysis on tourist satisfaction with rural homestay inns based on reviews from the website of online travel agency. International Journal of Sustainable Development and Planning, Vol. 15, No. 5, pp. 705-712, 2020
23) Quiaoit, M., & Furuya, H.: Analysis of Travel Reviews on the Philippine Tourism Experience by WORD2VEC Model, Journal of Japan Society of Civil Engineers. Ser. D3 (Infrastructure Planning and Management), Vol. 77, No. 5, Special Issue, I_253-I_268, 2022
24) Hishida, N., Hibino, N., Morichi, S.: Chinese Tourists in Japan: Time-Series Analysis in Their Destination Choices and its Multiplicities. Journal of Japan Society of Civil Engineers D3, Vol. 68(5), pp. I_667-I_677, 2012
25) Matsui, Y., Hibino, N., Morichi, S., Ieda, H.: Behavior analysis on inbound tourists to japan focus on their destinations and activities based on individual data. Journal of Japan Society of Civil Engineers D3, Vol. 72(5), pp. I_533-I_546, 2016
26) Furuya, H., Liu, Y.: Analysis of combination of visiting places by foreign visitors in japan by Latent Class Model. Journal of Japan Society of Civil Engineers D3, Vol. 72(5), pp. I_571-I_583, 2016
27) Furuya, H.: Study on identification of destinations combination pattern of foreign visitors to japan based on similarity by hPAM. Journal of Japan Society of Civil Engineers D3, Vol. 75(5), pp. I_507-I_517, 2019
28) Tatsumi, Y., Tsukai, M.: A tour of analysis of foreign visitors to Japan by applying topic model. Transport Policy Studies' Review, Vol. 23, pp. 20-34, 2021. Available at: https://www.jstage.jst.go.jp/article/tpsr/advpub/0/advpub_TPSR_23R_04/_pdf/-char/ja
29) Song, Z., Furuya, H.: Analysis of Chinese tourists' travel literatures by Latent Dirichlet Allocation Model. Proceedings of Research Meeting on Civil Engineering Planning, 2018. Available at: http://lbrary.jsce.or.jp/jsce/open/00039/201806_no57/57-01-01.pdf

30) Guizhou China Homepage, 2018. Available at: https://www.guizhou.gov.cn/home/gzyw/202109/t20210913_70110352.html
31) Blei, D.M., Ng, A.Y., Jordan, M.I.: Latent Dirichlet Allocation. Journal of Machine Learning Research, 3(Jan), pp. 993-1022, 2003
32) Petterson, J., Buntine, W., Narayanamurthy, S., Caetano, T., Smola, A.: Word features for Latent Dirichlet Allocation. Advances in Neural Information Processing Systems 23, 2010
33) Mikolov, T., Chen, K., Corrado, G., Dean, Jeffrey.: Efficient estimation of word representations in vector space, 2013. Available at: https://arxiv.org/abs/1301.3781
34) Selva, Prabhakaran, S.: Topic Modeling with Gensim (Python), 2018. Available at: https://www.machinelearningplus.com/nlp/topic-modeling-gensim-python/#17howtofindtheoptimalnumberoftopicsforlda
35) Latent Dirichlet Allocation (n.d.), Gensim: Topic modelling for humans. Available at: https://radimrehurek.com/gensim/models/ldamodel.html
36) Hu, Z., Ma, X., Liu, Z., Hovy, E., Xing, E.: Harnessing Deep Neural Networks with Logic Rules, 2016. Available at: https://arxiv.org/abs/1603.06318
37) Jatnika, D., Bijaksana, M.A., Suryani, AA.: Word2vec Model Analysis for Semantic Similarities in English Words. Procedia Computer Science, Vol. 157, pp. 160-167, 2019
38) Word2vec embeddings (n.d.), Gensim: Topic modelling for humans. Available at: https://radimrehurek.com/gensim/models/word2vec.html
39) Visit Rate Ranking by Prefecture, Japan National Tourism Organization, 2015-2019. Available at: https://statistics.jnto.go.jp/en/graph/#graph--inbound--prefecture--ranking

CHAPTER 3
The differential impacts of the covid-19 pandemic on outbound tourists from different groups

3.1 Introduction

Since the outbreak of COVID-19 in early 2020, the global tourism industry has suffered a huge impact [1]. As the pandemic has spread globally, almost all countries and regions have implemented border closures and restrictive measures [2, 3], and industries such as aviation, hotels, restaurants, and tourism operations have been devastated [4, 5]. According to the data from the UNWTO, the number of international tourists decreased by 72%, and international tourism export earnings declined by 63% in 2020 compared to figures from the previous year, equivalent to a loss of $1.1 trillion in annual revenue [6]. As one of the world's important economic pillar industries, the tourism industry's impact is not only short-term direct economic losses but may also have long-term effects on the psychology, demand and tourism preferences of tourism consumers. This may subsequently affect their behavior and decision-making when it comes to participating in outdoor activities.

The spread of the coronavirus has brought safety and health considerations to tourists. Tourists may pay more attention to the hygiene and safety conditions of travel destinations and tourism service providers [7 - 11, 19]. This has led some tourism consumers to be more willing to choose destinations with smaller COVID-19 impact or that can provide space from others, such as natural scenic areas, outdoor activities, and other low-contact tourism destinations (rural areas, etc.), instead of densely populated urban areas with high human traffic [12 - 20]. Secondly, because there is a higher risk of contracting the virus in densely crowded environments, some tourism consumers are likely to avoid large tour groups when choosing their travel methods and instead prefer small groups or individual travel [21, 22]. Additionally, owing to the pandemic situation at travel destinations and the uncertainty brought about by travel restrictions, some tourism consumers prefer to choose destinations closer to home [9, 23 - 24].

On the other hand, the behavior and decision-making of tourism consumers may vary due to differences in demographic characteristics. For example, research by Roman [25] suggests that the majority of high-income tourism consumers choose to spend their vacations in their own country, while people with higher levels of education choose to travel abroad. In addition, owing to differences in cultural backgrounds, economic levels, tourism development stages, as well as variations in the infection situation and prevention and control measures of the COVID-19 pandemic in different countries and regions, the demands, preferences, and consumption habits of the tourism market may also differ [26, 27]. This may lead to differences in the behavior and decision-making of tourist consumers from different countries. For example, Łapko's [28] research has shown that country-

specific travel restrictions impact sailing tourism and cruise decisions. Tourism can be considered as the concerted behavior between tourists and tourism industry. The tourism satisfaction arises as an interaction between the hospitality of the tourism industry and the cognitive state of tourists. Therefore, when formulating tourism promotion policies, related decision-makers in the tourism industry need to comprehensively consider the tourism needs and preferences of different tourism consumer groups, especially when formulating inbound tourism policies, they also need to understand the behavioral and decision-making differences of tourism consumer groups from different countries, as well as the constraints behind them.

This study conducted online surveys in August and September in 2022 among respondents in China and Japan who had previous experience of outbound travel. The survey specifically investigated the respondents' perceptions and attitudes towards the COVID-19 pandemic at different times, their perception of the risks associated with outbound tourism, and changes in their travel preferences compared to before the pandemic. Descriptive statistics (such as mean and standard deviation et al.), ordinal logistic regression and t-tests were used to analyze and compare the survey data of the respondents from both countries. The purpose of this study is to understand the constraining factors (risk perceptions) and preference changes for different groups in choosing outbound tourism during the COVID-19 pandemic, as well as the differences in the impact among different groups. Through this research, we hope to provide valuable information for related decision-makers in the tourism industry, in order to better meet the needs of tourism consumers and accelerate the recovery of the tourism industry.

3.2 Literature review

The tourism industry has always been an important part of global economic growth and national economic development [29]. However, the outbreak of COVID-19 in early 2020 has restricted or completely suspended tourism activities worldwide globally [30]. To understand the impact of the COVID-19 pandemic on potential tourism consumers, scholars have conducted extensive research around tourism behavior, tourism market, and tourism policies, aiming to understand the attitude and behavior changes of tourism consumers and seek new strategies for tourism development.

Some studies indicate that the impact of the COVID-19 pandemic on potential tourist groups varies depending on their demographic characteristics. For example, research conducted by Zhan [31] indicates that the level of risk awareness of the adult population towards the COVID-19 pandemic is higher than that of teenagers. Jin's [32] research findings indicate that the likelihood of low-income groups engaging in leisure travel is lower than that of high-income groups, and the group with the highest income shows a significant preference for overseas destinations compared to all other groups.

Furthermore, Perić's [33] research also indicates that individuals with higher monthly incomes are more likely to choose opportunities to travel abroad.

On the other hand, some scholars have investigated the impact of the COVID-19 pandemic on changes in the travel intentions, restricting factors, and preferences of tourism consumers in certain countries. Liu [34] found that the willingness of Chinese residents to travel was generally low in June and July 2020. The perception of COVID-19 had a negative impact on Chinese residents' outbound travel, while risk tolerance had a negative moderating effect on the association between perception of COVID-19 after the pandemic and travel intentions. Wen [21] summarized the constantly changing consumption patterns of Chinese tourists based on news related to "COVID-19" and "tourism" reported by multiple media outlets in 2020, supplemented by an overview of literature on tourism marketing, management, and behavior. The authors concluded that the consumption patterns of Chinese tourists will include travel choices based on nature and a trend towards "slow travel," which emphasizes longer stays and more fulfilling travel experiences. However, a survey conducted on Costa Rican residents from March 15 to April 15, 2020, showed that, in addition to concerns about healthcare issues related to COVID-19 when deciding to visit a travel destination, potential domestic tourism consumers were more inclined to travel with family or friends for short trips of 2-3 days [35]. Furthermore, the studies of Santos [36] and Marques [17] indicate that, after the lifting of domestic mobility restrictions in Portugal, tourists have shown a significantly increased preference for tourism in low-density areas (rural areas), and regions that are best suited to provide rural accommodation have experienced a stronger and more sustained recovery in domestic demand. However, Orden-Mejía [37] conducted a survey on the preferences, attitudes, and expectations of Guayaquil residents in Ecuador when planning travel after the pandemic (April 2020 to May 2020). The results indicate that, after the outbreak, residents of the area showed a greater preference for urban tourism, followed by cultural tourism and family tourism. In addition, Park [38] recruited 197 adult participants from the United States to investigate how the COVID-19 pandemic affected their tourists' preferences for crowded and non-crowded options. The results indicated that, when the pandemic is prominent, tourists tend to reduce their preferences for crowded options and choose non-crowded ones instead. However, for participants with high sensory-seeking and uniqueness needs, the tendency to avoid crowded options is significantly reduced. NAKAMURA [39] conducted a survey on the intention of Japanese people to travel abroad in February 2021, which showed that before the COVID-19 pandemic stabilizes, most Japanese people try to avoid traveling abroad. Even for those who choose to travel abroad, they tend to choose countries/regions that are closer to Japan or English-speaking countries/regions. Additionally, they expect to have travel experiences that are close to normal with fewer restrictions, and they are less willing to accept various restrictions and conditions when traveling overseas.

Based on the aforementioned research, we have found that the COVID-19 pandemic has varying impacts on the travel preferences and attitudes of tourists in different countries and regions. Moreover, people's confidence may gradually be restored due to the increased availability of vaccines and implementation of stricter prevention and control measures, which could lead to changes in their attitudes towards travel depending on the timing. Therefore, it is necessary to continuously update surveys to keep up with changing trends of travel consumers in different countries and regions to better understand their needs and thoughts. In this study, we conducted surveys of respondents who had outbound travel experiences in China and Japan in August and September in 2022 respectively. The aim was to understand the differences in the impact of the COVID-19 pandemic on factors limiting the choices of outbound travel and changes in preferences among different groups, in order to provide valuable information for decision-makers in the tourism industry.

3.3 Methods

This study conducted an online survey on the impact of the COVID-19 pandemic on outbound travel using the same questionnaire for respondents from China and Japan, each collecting 300 responses (equally split between male and female). In China, we selected the Wenjuanxing network questionnaire survey company to conduct the survey because the company is commonly used by universities, research institutions, and government departments in China for collecting survey-related data such as questionnaires, evaluations, and voting. The survey was conducted from August 4th to 11th, 2022. In Japan, we also selected a market research company, Cross Marketing Inc., which is frequently used by companies, governments, and universities. The survey was conducted on September 5th and 6th, 2022. In addition, a previous survey conducted about the experiences of Chinese individuals traveling to Japan showed that the age of Chinese tourists is mainly between 20 and 30 years old, with those over 40 accounting for less than 3%. Additionally, most tourists were from major Chinese cities such as Shanghai, Beijing, and Guangdong [40]. Therefore, this study limited the survey subjects in China to those under 40 years of age (including 40), living in metropolises (Beijing, Shanghai, Guangzhou, Shenzhen), and having outbound travel experience. To facilitate a more effective comparative analysis, the same restrictions were applied to the survey subjects in Japan. Tokyo, Chiba, Kanagawa, Saitama, Nagoya, Osaka, Kyoto, and Hyogo were designated as metropolises in Japan. To ensure that all respondents met these criteria, a separate questionnaire was conducted before the survey began, and if they met the requirements, they continued to answer the questions, otherwise, the survey was terminated. The survey consisted of 11 questions distributed across the following six sections: 1) demographic characteristics; 2) perception of the COVID-19 pandemic at different times; 3) the impact of infection conditions and inbound/outbound management measures (quarantine, nucleic acid testing) at the destination on the choice of outbound

travel; 4) risk perception of outbound travel; 5) changes in travel preferences and destination choices compared to before the outbreak of COVID-19; 6) satisfaction with virtual outbound travel.

We used R (programming language) to analyze the survey data. Specifically, descriptive statistical (frequencies, percentages, means, standard deviations, etc.) were used to analyze Sections 1), 2), 3), and 4) of the survey, and these results were used to describe and discuss the characteristics of Chinese and Japanese respondents, their understanding of the COVID-19 pandemic during different periods, and their risk perception during outbound travel. Moreover, the differences in the effects of destination infection status and entry and exit management measures on the choice of outbound travel for Chinese and Japanese respondents were compared. In addition, ordinal logistic regression was used to analyze the differential in the impact of demographic characteristics on respondents' understanding of the COVID-19 pandemic, risk perceived and preference changes of outbound tourism. Furthermore, to test the changes in travel preferences of Chinese and Japanese respondents under different conditions, the variables in Section 5 were paired, and paired t-tests were performed on each pair of data. Independent sample t-tests were also used to compare the differences in the effects of the COVID-19 pandemic on Sections 2), 4), and 5) for Chinese and Japanese respondents. Both paired and independent sample t-tests were conducted using the t.test () package in R. p-value less than 0.05 indicates a significant difference between the two samples being tested.

3.4 Results

3.4.1 The results of descriptive statistical analysis

Demographic characteristics

Questions 1 to 6 investigated the characteristics of the respondents. As shown in **Table 3.1**, the age distribution of the 300 respondents from China is mainly concentrated in the age ranges of 21-30 and 31-40, accounting for over 98% in total. 92% of the respondents have a bachelor's degree or above, and 47.7% of the respondents have a monthly income between 8001 and 15000 yuan, while 27.3% of the respondents have an income exceeding 15000 yuan. Therefore, the sample can be regarded as a population with good education and higher income. In addition, 86.7% of the respondents live with their families. 88.6% of the respondents have traveled abroad two or more times, with the highest proportion being 52% for 2-3 times. On the other hand, the age distribution of the 300 respondents from Japan is mainly concentrated in the age range of 31-40, accounting for 76.3%, which is relatively older than the Chinese respondents. 73% of the respondents have a bachelor's degree or above, and the sample can also be regarded as well-educated. In addition, 73.3% of the respondents live with their families. The statistical results of income and outbound tourism times are evenly distributed in each stage, as detailed in **Table 3.1**.

Table 3.1 Demographic characteristics (Chinese: n = 300; Japanese: n = 300)

Personal information	Options	Chinese		Japanese	
		n	%	n	%
Gender	Male	150	50.0	150	50.0
	Female	150	50.0	150	50.0
Age	~ 20	5	1.7	4	1.3
	21 ~ 30	166	55.3	67	22.3
	31 ~ 40	129	43.0	229	76.3
Education	No high school	1	0.3	5	1.7
	High school	6	2.0	37	12.3
	Vocational school	17	5.7	20	6.7
	Undergraduate degree	242	80.7	183	61.0
	Graduate degree	34	11.3	36	12.0
	Other degree	0.0	0.0	19	6.3
Companion	Myself	40	13.3	68	22.7
	Family	260	86.7	220	73.3
	Other people	0.0	0.0	12	4.0
Revenue (yuan / thousand yen)	~ 3000 / ~100	7	2.3	54	18.0
	3001 ~ 5000 / ~ 200	14	4.7	53	17.7
	5001 ~ 8000 / ~ 300	54	18.0	94	31.3
	8001 ~ 15000 / ~ 400	143	47.7	62	20.7
	15001 ~ / 410 ~	82	27.3	37	12.3
Number of outbound tourisms	1	34	11.3	68	22.7
	2 ~ 3	156	52.0	89	29.7
	4 ~ 6	76	25.3	51	17.0
	7~	34	11.3	92	30.7

The respondents' understanding of the COVID-19 pandemic at different times

Question 7 investigated the respondents' perceptions of the COVID-19 pandemic at different periods, with six response options ranging from -3 (not afraid at all) to 3 (very afraid) in rank order (excluding 0). We calculated the mean and standard deviation of each variable. The results show that over time, the level of fear towards the COVID-19 pandemic has decreased among Chinese and Japanese participants. As shown in **Table 3.3**, the average fear score of Chinese participants was 1.76 in the early stage, 1.05 in the middle stage, and now is -0.18. The average fear score of Japanese participants was 1.03 in the early stage, 0.7 in the middle stage, and now is 0.26. However, the standard deviation

of Japanese participants at each stage was higher than that of Chinese participants, indicating significant variations in attitudes towards the COVID-19 pandemic among different Japanese respondents.

The factors that constrain respondents from choosing to travel abroad

Question 8 investigated the constraining factors that affect respondents' choices of traveling abroad. Specifically, the infection status of the destination, centralized quarantine measures for entry and exit, and nucleic acid testing policies were investigated. As shown in **Fig. 3.1**, both Chinese and Japanese respondents expressed a strong desire to cancel entry and exit quarantine measures. Regarding nucleic acid testing for entry and exit, most Chinese respondents wanted to reduce the frequency of testing, while Japanese respondents were more likely to want to cancel nucleic acid testing. In addition, regarding expectations after the disappearance of the COVID-19 epidemic, Chinese respondents were more inclined to choose outbound tourism after the epidemic has completely disappeared.

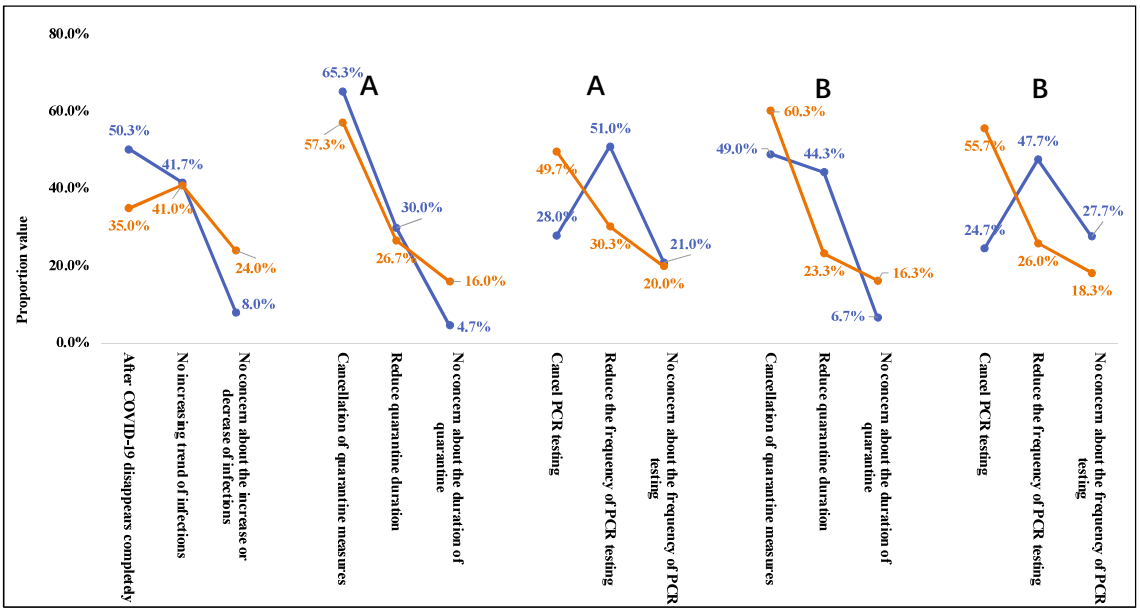


Fig. 3.1. The proportion of meet the conditions for outbound tourism (blue represents Chinese, orange represents Japanese; A refers to the inbound management measures of the tourism destination country, while B refers to the inbound entry management measures for returning home)

The risk perception of outbound travel by respondents

Question 9 investigated the risk perception of outbound tourism among respondents during the pandemic, with six answer options ranked from -3 (not worried at all) to 3 (extremely worried) in rank order (excluding 0). As shown in **Table 3.3**, the Chinese respondents' biggest concern was getting infected with COVID-19 while traveling abroad (M=2.01), followed by whether they could receive appropriate medical care after infection (M=1.87), and the increase in infected individuals during

travel (at the destination) that forces them to change their original travel plans (M=1.78). Japanese respondents shared similar concerns, with their top concerns being getting infected (M=1.11) and could receive appropriate medical care (M=1.12). In addition, they all expressed concern about the impact of pandemic prevention measures on their travel experience, such as not achieving the expected travel effects and safety issues related to food and accommodation facilities. However, they were less concerned about discrimination against them by local residents in COVID-19-related matters. Furthermore, the standard deviations of all variables were larger, indicating significant differences in perceptions of outbound tourism risks among different respondents. Compared to Chinese participants, the standard deviation range (1.72 to 1.91) of each variable for Japanese participants is smaller, indicating that the degree of difference between variables is relatively stable among Japanese participants.

The Impact of the COVID-19 pandemic on interviewees' travel preferences

Question 10 investigated whether the interviewees' travel preferences and destination choices have changed compared to before the pandemic. There are 7 answer options, ranging from -3 (strongly avoid) to 3 (very like), and 0 means no change. We paired the interviewees' travel preferences under different conditions, such as domestic travel and outbound tourism, individual travel and group travel, low-density and high-density destination, indoor and outdoor travel, and multi-destination and single-destination travel. This way, each pair of paired data constitutes a difference value (i.e., the difference in travel preferences of interviewees under two conditions), and we can perform a paired t-test on these difference values. As shown in **Table 3.2**, where the p-values for all five pairs of data are less than 0.001, indicating that the differences in travel preferences among Japanese and Chinese respondents under these conditions are significant. Specifically, both groups showed a preference for domestic travel, individual travel, and low-density and outdoor destinations. Moreover, respondents from both groups preferred to travel to as few destinations as possible within the same period. The study found that travel preferences and destination choices have changed significantly for both Chinese and Japanese interviewees before and after the pandemic.

Table 3.2 The matching t-test results for tourism preferences

Tourism preferences	Chinese (n=300)		Japanese (n=300)	
	t	p	t	p
Domestic travel / Outbound travel	16.89	0.000	10.53	0.000
Individual travel / Group travel	15.46	0.000	7.67	0.000
Destinations with lower population density / Destinations with higher population density	16.03	0.000	7.47	0.000
Indoor / Outdoor	-9.27	0.000	-3.45	0.001
Multi-destination travel / less-destination travel	-2.45	0.015	-4.95	0.000

Table 3.3 The results of descriptive statistics and independent samples t-test (Chinese /Japanese: n = 300)

Options	Chinese (sample 1) / Japanese (sample 2)			
	M	SD	t	p
Understanding of the covid-19				
Initial period (Mar to Aug 2020)	1.76 / 1.03	1.41 / 1.95	5.27	0.000
Mid-term (Sep 2020-Dec 2021)	1.05 / 0.7	1.29 / 1.88	2.65	0.008
Now (Jan 2022 - Aug 2022)	-0.18 / 0.26	1.53 / 1.93	-3.04	0.002
Risk perception				
Increased travel costs	0.84 / 0.91	1.45 / 1.72	-0.49	0.627
Infected with COVID-19	2.01 / 1.11	1.12 / 1.88	7.15	0.000
Availability of appropriate medical care after infection	1.87 / 1.12	1.21 / 1.9	5.81	0.000
Not achieving the expected travel effects	1.56 / 0.79	1.23 / 1.77	6.12	0.000
Adequacy of prevention and control measures at tourist attractions	1.16 / 0.72	1.46 / 1.87	3.23	0.001
Safety issues with food and accommodation facilities	1.51 / 0.73	1.43 / 1.85	5.82	0.000
Convenience of destination transportation	0.74 / 0.68	1.58 / 1.83	0.43	0.668
Increase in infected individuals at travel destination, forcing changes to original travel plans	1.78 / 0.94	1.19 / 1.83	6.7	0.000
Discriminated against with respect to COVID-19	0.87 / 0.52	1.72 / 1.91	2.33	0.020
Tourism preferences				
Domestic travel	1.64 / -0.28	1.11 / 1.34	19.05	0.000
Outbound travel	-0.13 / -1.12	1.75 / 1.46	7.49	0.000
Individual travel	1.43 / -0.40	1.22 / 1.29	17.84	0.000
Group travel	-0.2 / -0.92	1.44 / 1.39	6.16	0.000
Multi-destination travel	0.56 / -0.73	1.53 / 1.31	11.10	0.000
Less-destination travel	0.85 / -0.47	1.16 / 1.34	12.81	0.000
Destinations with lower population density	1.54 / -0.26	1.19 / 1.31	17.57	0.000
Destinations with higher population density	-0.25 / -0.75	1.45 / 1,30	4.44	0.000
Indoor	0.69 / -0.54	1.36 / 1.28	11.36	0.000
Outdoor	1.52 / -0.38	1.17 / 1.28	18.96	0.000

3.4.2 The results of ordinal logistic regression

We regarded demographic characteristics as independent variables and the respondents' understanding of COVID-19 at different times, risk perceptions associated with outbound travel, and changes in travel preferences as dependent variables. Ordered logistic regression analysis was conducted to understand the differences in the effects of demographic characteristics on each variable. The analysis results are shown in **Table 3.4**, where each row represents a different variable, and each column represents a different category of demographic characteristics, including gender, age, education level, cohabiting status, income, and frequency of outbound travel. The regression coefficients and p-values were output for each category (influencing factor). The regression coefficient represents the size and direction of the relationship between the category and the analyzed variable. A negative coefficient indicates a correlation with lower levels of understanding, while a positive coefficient indicates a correlation with higher levels of understanding. In addition, if the p-value for a category is less than or equal to 0.05, it is considered to have a significant impact on the corresponding variable.

As shown in **Table 3.4**, demographic characteristics have no significant impact on the understanding of COVID-19 among Chinese and Japanese participants at different stages. However, they do have a significant impact on the risk perception related to outbound travel and travel preferences. For instance, Chinese respondents' income is negatively correlated with the risk perception of discrimination related to COVID-19 infection. Age is positively correlated with the failure to achieve expected travel outcomes, and female participants are more concerned about not achieving the expected travel outcomes. Cohabitants are positively correlated with the discrimination related to COVID-19, indicating that groups traveling with family members are less concerned about discrimination related to COVID-19. Additionally, in terms of travel preferences, Chinese female participants strongly avoid indoor and densely populated destinations, and they expect to visit more attractions during a certain period of travel. There is a negative correlation between the age and income of respondents and indoor travel. The education level and number of outbound trips of respondents are positively correlated with travel to fewer destinations. The income of respondents is negatively correlated with group travel. The number of outbound trips is positively correlated with personal travel, low population density destinations, and indoor and outdoor travel. On the other hand, male Japanese respondents are particularly concerned about whether they can receive appropriate medical care after contracting COVID-19, achieving the expected travel effect, the effectiveness of epidemic prevention measures at tourist attractions, the safety of accommodation facilities, and the need to change their original travel plans due to an increase in infected individuals at a particular destination. In addition, the younger the respondents, the greater their concerns about contracting COVID-19, receiving appropriate medical care, the convenience of transportation at the destination, the need to change travel plans due to an increase in infected individuals at a particular destination, and discrimination related to COVID-19. The education level of respondents is positively correlated with receiving appropriate medical care

after contracting COVID-19 and the need to change the original travel plan due to an increase in infected individuals at a particular destination. In terms of travel preferences, Japanese female respondents strongly avoid outbound travel. Additionally, Japanese respondents prefer outdoor activities with family members.

Table 3.4 The results of ordinal logistic regression (1)

Options	Gender		Age		Education		Companion		Income		Frequency	
	coef	p	coef	p	coef	p	coef	p	coef	p	coef	p
Understanding of the covid-19												
Initial period	-0.22	0.31	0.19	0.39	-0.19	0.38	0.15	0.35	-0.05	0.70	-0.09	0.50
	0.41	**0.09**	**0.29**	**0.07**	**0.13**	**0.21**	**-0.04**	**0.72**	**-0.02**	**0.87**	**0.05**	**0.60**
Mid-term	-0.26	0.23	-0.21	0.36	-0.23	0.27	0.23	0.15	-0.17	0.21	-0.11	0.41
	0.31	**0.19**	**0.14**	**0.39**	**0.01**	**0.94**	**-0.03**	**0.76**	**-0.10**	**0.30**	**0.03**	**0.76**
Now	0.04	0.85	-0.08	0.72	-0.06	0.75	0.19	0.22	-0.06	0.66	-0.14	0.31
	0.09	**0.70**	**0.05**	**0.75**	**-0.04**	**0.69**	**-0.09**	**0.41**	**-0.09**	**0.33**	**0.02**	**0.80**
risk perception												
Increased costs	-0.05	0.80	0.24	0.28	-0.14	0.52	-0.04	0.83	-0.26	0.06	-0.22	0.11
	0.27	**0.25**	**0.23**	**0.15**	**0.13**	**0.20**	**0.05**	**0.66**	**0.06**	**0.56**	**0.14**	**0.15**
Infected covid-19	-0.15	0.50	0.07	0.75	-0.19	0.41	0.23	0.16	-0.29	0.04	0.02	0.89
	0.42	**0.07**	**0.45**	**0.01**	**0.19**	**0.07**	**-0.04**	**0.73**	**-0.14**	**0.14**	**0.08**	**0.43**
medical care after infection	-0.03	0.88	-0.08	0.72	-0.33	0.13	0.10	0.51	0.08	0.56	0.18	0.19
	0.55	**0.02**	**0.52**	**0.00**	**0.25**	**0.02**	**-0.03**	**0.81**	**-0.19**	**0.06**	**0.05**	**0.63**
Not achieving expected effects	-0.56	0.01	0.60	0.01	-0.14	0.50	-0.01	0.97	-0.03	0.82	0.01	0.96
	0.55	**0.02**	**0.42**	**0.01**	**0.15**	**0.14**	**-0.04**	**0.69**	**-0.02**	**0.82**	**0.04**	**0.63**
Precautions for attractions	-0.24	0.26	0.07	0.74	-0.34	0.11	0.07	0.67	-0.03	0.82	-0.05	0.68
	0.63	**0.01**	**0.30**	**0.06**	**0.13**	**0.22**	**-0.12**	**0.28**	**-0.08**	**0.39**	**0.13**	**0.16**
Accommodation facilities (safety)	-0.30	0.16	0.12	0.59	-0.20	0.35	0.22	0.16	-0.23	0.11	-0.01	0.94
	0.62	**0.01**	**0.40**	**0.01**	**0.12**	**0.26**	**-0.12**	**0.27**	**-0.04**	**0.65**	**0.06**	**0.52**
Destination transportation	-0.40	0.06	0.06	0.79	-0.29	0.17	0.20	0.21	-0.18	0.19	-0.18	0.18
	0.41	**0.09**	**0.50**	**0.00**	**0.15**	**0.15**	**-0.01**	**0.89**	**-0.05**	**0.63**	**0.03**	**0.75**
Forcing changes to travel plans	0.12	0.57	-0.18	0.42	0.04	0.85	0.05	0.79	-0.12	0.39	0.26	0.06
	0.68	**0.01**	**0.39**	**0.02**	**0.27**	**0.01**	**-0.02**	**0.88**	**0.03**	**0.77**	**0.15**	**0.12**
Discriminated with covid-19	0.03	0.90	0.01	0.95	-0.08	0.72	0.41	0.01	-0.43	0.00	-0.10	0.47
	0.18	**0.45**	**0.30**	**0.05**	**0.09**	**0.37**	**-0.09**	**0.41**	**-0.12**	**0.19**	**-0.07**	**0.44**

Table 3.4 The results of ordinal logistic regression (2)

Options	Gender		Age		Education		Companion		Income		Frequency	
	coef	p	coef	p	coef	p		coef	p	coef	p	coef
Tourism preferences												
Domestic travel	-0.41	0.06	-0.02	0.94	-0.29	0.21	0.13	0.42	0.00	0.98	0.24	0.08
	-0.14	**0.58**	**0.04**	**0.83**	**-0.13**	**0.20**	**-0.08**	**0.47**	**0.07**	**0.44**	**0.01**	**0.91**
Outbound travel	-0.24	0.25	0.16	0.47	0.31	0.14	-0.16	0.31	-0.14	0.29	0.22	0.10
	-0.62	**0.01**	**0.02**	**0.90**	**-0.16**	**0.12**	**-0.10**	**0.36**	**0.04**	**0.65**	**-0.11**	**0.23**
Independent travel	-0.33	0.12	0.30	0.17	-0.08	0.70	0.05	0.77	-0.07	0.59	0.30	0.03
	0.14	**0.60**	**0.17**	**0.34**	**-0.04**	**0.68**	**-0.13**	**0.24**	**0.20**	**0.06**	**0.00**	**0.97**
Group travel	-0.27	0.20	0.02	0.93	0.22	0.31	0.02	0.91	-0.35	0.01	0.02	0.90
	-0.17	**0.48**	**-0.05**	**0.76**	**-0.09**	**0.39**	**-0.05**	**0.67**	**0.10**	**0.33**	**-0.01**	**0.90**
Multi-destination travel	-0.19	0.37	-0.30	0.17	-0.08	0.72	0.02	0.88	-0.10	0.47	0.03	0.85
	-0.29	**0.25**	**0.02**	**0.93**	**-0.07**	**0.50**	**0.08**	**0.46**	**0.07**	**0.45**	**-0.08**	**0.41**
Less-destination travel	-0.63	0.00	0.36	0.10	0.46	0.03	0.28	0.08	-0.18	0.17	0.32	0.02
	0.10	**0.68**	**0.10**	**0.56**	**0.05**	**0.62**	**-0.02**	**0.86**	**0.03**	**0.77**	**0.03**	**0.74**
Lower density (population)	-0.22	0.31	0.31	0.16	-0.17	0.44	-0.05	0.78	0.14	0.29	0.34	0.02
	0.27	**0.28**	**0.15**	**0.39**	**-0.03**	**0.79**	**0.07**	**0.51**	**0.01**	**0.90**	**0.08**	**0.44**
Higher density (population)	-0.47	0.03	0.04	0.84	0.22	0.29	-0.30	0.06	-0.20	0.15	0.03	0.80
	-0.27	**0.28**	**-0.09**	**0.60**	**-0.01**	**0.92**	**0.05**	**0.65**	**0.12**	**0.23**	**-0.09**	**0.37**
Indoor	-0.53	0.01	-0.58	0.01	0.11	0.60	-0.10	0.55	-0.28	0.04	0.39	0.01
	-0.08	**0.74**	**-0.01**	**0.95**	**0.04**	**0.74**	**0.05**	**0.67**	**0.08**	**0.43**	**0.06**	**0.56**
Outdoor	0.20	0.35	0.08	0.72	-0.21	0.32	-0.08	0.61	-0.09	0.51	0.38	0.00
	0.09	**0.74**	**-0.08**	**0.64**	**0.06**	**0.58**	**0.22**	**0.05**	**0.10**	**0.32**	**-0.03**	**0.79**

3.4.3 The differential impacts of the COVID-19 pandemic on outbound tourists from China and Japan

To understand the differences in perceptions of COVID-19 between Chinese and Japanese respondents at different stages, independent sample t-tests were conducted on their responses to questions related to their perception of the severity of the COVID-19 pandemic in the initial period, mid-term, and now. As shown in **Table 3.3**, the absolute values of the t-values for each stage were large, and the corresponding p-values were all less than 0.01, indicating a significant difference in the understanding of the coronavirus among Chinese and Japanese respondents at different times. Since the t-values for the early and middle stages were positive, it indicates that Chinese respondents gave higher ratings to the severity of the coronavirus than Japanese respondents in the early and middle stages of the

pandemic. In addition, the current t-value was negative, indicating that Japanese respondents are now more afraid of the coronavirus. This may be because at the time of conducting our online survey of outbound tourists, Japan was experiencing its seventh wave of COVID-19 outbreak, while the situation in China was relatively more stable [41].

In addition, to understand potential differences in perceptions of risks associated with outbound tourism between Chinese and Japanese respondents, we conducted independent samples t-tests on their scores for each perceived risk. As shown in **Table 3.3**, the p-values corresponding to the increase in travel cost and convenience of destination transportation are much greater than 0.05, indicating that there were no significant differences in the perceived risks of these variables between Chinese and Japanese respondents. However, the P values corresponding to other perceived risks are all less than 0.05, and the range of t values is between 3.23 and 7.15, indicating significant differences in perceived risks for these variables between Chinese and Japanese respondents, and that Chinese respondents perceive greater risk for these variables compared to Japanese respondents, particularly for contracting COVID-19, whether could receive opportune medical care after infection, having to change the original travel schedule due to an increase in infected persons at the travel destination, not achieving the expected travel outcome, and safety issues with food and accommodation facilities.

Furthermore, to understand the differences in the changes of preferences for travel mode and destination among Chinese and Japanese respondents compared to before the pandemic, independent sample t-tests were conducted on the rating data for each variable of Question 10. As shown in **Table 3.3**, the p-values for all variables are much less than 0.05, and the t-values are larger, indicating significant differences in preferences for these variables between Chinese and Japanese respondents. Based on their t-values, it can be found that Chinese respondents have higher preferences for all these variables than Japanese respondents, especially for domestic travel, individual travel, low population density destination, and outdoor activities.

3.5 Conclusion and Discussion

In chapter 3, conducted a survey on the impact of the COVID-19 pandemic on outbound travel among 300 respondents with travel experience from China and Japan respectively. The survey data was analyzed using descriptive statistics, ordinal logistic regression and t-tests. Aimed at understanding the impacts of the COVID-19 pandemic on outbound tourists from different groups.

The results indicated that the fear of COVID-19 among Chinese and Japanese respondents has decreased over time. However, there are significant differences in attitudes towards the pandemic among different respondents. Therefore, more specific investigations and analyses are needed to better understand the needs and concerns of travel consumers regarding these issues.

In terms of entry and exit management measures, both Chinese and Japanese respondents expressed a strong desire to cancel entry and exit quarantine measures. Additionally, most Chinese respondents hoped to reduce the number of nucleic acid tests required for entry and exit, while most Japanese respondents hoped to cancel nucleic acid testing. Therefore, it can be assumed that Japanese respondents had a lower willingness to accept various restrictions and conditions when traveling abroad.

Regarding the perception of risk in outbound tourism, Chinese and Japanese respondents are most concerned about being infected with COVID-19 during the trip, followed by whether they can receive appropriate medical care after infection, and the increase in infected individuals during travel (at the destination) that forces them to change their original travel plans. Therefore, future tourism consumers who choose to travel abroad may be more inclined to destinations with well-developed infrastructure and high-quality medical facilities. Additionally, businesses associated with the tourism industry, such as airlines, hotels, and restaurants, should also improve the convenience of their rebooking and cancellation services to enhance the satisfaction and loyalty of tourism consumers. Moreover, the results of the independent t-test on two samples showed that, Chinese respondents are more concerned about the risks of outbound travel.

In addition, analyzing the changes in the travel preferences of respondents from the two countries, it was found that Chinese and Japanese respondents tend to choose domestic travel, individual travel, destinations with low population density, outdoor activities, and travel to as few destinations as possible for the same amount of time. Moreover, the results also showed that Chinese respondents had greater preferences than Japanese respondents for all modes of travel and destinations, especially for domestic travel, independent travel, destinations with lower population density, and outdoor activities. Since all t-values on tourism preferences in **Table 3.3** are positive, it seems that Japanese respondents were more cautious or negative toward all types of tourism.

According to the analysis results of ordered logistic regression, demographic characteristics do not have a significant impact on respondents' understanding of COVID-19 at different periods, but they can affect respondents' risk perceptions and travel preferences related to outbound tourism. For example, older female Chinese respondents are more concerned about not achieving the expected travel experience. In addition, female respondents strongly avoid indoor and densely populated destinations and expect to visit more scenic spots. Furthermore, respondents with higher education and more outbound travel experience prefer to choose a small number of destinations for a fulfilling travel experience. On the other hand, younger male Japanese respondents are particularly concerned about the risks of outbound travel. Respondents with higher education have lower concerns about the risks of outbound travel. Japanese female respondents strongly avoid outbound travel. Respondents prefer outdoor activities with their families.

In addition, during the COVID-19 pandemic, differing social and political environments in China and Japan may have led to variations in their perceptions of travel risks. The following is a possible political and social background of these differences in risk perception during the pandemic.

From a political perspective, China and Japan adopted different strategies and measures in response to the COVID-19 virus. After the outbreak, the Chinese government quickly implemented strict preventive measures, including city lockdowns, restrictions on movement, and quarantine policies. These strong measures might have led to a more unified perception of travel risks among the public during the early stages of the pandemic. People relied heavily on government directives and regulations for their travel decisions, which left little room for personal judgment. As the pandemic has come under control, the Chinese government has gradually relaxed its preventive measures. However, the stringent policies implemented earlier may have left a lasting impression on the public, leading to continued high levels of concern about travel risks. In contrast, the Japanese government initially adopted relatively lenient preventive measures, avoiding mandatory actions as much as possible and relying more on the public's self-awareness and sense of social responsibility. This approach somewhat reduced public concern about travel risks but also increased the risk of virus spread. As the pandemic progressed, the Japanese government gradually strengthened preventive measures, including restricting international travel and enhancing entry management. However, Japanese citizens largely relied on personal judgment and voluntary compliance when deciding whether to travel, which could lead to significant differences in travel risk perception among different groups.

From a social perspective, it is possible that China's collectivist culture and high trust in government authority made it easier for the public to accept the government's preventive measures, including travel restrictions. Chinese society may place greater emphasis on public safety and social responsibility, so individuals' perceptions of travel risks are more influenced by government propaganda and public health advice. In contrast, while Japanese culture also has collectivist elements, it may place greater emphasis on personal freedom and autonomous decision-making rather than complete reliance on government directives. Additionally, Japanese society exhibited a desire to maintain normal daily life, exemplified by the implementation of the “Go to Travel” campaign during the pandemic, which encouraged domestic travel and potentially alleviated concerns about traveling within the country to some extent.

In summary, the high trust in government and collectivist culture in Chinese society may lead to a more unified perception of travel risks, with the public strictly adhering to government instructions. On the other hand, Japanese society's emphasis on personal freedom and autonomous decision-making, which can result in a more individual autonomy and diversity in dealing with travel risks.

References

1. Uğur, N. G., Akbıyık, A.: Impacts of COVID-19 on global tourism industry: A cross-regional comparison. Tourism management perspectives, Vol. 36, 100744 (2020)
2. CONNOR, P.: More than nine-in-ten people worldwide live in countries with travel restrictions amid COVID-19. (2020)
3. Luković, S., Stojković, D.: Covid-19 pandemic and global tourism. Hotel and Tourism Management, Vol. 8(2), 70-88 (2020)
4. Gössling, S., Scott, D., Hall, C.M.: Pandemics, tourism and global change: A rapid assessment of COVID-19. Journal of Sustainable Tourism, Vol. 29(1), pp.1–20 (2021)
5. Sharma, A., Nicolau, J. L.: An open market valuation of the effects of COVID-19 on the travel and tourism industry. Annals of Tourism Research, Vol. 83, 102990 (2020)
6. UNWTO.: IMPACT ASSESSMENT OF THE COVID-19 OUTBREAK ON INTERNATIONAL TOURISM, https://www.unwto.org/impact-assessment-of-the-covid-19-outbreak-on-international-tourism
7. Kupi, M., Szemerédi, E.: Impact of the COVID-19 on the Destination Choices of Hungarian Tourists: A Comparative Analysis. Sustainability, Vol. 13(24), 13785 (2021)
8. Ivanova, M., Ivanov, I. K., Ivanov, S.: Travel behaviour after the pandemic: The case of Bulgaria. Anatolia, Vol. 32(1), 1-11 (2021)
9. Mirzaei, R., Sadin, M., Pedram, M.: Tourism and COVID-19: changes in travel patterns and tourists' behavior in Iran. Journal of Tourism Futures, Vol. 9(1), ISSN: 2055-5911 (2021)
10. Riestyaningrum, F., Ferdaos, E., Bayramov, B.: Customer behavior impact on international tourist's travel intention due to Covid-19. Journal of Sustainable Tourism and Entrepreneurship, Vol. 1, No. 3, pp. 231-243 (2020)
11. Orîndaru, A., Popescu, M.-F., Alexoaei, A.P., Căescu, Ș.-C., Florescu, M.S., Orzan, A.-O.: Tourism in a Post-COVID-19 Era: Sustainable Strategies for Industry's Recovery. Sustainability, Vol. 13(12), 6781 (2021)
12. Huang, S., Shao, Y., Zeng, Y., Liu, X., Li, Z.: Impacts of COVID-19 on Chinese nationals' tourism preferences. Tourism Management Perspectives, Vol 40, ISSN 2211-9736 (2021)
13. Li, Z., Zhang, X., Yang, K., Singer, R., Cui, R.: Urban and rural tourism under COVID-19 in China: research on the recovery measures and tourism development. Tourism Review, Vol. 76(4), pp. 718-736 (2021)
14. Kinczel, A., Müller, A.: STUDY ON TRAVEL HABITS AND LEISURE ACTIVITIES IN THE LIGHT OF COVID-19 TRIGGERED CHANGES IN ROMANIA AND HUNGARY. GeoJournal of Tourism and Geosites, Vol. 41, No. 2, pp. 440-447 (2022)
15. Umukroro, G. M., Odey, V.E., Yta, E.M.: The effect of pandemic on homebased tourism: post Covid-19. International Journal of Humanities and Innovation (IJHI), Vol. 3, No. 3, pp. 115-120 (2020)

16. Fan, X., Lu, J., Qiu, M., Xiao X.: Changes in travel behaviors and intentions during the COVID-19 pandemic and recovery period: A case study of China. Journal of Outdoor Recreation and Tourism, Vol. 41, 100522 (2022)
17. Marques, C.P., Guedes, A., Bento, R.: Rural tourism recovery between two COVID-19 waves: the case of Portugal. Current Issues in Tourism, Vol. 25(6), 857-863 (2022)
18. Wachyuni, S.S., Kusumaningrum, D.A.: The Effect of COVID-19 Pandemic: How are the Future Tourist Behavior?. Journal of Education, Society and Behavioural Science, Vol. 33(4), 67-76 (2020)
19. Jafari, K., Saydam, M. B., Erkanlı, E., Olorunsola, V. O.: The impacts of the COVID-19 pandemic on the consumer behavior of Turkish tourists. Revista Turismo Estudos e Práticas-RTEP/UERN, No. 5, 1-17 (2020)
20. Popescu, A.: The impact of COVID-19 pandemic on Romania's tourist flows in the year 2020. Scientific Papers Series Management, Economic Engineering in Agriculture and Rural Development, Vol. 21(1), 655-666 (2021)
21. Wen, J., Kozak, M., Yang, S., Liu, F.: COVID-19: potential effects on Chinese citizens' lifestyle and travel. Tourism Review, Vol. 76(1), 74-87 (2021)
22. Afanasiev, O.E., Afanasieva, A.V.: The Effect of COVID-19 Pandemic on Russian Tourists' Demand and Preferences. Tourism Through Troubled Times, pp. 203-230 (2022)
23. Li, X., Gong, J., Gao, B., Yuan, P.: Impacts of COVID-19 on tourists' destination preferences: Evidence from China. Annals of Tourism Research, Vol. 90, 103258 (2021)
24. Lebrun, AM., Corbel, R., Bouchet, P.: Impacts of Covid-19 on travel intention for summer 2020: A trend in proximity tourism mediated by an attitude towards Covid-19. Service Business, 1-33 (2021)
25. Roman, M., Niedziółka, A., Krasnodębski, A.: Respondents' involvement in tourist activities at the time of the COVID-19 pandemic. Sustainability, Vol. 12(22), 9610, 2020
26. Aronica, M., Pizzuto, P., Sciortino, C.: COVID-19 and tourism: What can we learn from the past?. The World Economy, Vol. 45(2), pp. 430-444 (2022)
27. Chemli, S., Toanoglou, M., Valeri, M.: The impact of Covid-19 media coverage on tourist's awareness for future travelling. Current Issues in Tourism, 25:2, 179-186 (2022)
28. Łapko, A., Hącia, E., Strulak-Wójcikiewicz, R., Çınar, K., Panai, E., Lučić, L.: Eco-friendly tourism decision making during COVID-19—sailing tourism example. Sustainability, Vol. 14(1), 134, 2022
29. Song, H., Li, G., Cao, Z.: Tourism and economic globalization: an emerging research agenda. Journal of Travel Research, Vol. 57(8), 999-1011 (2018)
30. UNWTO: UNWTO TOURISM DATA DASHBOARD. https://www.unwto.org/tourism-data/global-and-regional-tourism-performance
31. Zhan, L., Zeng, X., Morrison, A. M., Liang, H., Coca-Stefaniak, J. A.: A risk perception scale for travel to a crisis epicentre: Visiting Wuhan after COVID-19. Current Issues in Tourism, Vol. 25(1), pp. 150-167, 2022

32. Jin, X., Bao, J., Tang, C.: Profiling and evaluating Chinese consumers regarding post-COVID-19 travel. Current Issues in Tourism, Vol. 25(5), pp. 745-763, 2022
33. Perić, G., Dramićanin, S., Conić, M.: The impact of Serbian tourists' risk perception on their travel intentions during the COVID-19 pandemic. European Journal of Tourism Research, Vol. 27, 2705, 2021
34. Liu, Y., Shi, H., Li, Y. and Amin, A.: Factors influencing Chinese residents' post-pandemic outbound travel intentions: an extended theory of planned behavior model based on the perception of COVID-19. Tourism Review, Vol. 76(4), 871-891 (2021)
35. Moya Calderón, M., Chavarría Esquivel, K., Arrieta García, M. M., Lozano, C. B.: Tourist behaviour and dynamics of domestic tourism in times of COVID-19. Current Issues in Tourism, Vol. 25(14), pp. 2207-2211 (2022)
36. Santos, N., Moreira, C. O.: Uncertainty and expectations in Portugal's tourism activities. Impacts of COVID-19. Research in Globalization, Vol. 3, 100071 (2021)
37. Orden-Mejía, M., Carvache-Franco, M., Huertas, A., Carvache-Franco, W., Landeta-Bejarano, N., Carvache-Franco, O.: Post-COVID-19 tourists' preferences, attitudes and travel expectations: A Study in Guayaquil, Ecuador. International Journal of Environmental Research and Public Health, Vol. 19(8), 4822 (2022)
38. Park, I. J., Kim, J., Kim, S. S., Lee, J. C., Giroux, M.: Impact of the COVID-19 pandemic on travelers' preference for crowded versus non-crowded options. Tourism Management, Vol. 87, 104398 (2021)
39. NAKAMURA, T.: Intention of Japanese to Travel Abroad During and After the COVID-19 Pandemic. The journal of Tamagawa University College of Tourism and Hospitality, No. 9, pp. 43-61 (2022) (Japanese)
40. Ming, H., Hiroyuki, S.: ANALYSIS ON THE BAHAVIORS OF INDIVIDUAL CHINESE TOURISTS IN JAPAN USING WEB QUESTIONNAIRE SURVEY IN CHINA. The 63rd Research Presentation・Spring Conference of Infrastructure Planning and Management (2021). https://onsite.gakkai-web.net/jsce/ip_au/time_table.html
41. Center for Systems Science and Engineering.: Tracking COVID-19. https://www.arcgis.com/apps/dashboards/bda7594740fd40299423467b48e9ecf6

CHAPTER 4
Summarize

The tourism data before the outbreak of the COVID-19 pandemic provided insights on the visit patterns, tourism themes and their trends. These data could be utilized for forecasting tourism demand and flow under normal circumstances, assessing post-pandemic recovery, and serving as a benchmark for measuring progress. However, relying solely on this data is risky, it is essential to complement it with the latest travel data and changes in tourist preferences to formulate flexible and adaptive strategies. This paper analyzes the annual changes in the tourism behavior of Chinese tourists to Japan before the pandemic and explores the impact of the COVID-19 pandemic on the attitudes and preferences of potential outbound Chinese tourists. Its aim is to provide effective policy recommendations to local governments in Japan to aid in the recovery of their tourism industry post-pandemic.

In Chapter 2, we used the LDA model and Word2vec model of natural language processing to analyze the web travel notes posted on the Chinese visitor website Mafengwo written by Chinese tourists who visited Japan before the pandemic. Using the analytical results, we identified the tourism themes and the visit patterns of the Chinese tourists travelling to Japan and how they have changed. Furthermore, we also elucidated impressive and attractive local and regional tourism resources in some regions in Japan. On the other hand, considering the safety and health concerns brought about by the COVID-19 pandemic, visitors may be more inclined to choose localized and smaller-scale tourist destinations, avoiding crowded attractions and large events to reduce the risk of infection. Therefore, post-pandemic recovery strategies for the tourism industry need to be developed and adjusted in conjunction with new market demands and tourism behaviors. Relying solely on pre-pandemic data may not comprehensively address the complexities of the post-pandemic tourism environment. In Chapter 3, we conducted an online survey with Chinese respondents who had previous experience with outbound tourism. The purpose was to timely understand the attitude and preference changes of Chinese travel consumers towards outbound travel after the COVID-19 pandemic.

The analysis results of Chapter 2 and Chapter 3 are discussed in Sections 4.1 and 4.2 below, respectively.

4.1 Discussion on the policy to promote inbound tourism in local regions

In this section, based on the analytical results from LDA and Word2vec models, the policy to promote inbound tourism in local regions in Japan is discussed.

The local regions far from metropolitan regions and major tourism sites are considered to be at a relative disadvantage in terms of inbound tourism. However, as shown in Fig. 2.3, the Chinese tourists mentioned not only the topics on the metropolitan regions, but also those on the local regions such as

Seto Inland Sea /Takamatsu and Hiroshima_Okayama_Tottori. Figure 2.6 also shows that the percentage of mentions of the visit patterns in many local regions is increasing.These results suggest that other local regions also have the opportunity to promote inbound tourism from China.

From Table 2.2, Seto Inland Sea /Takamatsu visit pattern included “art gallery,” and Hiroshima_Okayama_Tottori included “Conan,” together with the major tourism sites in these regions. These are unique local resources that cannot be experienced elsewhere. Furthermore, Table 2.4 shows that Chinese tourists mentioned small municipalities and local foods, and festivals, etc. Additionally, we can focus not only on the mentioned words, but also on the unmentioned words. In other words, we can determine which travel information within specific region has been extracted and which has not, allowing us to predict the preferences of Chinese tourists traveling to the local regions in Japan. The information on the preferences of Chinese tourists can be useful to propose an effective tourism promotion.

4.2 The differences impact of the COVID-19 pandemic on tourism consumers in different countries

Based on the analytical results from Chapter 3, we found that Chinese and Japanese respondents showed consistency in their concerns and changes in travel preferences when choosing outbound travel. However, Chinese respondents showed greater concern for the risks brought by the COVID-19 pandemic and were more inclined to choose domestic travel, individual travel, destinations with lower population density, and outdoor activities.

Furthermore, owing to the spread of the coronavirus, safety and health considerations have been brought to tourists, if some countries or regions fail to effectively respond to the COVID-19 pandemic and lack corresponding health security measures, they may lose market share from risk-aware tourism consumers. At the same time, destinations that can provide high-level health protection and effective risk management measures may attract more tourism consumers.

4.3 Conclusions

The findings obtained from the analysis in Chapters 2 and 3 were summarized as follows.

The Findings:

1. Chapter 2 of the LDA model analysis results indicate that the proportion of high-representation tourism themes (such as Gastronomy, Hot-Spring Hotel, Shrine/temple, and Shopping) is gradually decreasing year by year, while the proportion of low-representation tourism themes (such as Family trip, Cruise, and Fireworks) is gradually increasing.

2. The analysis results of Chapter 2 indicate that the annual corresponding ratio (R'_y values) outside of Hokkaido has generally shown an upward trend, particularly with significant increases in the Seto Inland Sea / Takamatsu, Tohoku region, and Hiroshima_Okayama_Tottori, etc. Moreover, the analysis results from Chapter 3 indicate that, after the COVID-19 pandemic, Chinese tourists are more likely to choose destinations (such as local regions) with lower COVID-19 impact or less contact.
3. According to tourism statistics from the Japan Tourism Agency, it was found that the proportion of individual Chinese tourists visiting Japan was only 25.6% in 2015, but by 2019, this proportion had increased to 62.3%. Moreover, the survey results of Chapter 3 indicate that Chinese tourists are more inclined to choose individual travel after the pandemic.
4. In Chapter 2 the analysis results of the LDA and Word2Vec models elucidate that the regional tourism resources of certain areas. For example, Aomori's Nebuta Festival, Okayama's Momotaro, Tottori's Conan, Seto Inland Sea/Takamatsu's art festival, etc. These are unique local resources that cannot be experienced elsewhere. Based on the analysis results, we can determine which travel information within a specific region has been extracted and which has not, allowing us to estimate about the preferences of Chinese tourists traveling to local regions in Japan.
5. This study utilizes the Word2Vec model to analyze word vectors related to specific place names and extract tourism information pertaining to these regions. As shown in Table 2.4, this tourism information includes visited places, attractions, tourism resources, and shop names, etc.
6. Chapter 3 investigated the necessary conditions by Chinese and Japanese respondents when choosing to travel abroad. The results indicated that both Chinese and Japanese respondents expressed a strong desire for the cancellation of quarantine measures for entry and exit. Additionally, Chinese respondents hoped for a reduction in the number of PCR tests required upon entry and exit.
7. Chapter 3 survey results on the risk perception of outbound travel among Chinese and Japanese respondents indicate that the biggest concern for both groups is the risk of contracting COVID-19 during their travels. Chinese and Japanese respondents are most concerned about being infected with COVID-19 during the trip, followed by whether they can receive appropriate medical care after infection, and the increase in infected individuals during travel (at the destination) that forces them to change their original travel plans.

According to these conclusions, the possibilities on the future trends of Chinese tourists traveling to Japan can be shown as follows.

Considerations based on findings obtained from analysis:

1) Based on Finding 2, the number of Chinese tourists visiting local regions in Japan may increase after the pandemic.

2) Based on Finding 3, the proportion of Chinese individual travelers to Japan will continue to increase after the pandemic.
3) Based on Finding 5, using the geographical information extracted from Table 2.4 can be used to plan or update local tourism routes and provide suggestions for sightseeing transportation routes to the relevant departments.
4) Based on Finding 6, lifting quarantine measures for entry and exit, as well as reducing the number of PCR tests required for travel, can be necessary conditions for Chinese tourists when choosing to travel abroad.
5) Based on Finding 7, improving local medical facilities, enhancing the hygiene standards of tourism infrastructure, and establishing emergency medical plans for foreign tourists may attract more tourism consumers after the pandemic.

Finally, strategies for promoting the recovery of inbound tourism from China in local regions of Japan can be discussed as follows.

Discussions on Recovery Strategies of Inbound Tourism from China in Local Regions in Japan After COVID-19:

A. Based on Finding 1, it can be observed that the travel purposes of Chinese tourists are becoming more diversified. Therefore, developing new types of travel experiences can become an important strategy for attracting Chinese tourists in the future.
B. To enhance mobility convenience for tourists traveling locally, particularly by visit patterns with increased visitation rates (R'_y values) identified in Chapter 2 analysis, such as those in the Seto Inland Sea, Takamatsu, Tohoku region, and Hiroshima-Okayama-Tottori, etc.
C. Based on the analysis results from Table 2.2, Figure 2.8, and Table 2.4, it is understood that Chinese tourists are inclined towards experiences in which they can directly engage, such as participating in seasonal events (e.g., Sakura viewing, Fireworks viewing, Nebuta Festival), visiting theme parks (e.g., USJ, Huis Ten Bosch), skiing (in the Tohoku region), enjoying local gastronomy (e.g., Sendai's beef tongue), relaxing in hot springs (e.g., Akiho Onsen, Beppu Hatto), and visiting anime pilgrimage sites (e.g., Conan in Tottori). By developing and promoting tourism resources that offer such experiences within the local region, it is expected that the interest of Chinese tourists can be effectively attracted.
D. Using the geographical information extracted from Table 2.4, we can analyze the tourist attractions of particular interest to Chinese tourists, such as famous shrines, Sakura spots, and natural scenic areas. This will help us formulate corresponding development and conservation policies to ensure that these resources receive proper maintenance and protection, thereby promoting the sustainable development of the local tourism industry.

E. In some regional areas of Japan, especially those with unique tourism resources, providing easily accessible medical services and health information, including multilingual health guides and emergency contact information, can enhance tourists' confidence in health protection.

4.4 Limitations and future research

The online travel notes data are mainly generated by tourists, but it is inevitable that some diaries are written by professional advertisers. Designing recognition rules to identify and remove such travel diaries are urgently needed to improve the quality of the data. Therefore, the analysis results of this study are for reference only and cannot represent the tourism behavior of all Chinese tourists traveling to Japan. Furthermore, as the user groups of the different travel websites can be different, it is necessary to use the travel data of other websites for comparative analysis so as to better understand the travel behavior of Chinese tourists traveling to Japan. In addition, the tourists' attribute data in the website used in this study does not include the number of trips to Japan. Therefore, this study cannot explain whether repeat visits to Japan by Chinese tourists have influenced the increase in local travel in recent years. In the future, researchers might consider supplementing this data to analyze the visit patterns and changes among Chinese tourists who travel to Japan multiple times.

The survey limited the age, residence, and outbound tourism experience of the respondents, which may have limitations. Moreover, the respondents' attitudes towards the COVID-19 epidemic, concerns about outbound tourism risks, travel preferences, and destination choices are subjective and are also affected by factors such as survey timing, location, and sample characteristics. Therefore, this data can only serve as a reference for overall trends and cannot represent the views and actual situations of all individuals. On the other hand, this study only investigated the impact of the COVID-19 epidemic on domestic respondents in China and Japan and cannot represent the situation of tourism consumers in other countries or regions. To gain a more comprehensive understanding of tourism consumers, future research could consider expanding the range of respondents to include more tourism consumers from different countries and regions to obtain more comprehensive and objective data and conclusions.

ACKNOWLEDGMENTS:

I would like to express my gratitude to all those who helped me during the writing of this thesis.

Firstly, I would like to express my sincere and heartfelt gratitude to my supervisor, Professor Hiroyuki Sakakibara, for his meticulous guidance on this research at regular seminars. He has walked me through all the stages of the writing of this thesis. Without his consistent and illuminating instruction, this thesis could not have reached its present form.

Secondly, I would like to express my gratitude to Associate Professor Makoto Tsukai from Hiroshima University for providing valuable insights on the utilization of Latent Dirichlet Allocation and Word2vec models.

Thirdly, I would like to extend my sincere gratitude to everyone at the Urban and Social Systems Engineering Laboratory for their tremendous assistance and cooperation throughout the research process.

Finally, this research was conducted with the support of a Grant-in-Aid for Scientific Research (JPMJSP2111), and we express our gratitude for this support.