

# Self-organizing map analysis and classification of consumption trends of foreigners visiting Japan using a questionnaire survey

Tadanari Taniguchi (IT Education Center, Tokai University, taniguchi@tokai-u.jp)

## Abstract

*This study conducts analysis and classification of consumption trends for foreigners visiting Japan using a questionnaire survey of the Japan Tourism Agency. The agency published a topic analysis of tourist consumption trends. However, the analysis only considered the consumption trends of tourists from each country. This study analyzes and visualizes the relationship between consumption trends, visiting rates to Japanese prefectures, and tourist nationalities. We use the questionnaire survey of 19 countries who visited Japan from 2015 to 2017. The questionnaire results are analyzed using self-organizing maps (SOMs). A SOM can visualize the similarity relations between tourist nationalities and questionnaire items, which is difficult to comprehend with simple or cross tabulations. This study illustrates the relationship between foreign visitors, the visiting prefectures, and the consumption trends and expects the future consumption trends.*

## Keywords

*self-organizing map, questionnaire analysis, tourism consumption trend, clustering, unsupervised learning*

## 1. Introduction

The number of foreigners visiting Japan is increasing every year. Leading up to the Tokyo Olympic Games in 2020, the importance of tourists in Japan from the perspective of economic and international mutual understanding has been continually growing. However, some infrastructural problems exist, such as lack of accommodation facilities and reduction of inbound consumption. The Japan Tourism Agency conducts a consumption trend survey for foreigners visiting Japan and reports the analysis results every year. The reports denoted the relationship between the number of visitors in Japan and the consumption trends of visitors by country. It was very useful, but most analyses only included cross tabulations. However, there were a few analyses that considered the relationship between questionnaire items and tourist nationalities.

Previous studies of tourist consumption have adopted analytical methods [Shaw et al., 2004; Tremblay, 2001; Stone 2005] to examine historical and cultural aspects. To analyze consumption, there are simple tabulation methods [Bruwer and Alant, 2009] and cross tabulation methods [Richards, 1996]. Woodside and Dubelaar [2002] proposed a theory and framework of tourism consumption systems (TCSs).

This study considers not only an analysis of tourist consumption trends but also the relationship between consumption trends, visiting rates of Japanese prefectures, and tourist nationalities. To analyze the relationship between data, many analytical methods can be used, including principal component analysis, factor analysis, multidimensional scaling, k-means, and the self-organizing map (SOM). This study uses SOM as the analysis method. SOM [Kohonen, 1998] is an unsupervised neural network method. It is suitable for analyzing complicated data, and its calculation is not difficult. The SOM method is an

efficient tool for visualizing multidimensional data. Tourism analysis using the SOM method involves, market segmentation approaches [Kuo, 2002; Hung and Tsai, 2008], position analysis [Mazanec, 1995], and image analysis [Mazanec, 1994].

This study uses an SOM to organize data from computer-assisted personal interviewing (CAPI) and paper-and-pencil interviewing (PAPI), regarding a consumption trend survey of foreigners visiting Japan [Japan Tourism Agency, 2015-2017]. The survey was given to tourists from 19 countries during 2015 to 2017. A SOM was used to visualize the similarity relations between consumption trends, visiting rates of Japanese prefectures, and activities of travelers as well as the relationship among the countries.

## 2. Methods

### 2.1 Participants

This study analyzes a questionnaire [Japan Tourism Agency, 2015-2017], which was surveyed by the Japan Tourism Agency. The agency surveyed foreigners who departed from Japan. Note that permanent residents, long-term residents, and transit passengers were excluded. The questionnaire survey was conducted over four periods for each year. Table 1 shows the number of valid participant responses for each questionnaire item. Figure 6 shows the age distribution of respondents in each country. First, second, and third bars show the distribution of women in 2015, 2016, and 2017, respectively. Fourth, fifth, and sixth bars show the distribution of men in 2015, 2016, and 2017,

Table 1: Valid responses from 2015 to 2017

Questionnaire items	2015	2016	2017
Q1	26,068	27,812	29,092
Q2	21,556	23,811	24,892
Q3	20,815	22,154	23,540
Q4	22,080	23,133	24,405
Q5	21,243	22,341	23,808

respectively.

## 2.2 Procedure

A surveyor who could speak foreign languages took questionnaire data using the CAPI and PAPI techniques in departure lobbies of Japanese international airports. The surveyors used a table PC with 12 languages (English, Korean, Simplified and Traditional Chinese, Thai, Indonesian, Vietnamese, German, French, Italian, Spanish, and Russian) and a paper questionnaire written in the 12 foreign languages.

## 2.3 Questionnaire

The Japan Tourism Agency conducted a questionnaire survey for foreigners visiting Japan. We used the following five questionnaire items from 2015 to 2017 in the survey.

- Q1) Where did you go in Japan?
- Q2) Put checkmarks all of the expense items during stay in Japan.
- Q3) What did you want to do before visit to Japan?
- Q4) What did you do during your current stay in Japan?
- Q5) What would you like to do if you were to come to Japan in future?

In the released questionnaire results by the Japan Tourism Agency, the participants have been classified into tourists from 20 different countries and the others. However, this study considered 19 countries, not including India because there were few travelers for sightseeing purposes from India.

## 2.4 Data analysis

SOM [Kohonen, 1998] is an unsupervised neural network method, which maps high-dimensional data into low-dimensional space, preserving the topological structure of the data distribution. It is suitable for analyzing complicated data, and its calculation is not difficult. The SOM is also an efficient tool for visualizing multidimensional data. In this study, the SOM was used to visualize the similarity relations between the participant nationalities and questionnaire items. In this analysis, 47 prefectures (Figure 2), 14 expense items (Figure 3), and 5 activities (Figure 4) were used as SOM variables.

## 3. Results

This study used the SOM-Toolbox [Vesanto, 1999] in MATLAB to make and visualize the SOM for datasets. After the SOM was initialized and trained by principal component analysis, the SOM results were obtained. The number of map units was 600. In this calculation, the final quantization, topographic, and combined errors are 0.330, 0.000, and 0.691, respectively. Figures 1 and 2 show the U-matrix (unified distance matrix) and the clusters. In Fig. 1, the difference of reference vectors (dot sequence) represents the cluster boundary. This figure illustrates the relationship among 19 countries from 2015 to 2017. Figure 3 shows the component planes for 47

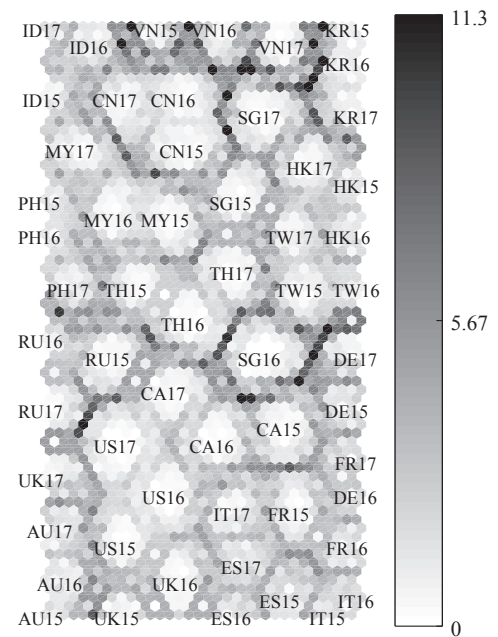


Figure 1: U-matrix

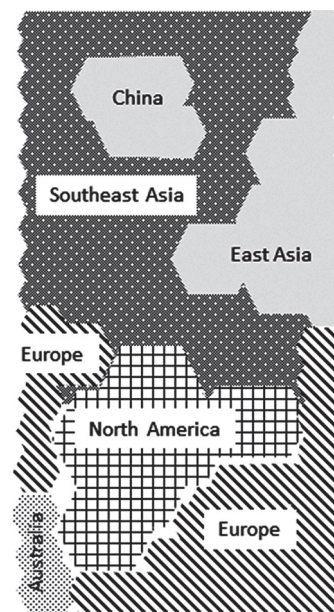


Figure 2: Clusters in U-matrix

prefectures, visited by participants during their stay in Japan. Figure 4 shows the component planes for expense items. Figure 5 shows the component planes for the activities that tourists wanted to do before visiting Japan, did during their trip to Japan, and wanted to do on their next trip to Japan.

Analysis using SOM is suitable for two-dimensional clustering of multidimensional data such as questionnaire results. The results show the relationship between foreign visitor countries, the visiting prefectures, and the consumption trends. The analysis results are not different from the analysis result conventionally mentioned. However it might be difficult to distinguish whether or not the correct result since clustering method is based on unsupervised learning.

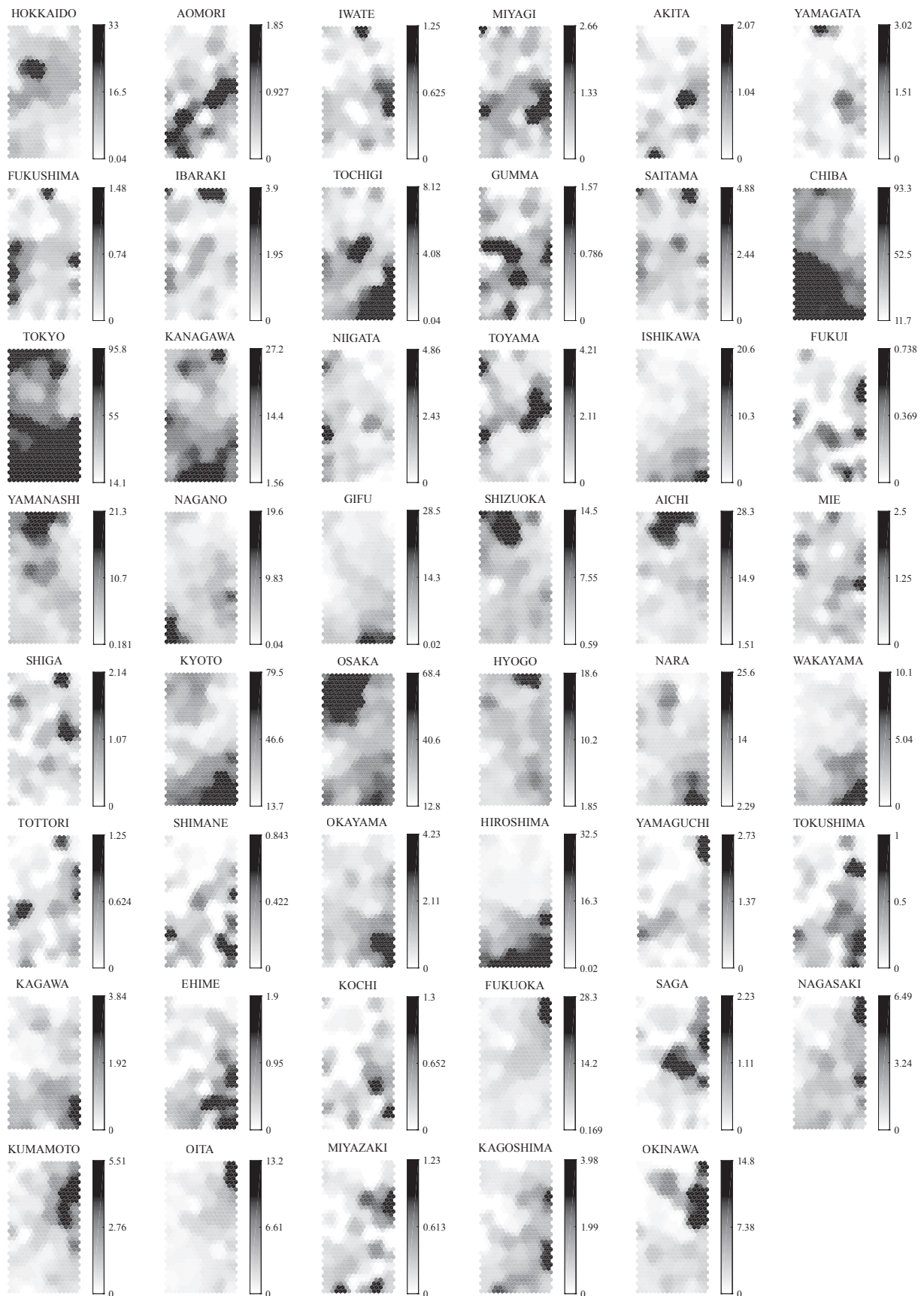


Figure 3: Component planes for 47 prefectures which participants visited during their stay in Japan

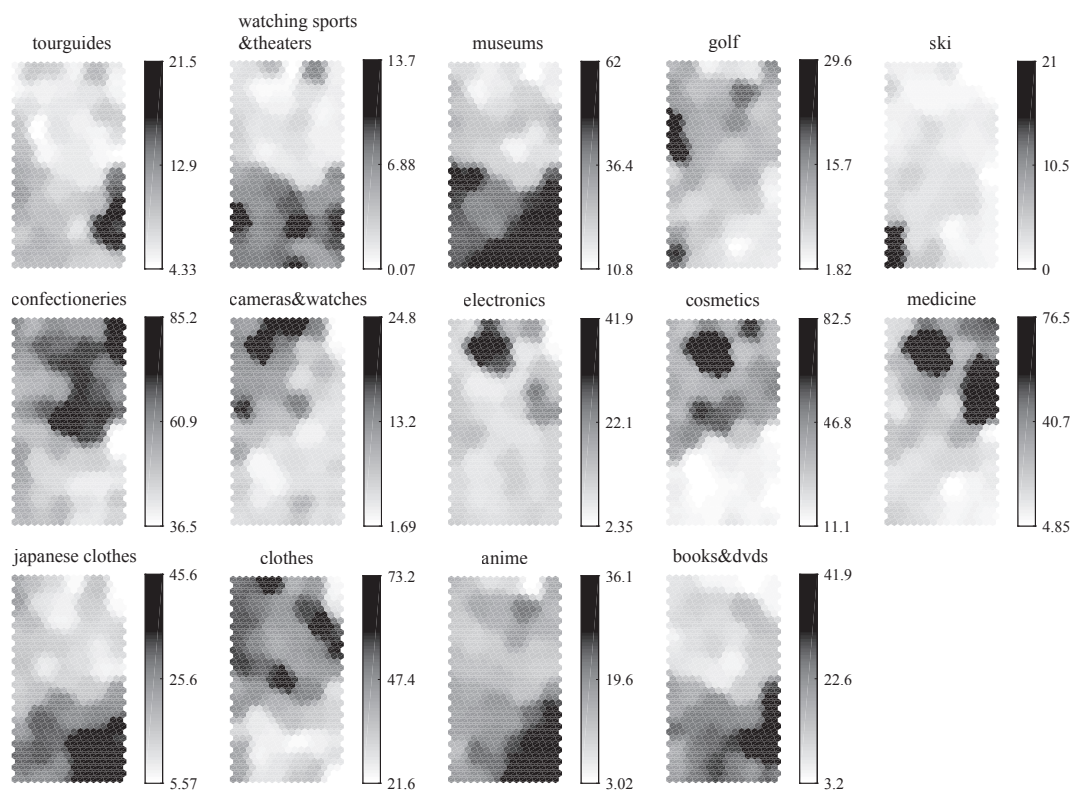


Figure 4: Component planes for expense items during tourist stay in Japan

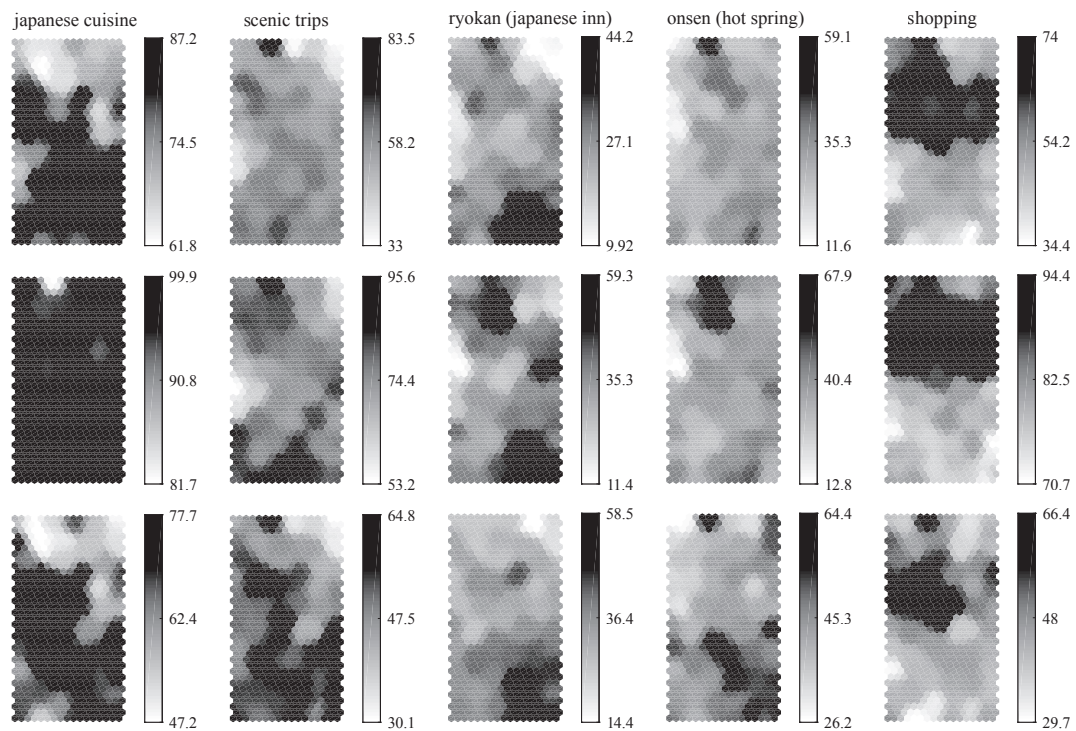


Figure 5: Component planes for activities tourists wanted to do before their visiting to Japan in the upper row, activities during their stay in the middle row, and activities tourists want to do in future in the lower row

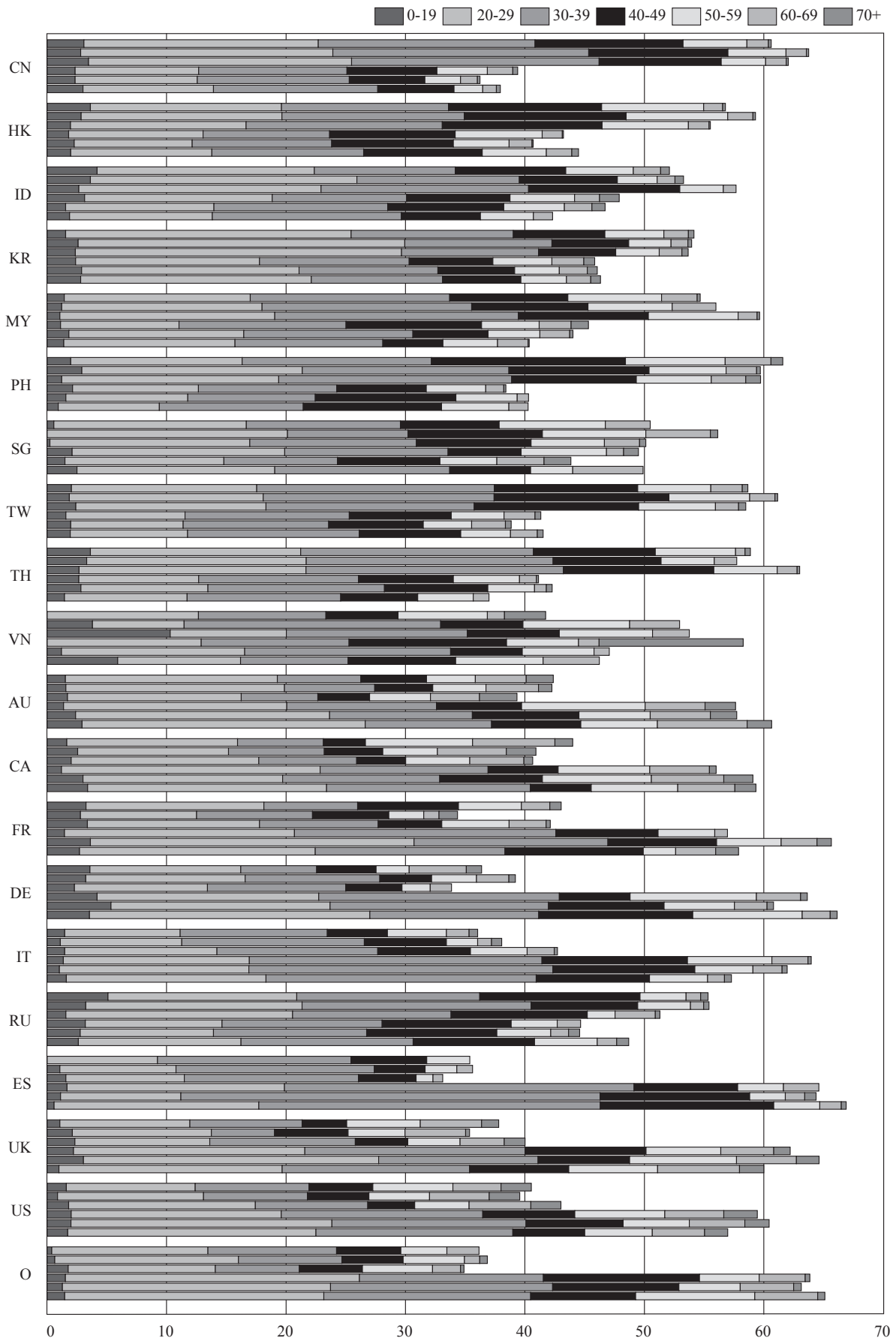


Figure 6: Age distribution of respondents in each country

Notes: First, second, and third bars show the distribution of women in 2015, 2016, and 2017, respectively. Fourth, fifth, and sixth bars show the distribution of men in 2015, 2016, and 2017, respectively. The horizontal axis represents the percentage of the age distribution.

#### 4. Discussion

The results are summarized as follows:

- Despite many tourists coming prefectures with famous sightseeing spots, there were not many visits to the adjacent prefectures.
- European and American tourists liked cultural activities Japanese culture.
- Asian tourists preferred shopping
- Tourists' interests of shopping, Japanese cuisine, and scenic trips were decreasing.
- The interests in accommodation at ryokan and hot springs were increasing.

The visited prefectures, expense trends, and activities of most countries are almost the same over three years (2015, 2016, and 2017), as the three-year positions are very close in the U-matrix. Table 2 shows the legend of symbols used in Figure 1. Asian and South Asian countries are at the top and European and American countries are at the bottom. Some European countries (Germany, France, Italy, and Spain) are at the bottom right. English-speaking countries in the West are at the bottom left. Asian socialist countries (China and Vietnam) are very close at the top center.

Figure 3 shows the component planes for 47 prefectures, which were visited by participants during their stay in Japan. Some prefectures (e.g., Tokyo and Chiba) with large international airports and some prefectures (e.g., Kyoto and Osaka) with famous sightseeing spots have high absolute visiting rates.

Table 2: Legend of symbols used in Figure 1

Country/Region	2015	2016	2017
China	CN15	CN16	CN17
Hong Kong	HK15	HK16	HK17
Indonesia	ID15	ID16	ID17
Korea	KR15	KR16	KR17
Malaysia	MY15	MY16	MY17
Philippines	PH15	PH16	PH17
Singapore	SG15	SG16	SG17
Taiwan	TW15	TW16	TW17
Thailand	TH15	TH16	TH17
Vietnam	VN15	VN16	VN17
Australia	AU15	AU16	AU17
Canada	CA15	CA16	CA17
France	FR15	FR16	FR17
Germany	DE15	DE16	DE17
Italy	IT15	IT16	IT17
Russia	RU15	RU16	RU17
Spain	ES15	ES16	ES17
United Kingdom	UK15	UK16	UK17
United States	US15	US16	US17

Therefore, we consider not the absolute visiting rates but the relative visiting rate. In the component planes black means high visiting rate and white means low visiting rate.

Figure 4 shows the component plane for each consumption item. European and American tourist spent on items related to Japanese culture such as theaters, museums, Japanese clothes, anime, and books. Meanwhile, Asian tourists spent on electric appliances, cosmetics, and confectioneries. French and German tourists tended to ask for tour guides. Filipinos, Singaporeans, and Australians tourists liked to play golf. Australians also enjoyed skiing compared with tourists from other countries.

Figure 5 shows activities that tourists wanted to do before their visiting to Japan, activities during their stay, and activities that tourists wanted to do in a future visit. Europeans were very interested in Japanese cuisine. Most tourists visited sightseeing spots. However, some Asian tourists did not want to visit there again so much compared with European tourists. Staying at a ryokan (Japanese Inn) was preferred by Europeans, however most Asian tourists did not care for it. Chinese tourists took a hot spring (onsen) but did not want to do so the next time. Meanwhile, most Europeans did not take an onsen but wanted to take one. Many Asian tourists went shopping. Tourists from Thailand, Malaysia, and the Philippines wanted to go shopping next time. However, tourists who travel for shopping, including Chinese tourists, decreased year after year.

Despite many tourists coming Kyoto and Osaka, there were not many visits to Hyogo and Nara, which were adjacent. Therefore, it is necessary to guide tourists from famous sightseeing spots to nearby other prefectures and it is insightful to track tourists who traveled the above itinerary. Figure 1 to Figure 4 show that European and American tourists liked cultural activities and were also interested in Japanese culture. Asian tourists tended to prefer shopping, but travelers for shopping continually decreased year after year. As shown in these figures, it is possible to distinguish visited prefectures and consumption trends for each country, facilitating the improvement of service provision methods for each foreigner.

From these results, it is easier to understand the consumption trends by country, the consumption and visiting trends in similar countries, and the future consumption trends compared with simple or cross tabulations. Therefore we can prepare for languages and environmental problems for tourist. This approach is difficult to expect visiting destination trends. In addition, there are limitations on the number of questionnaire items and we might not consider what the tourists need. Questionnaire survey and analysis to investigate the needs of tourist will be necessary in the future.

#### 5. Conclusion

This study has conducted an analysis and classification of consumption trends for foreigners visiting Japan using a questionnaire survey from the Japan Tourism Agency. The agency had published a topic analysis of tourist consumption trends. However, the analysis only considered consumption of tourists

from each country. This study has analyzed the relationship between consumption trends, visiting rates of Japanese prefectures, and tourist nationalities. We have used the questionnaire survey of 19 countries who visited Japan from 2015 to 2017. The questionnaire results have been analyzed using self-organizing maps (SOMs) to visualize the similarity relations of the tourist nationalities and questionnaire items. This study has showed the relationship between foreign visitors, the visiting prefectures, and the consumption trends and expects the future consumption trends. In the future, we will try to suggest visiting destination trends based on consumption trends using a machine learning method or a deep learning method.

an empirical exploration. *Journal of Travel Research*, Vol. 41, 120-132.

(Received October 9, 2018; accepted November 13, 2018)

## References

- Bruwer, J. and Alant, K. (2009). The hedonic nature of wine tourism consumption: An experiential view. *International Journal of Wine*, Vol. 21, No. 3, 235-257.
- Hung, C. and Tsai, C.-F. (2008). Market segmentation based on hierarchical self-organizing map for markets of multimedia on demand. *Expert Systems with Applications*, Vol. 34, 780-787.
- Japan Tourism Agency Homepage (2015-2017). Consumption trend survey for foreigners visiting Japan. <http://www.mlit.go.jp/kankocho/siryoutoukei/syouthityousa.html>.
- Kohonen, T. (1998). The self-organizing map. *Neurocomputing*, Vol. 21, 1-6.
- Kuo, R. J., Ho, L. M., and Hu, C. M. (2002). Integration of self-organizing feature map and k-means algorithm for market segmentation. *Computer & Operations Research*, Vol. 29, 1475-1493.
- Mazanec, J. A. (1995). Positioning analysis with self-organizing maps: An exploratory study on luxury hotels. *The Cornell Hotel and Restaurant Administration Quarterly*, Vol. 36, 80-95.
- Mazanec, J. A. (1994). Image measurement with self-organizing maps: A tentative application to Austrian tour operators. *The Tourist Review*, Vol. 49, 9-18.
- Richards, G. (1996). Production and consumption of European cultural tourism. *Annals of Tourism Research*, Vol. 23, No. 2, 261-283.
- Shaw, S., Bagwell, S., and Karmowska, J. (2004). Ethnoscapes as spectacle: Reimagining multicultural districts as new destinations for leisure and tourism consumption. *Urban Studies*, Vol. 41, No. 10, 1983-2000.
- Stone, P. R., (2005). Dark tourism consumption. *e-Review of Tourism Research*, Vol. 3, No. 5, 109-117.
- Tremblay, P. (2001). Wildlife tourism consumption: Consumptive of non-consumptive? *International Journal of Tourism Research*, Vol. 3, 81-86.
- Vesanto, J., Himberg, J., Alhoniemi, E., and Parhankangas, J. (2000). Self-organizing map in Matlab: The SOM Toolbox. *Proceedings of the Matlab DSP Conference 1999*, 35-40.
- Woodside, A. G. and Dubelaar, C. (2002). A general theory of tourism consumption systems: A conceptual framework and