Application of Google Trends to Forecast Tourism Demand

Kun-Huang Huarng[1], Tiffany Hui-Kuang Yu[2]

[1] Department of Product Innovation and Entrepreneurship, National Taipei University of Business, Taiwan
[2] Department of Public Finance, Feng Chia University, Taiwan
khhuarng@ntub.edu.tw, hkyu@mail.fcu.edu.tw

Abstract

Internet becomes a necessity in our modern lives. The rapid growth of Internet's popularity results in a huge amount of data. Hence, big data analytics is on its berth to handle the data. One interesting research track of the big data literature focuses on "search engine data." The analysis the search engine data is valuable because the business intelligence generated from the analysis offers insights for business opportunities. Google Trends is a popular target for studying search engine data, because it is readily available and easy to access. Because analyzing and forecasting things based on Google Trends can help various domain problems, this study proposes a systematic approach to obtain Google Trends search engine data, to explore usage of the data, and then to provide a forecast. We use Taiwan tourism demand as a study target, where both estimation and forecasting are done by our proposed method. The forecasting results are then compared with real data from the Taiwan Tourism Bureau.

Keywords: Big data analytics, Search engine, Tourism demand

1 Introduction

The Internet is extremely popular all over the world [1], with many business activities having moved from physical stores to various platforms on the web. For example, online shops [2-3], online auction [4], online social media [5], etc. are now major platforms for business transactions and interactions. The development of information technology has generated a massive amount of big data from users, including search queries [6]. For example, big data technologies are used to classify and recommend trips in New York City [7]. For enterprises, big data can be critical for decision making [8]. As people begin to note the importance and value of "big data", one of the interesting research strands of the related literature focuses on "search engine data." Many studies present the value and the possibility of uncovering information from search engine data [9].

Google Trends is a very popular target for studying search engine data, because the data are readily available and easy to access. Google Scholar shows that research on Google Trends has become more popular. First, over 1,000 publications per year have come out in the past three years. Second, the number of publications per year is increasing over the same period.

International tourism is very large industry [10]. International tourism is also a very competitive business [11]. Tourism also plays a significant role in national economies both directly and indirectly by providing employment and earning foreign exchange, explaining why so many small countries rely heavily on tourism. Hence, a need for accurate forecasts of the demand for tourism is widely recognized [12]. The abundant search engine data become the favorable sources for tourism forecasting in the big data era [13].

This study intends to explore the usage of Google Trends to obtain search engine data and then forecast Taiwan tourism demand. The results will be compared with the data from the Tourism Bureau, Taiwan. To that end, section 2 reviews the relevant studies on Google Trends and tourism forecasting. Section 3 describes the data and proposes the research algorithms. Sections 4 compares various empirical forecasting results by the proposed algorithms. Section 5 wraps up this paper.

2 Literature Review

2.1 Google Trends

Google Trends is used in many applications. Believed to be one of the earliest studies to use web-based search data, one study utilizes data on the U.S. unemployment rate [14] while another study uses Google search data to look at how job searches respond to extensions of unemployment payments [15]. On the other hand, Google Trends data are used as a predictor of inflation [16].

Search data from Google Trends are used to measure consumer sentiment [13, 17]. One of the studies find evidence that weekly transaction volumes of S&P 500

*Corresponding Author: Kun-Huang Huarng; E-mail: khhuarng@ntub.edu.tw
DOI: 10.3966/160792642019072004025

companies correlate with weekly search volume of corresponding company names [17]. Both studies see that both the time series of search volume and the time series of transaction volume exhibit recurring patterns.

In addition, one study shows how to use search engine data to forecast automobile sales, unemployment claims, travel destination planning, and consumer confidence [18]. Another study describes the predictability of Google Trends data itself, pointing out that a substantial amount of search terms is highly predictable by using simple seasonal decomposition methods [19].

2.2 Tourism Demand Forecasting

Tourism demand forecasting is a very popular research topic. Many studies probe into the problems of tourism demand that arise in different countries, as these are critical for national economy. Time series models are used to analyze and forecast tourism demand by considering past patterns [20-24]. In addition, forecast combination is a well-established and well-tested approach for improving forecasting accuracy [25-27].

One study forecasts monthly tourism numbers and their aggregate for inbound tourism to Egypt from 33 source countries [28]. They propose a novel combination strategy that produces improved forecasting accuracy. Advanced artificial intelligent tools, such as neural network based fuzzy time series model [29], are used to forecast international tourist arrivals. Neural networks are good at handling non-linear data, and fuzzy time series models outperform conventional time series models in solving time series problems. The numbers of tourist arrivals in Taiwan show a non-linear characteristic with a structural break. Applying a neural network-based fuzzy time series model outperforms some other models in forecasting these tourist numbers. As China is now one of the largest originators of outbound tourists in the world, it is interesting to explore what variables affect tourist arrivals to Taiwan from there and to forecast the corresponding tourism demand. Some other examples include neural networks to select proper models and to forecast tourism demand [30].

Some study uses a neural network based fuzzy time series model to forecast Taiwan's tourism demand. Fuzzy sets are for modeling imprecise data and neural networks are for establishing non-linear relationships among fuzzy sets [11]. That study performs both in-sample estimation and out-of-sample forecasting. Non-linear data are complicated to forecast, and it is even more difficult to forecast non-linear data with shocks. Their model outperforms other studies in forecasting tourism arrivals during the Severe Acute Respiratory Syndrome period of November 2002 to June 2003.

Due to the dynamic relationships among travelers as information searchers, it is necessary to capture the relationships between travelers and search engines to reflect travelers' needs better. A study synthesizes research related to search engine marketing in tourism and present a model that describes the evolving dynamics in such marketing [31]. Another study investigates the usefulness of search query data in forecasting demand for hotel rooms [32]. They use search volume data on five related queries to predict demand for hotel rooms in a specific tourist city and employ different models to evaluate the usefulness of the data. It is claimed that the widespread diffusion of new technologies in communication and business has impacted consumer and product/store knowledge, especially in tourism [33]. Their study conceptualizes strategies and operational marketing policies from an original standpoint, for the purpose of reconstructing and comprehending the dynamics of "integrated" marketing in tourism. They propose an exploratory model involving the main marketing mix levers and recommend the Internet as a "point of synergy" in the "promo-distribution" process of tourism.

3 Research Methods

3.1 Data

To demonstrate the analysis of tourism demand, this study takes tourist arrival data from the Taiwan Tourism Bureau (http://admin.taiwan.net.tw/statistics/year.aspx?no=134) as the research target. In addition, we collect Google Trends data for analysis. Both datasets cover January 2006 to December 2016. We further separate the two into in-sample (2006/01-2014/12) and out-of-sample (2015/01-2016/10) data, which are used for estimation and for forecasting, respectively.

3.2 Estimation Algorithm

Seasonality in tourism arrivals poses enormous challenges and it should be handled and analyzed carefully [34]. One study states that tourism seasonality is the main issue for this topic [35]. Seasonality significantly affects different aspects of the tourism economy and other operations [35-36]. Hence, we consider tourism arrivals as exhibiting seasonal effects.

This study first takes the average of the tourism arrivals of the same month in different years from the in-sample data. Second, we repeat this for the Google Trends data. Thus, we are able to obtain the ratio between the tourism arrival and the Google Trends data of the previous month. The algorithm for estimation is explained below.

Step 1. Calculate the average tourism arrival for each month, *m*, of all the years from the in-sample data of $tour_{y,m}$.

$$T_{m,A} = \frac{\sum_{y=1}^{j} tour_{y,m}}{j-i+1} \quad \textbf{(1)}$$

Here, y represents a specific year; i and j are the starting and ending years of the in-sample data, respectively.

Step 2. Calculate σ_m, or the standard deviation for the tourism of each month m, for all the years from the in-sample data of tourism data, $tour_{y,m}$.

$$\sigma_m = \sqrt{\frac{\sum_{y=1}^{j} (tour_{y,m} - T_{m,A})^2}{j-i+1}} \quad \textbf{(2)}$$

Step 3. Calculate the average for each month m of all the years from the in-sample data of Google Trends, $google_{y,m}$. For each m, the average is:

$$G_{m,A} = \frac{\sum_{y=1}^{j} google_{y,m}}{j-i+1} \quad \textbf{(3)}$$

Step 4. Calculate the ratio between the averages of tourism arrivals and Goggle Trends for each m:

$$R_{m+1,A} = \frac{T_{m+1,A}}{G_{m,A}} \quad \textbf{(4)}$$

Step 5. Use Google Trends in m from the in-sample data (of year y) to forecast tourism arrivals in m+1, $F_{y,m+1}$.

$$F_{y,m+1} = google_{y,m} \times R_{m+1,A} \quad \textbf{(5)}$$

Step 6. Calculate the root mean squared errors (RMSEs) for the estimation. Each forecast is compared with its actual tourism counterpart.

$$RMSE_y = \sqrt{\frac{\sum_{m=1}^{12} (F_{y,m} - tour_{y,m})^2}{12}} \quad \textbf{(6)}$$

3.3 Forecasting Algorithm

The ratios calculated in Estimation Algorithm can be used for forecasting. This study forecasts the tourism arrival of the next month by the ratio of the next month times the Google Trends of the current month. If we have the latest Google Trends, we can always forecast the upcoming month.

We use the in-sample data to obtain the ratios, and the forecasting target is the out-of-sample data. The algorithm for using Google Trends to forecast is explained below.

Repeat Steps 1, 3, and 4 in Estimation Algorithm as Steps 1 to 3 here.

Step 4. Use Google Trends in m from the out-of-sample data (of year y') to forecast tourism arrival in m+1, .

$$F_{y,m+1} = google_{y,m} \times R_{m+1,A} \quad \textbf{(7)}$$

Step 5. Calculate the root mean squared errors (RMSEs) for the forecasts.

$$RMSE_{y'} = \sqrt{\frac{\sum_{m=1}^{12} (F_{y',m} - tour_{y',m})^2}{12}} \quad \textbf{(8)}$$

3.4 Forecasting with a Heuristic Algorithm

The steps for out-of-sample forecasting with extra information are explained below.

Repeat Steps 1 to 4 in Estimation Algorithm.

Step 5. Calculate the differences in tourism arrivals between two consecutive months.

$$d_{y,m} = tour_{y,m} - tour_{y,m-1} \quad \textbf{(9)}$$

Note that y can be the year for the in-sample or out-of-sample data. If m is equal to 1, which means January, then m-1 means December of the previous year. If $d_{y,m}$ is larger than a threshold, then $d_{y,m}$ is considered to be positive (+); if $d_{y,m}$ is smaller than a negative threshold, it is considered to be negative (-); otherwise, $d_{y,m}$ is considered to be neutral (=).

Step 6. Determine the heuristic by observing the differences in the previous three consecutive months. In other words, determine $h_{y',m}$, the heuristic for month m by observing $d_{y',m-3}$, $d_{y',m-2}$, $d_{y',m-1}$, where y' represents the year for the out-of-sample data. Note that m-3, m-2, or m-1 can be smaller than or equal to zero. In this case, we add 12 to either one of them, which means that month of the previous year. If all the three differences are positive, then the heuristic is Rise; if all the differences are negative, then the heuristic is Fall; otherwise, the forecast remains the same.

Step 7. Adjust the forecasts. If the heuristic is Rise, then the forecast (equation (7)) is adjusted as:

$$F_{y',m+1} = google_{y',m} \times R_{m+1,A} + \sigma_m \quad \textbf{(10)}$$

If the heuristic is Fall, then the forecast (equation (7)) is adjusted as:

$$F_{y',m+1} = google_{y',m} \times R_{m+1,A} - \sigma_m \quad \textbf{(11)}$$

Otherwise, the forecast remains the same.

Step 8. Repeat Step 5 of the Forecasting Algorithm for RMSEs.

4 Empirical Analysis

We now present the estimation, forecasting, and forecasting with heuristic and compare the overall performances of these three algorithms.

4.1 Estimation

Step 1. Calculate average tourism arrivals. For example, the average tourism arrival for February is:

$$T_{2,A} = 263,643.00$$

Table 1 lists the average tourism arrivals of all the months.

Step 2. Calculate the standard deviation of tourism arrival. For example, the standard deviation for February is:

$$\sigma_2 = 52,930.33$$

Table 2 lists the standard deviations of tourism arrivals all the months.

Step 3. Calculate the average Google Trends of each month. For example, the average Google Trends for January is:

$$G_{1,A} = 42.78$$

Table 3 lists the average Google Trends of all the months.

Step 4. Calculate the ratios of tourism arrival and Google Trends. For example, the ratios for February are as follows:

$$R_{2,A} = \frac{T_{2,A}}{G_{1,A}} = 6,163.08$$

Table 4 lists all the ratios.

Step 5. Multiply the in-sample Google Trends data by the ratio to estimate the arrival for a certain month. For example, the estimate for February 2015 is:

$$\begin{aligned} F_{2014,2} &= google_{2014,1} \times R_{2,A} \\ &= 56 \times 6,163.08 \\ &= 345,132.48 \end{aligned}$$

Step 6. Calculate the RMSEs for various in-sample years. For example:

RMSE (2014) = 31052.28.

4.2 Forecasting

Repeat Steps 1, 3, and 4 in Estimation Algorithm as Steps 1 to 3 here.

Step 4. Use out-of-sample Google Trends to calculate the arrival forecast. For example, the February 2016 forecast is:

$$\begin{aligned} F_{2016,2} &= google_{2016,1} \times R_{2,A} \\ &= 462,231.23 \end{aligned}$$

Table 5 lists all the forecasts and their corresponding tourism data for 2015 and 2016, respectively. We depict the forecasts of 2015 and 2016 with their counterparts in Figure 1 and Figure 2, respectively.

Table 1. Average tourism for all the months

Month	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
Average	251,654.56	263,643.00	320,772.78	279,524.00	270,927.89	277,234.00	253,631.11	274,561.78	274,906.22	302,547.67	325,064.67	351,704.33

Table 2. Standard deviation of tourism arrivals for all the months

Month	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
Average	42,250.73	52,930.33	62,129.00	51,153.80	54,985.45	46,360.66	43,398.96	53,227.09	50,704.38	60,824.81	69,465.29	82,835.67

Table 3. Average Google Trends for all the months

Month	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
Average	42.78	43.67	45.44	45.56	45.78	46.00	45.78	46.33	45.00	45.11	45.33	47.56

Table 4. Ratios of tourism and Google Trends

Month	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
Average	5,291.80	6,163.08	7,345.94	6,150.89	5,947.20	6,056.08	5,513.72	5,997.71	5,933.23	6,723.28	7,205.87	7,758.18

Table 5. Forecasts

2015	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
Tourism	366,321.00	325,447.00	478,018.00	396,161.00	380,212.00	363,657.00	349,575.00	377,133.00	384,661.00	437,939.00	475,917.00	548,006.00
Forecast	333,383.49	375,948.07	433,410.55	424,411.75	404,409.43	417,869.69	385,960.39	431,835.03	444,992.09	477,352.99	518,822.42	574,105.60

2016	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
Tourism	415,503.00	410,711.00	519,779.00	433,110.00	440,856.00	414,718.00	409,773.00	462,919.00	449,140.00	517,136.00	574,545.00	654,830.00
Forecast	386,301.50	462,231.23	521,561.84	473,618.90	440,092.62	448,150.11	419,042.71	491,812.12	474,658.23	531,139.24	562,057.63	612,896.52

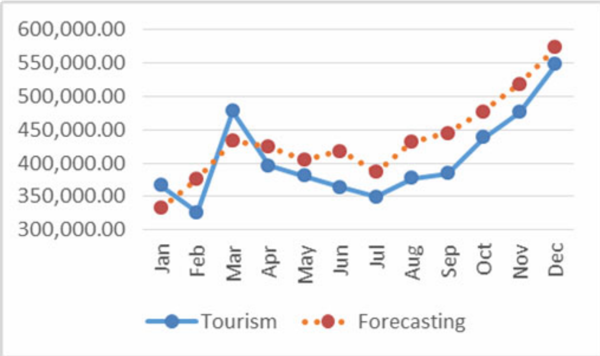

Figure 1. Forecasts of year 2015

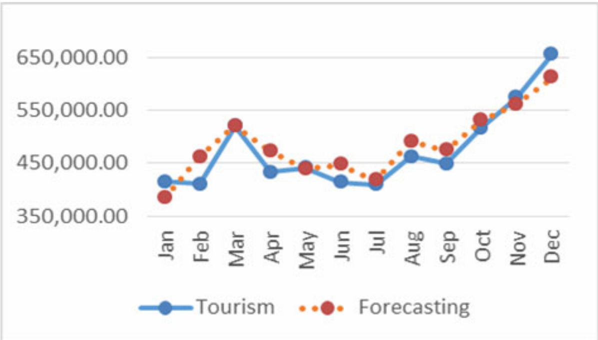

Figure 2. Forecasts of year 2016

Step 5. Calculate the root mean squared errors (RMSEs) for the forecasts. The RMSEs of years 2015 and 2016 are calculated below.

RMSE (2015) = 42,797.27
RMSE (2016) = 28,797.95

4.3 Forecasting with Heuristics

Repeat Steps 1 to 4 in Estimation Algorithm here.
Step 5. Calculate the difference between two consecutive tourism arrivals. This study also sets the threshold to be 10,000. To illustrate the process, one example is:

$$\begin{aligned} d_{2015,7} &= tour_{2015,7} - tour_{2015,6} \\ &= 349{,}575 - 363{,}657 \\ &= -14{,}082 \end{aligned}$$

Because the difference is smaller than the negative threshold, it is considered negative (-).

$$\begin{aligned} d_{2015,8} &= tour_{2015,8} - tour_{2015,7} \\ &= 377{,}133 - 349{,}575 \\ &= 27{,}558 \end{aligned}$$

The difference is larger than the threshold. Hence, it is considered positive (+).

$$\begin{aligned} d_{2015,9} &= tour_{2015,9} - tour_{2015,8} \\ &= 384{,}661 - 377{,}133 \\ &= 7{,}528 \end{aligned}$$

The difference is smaller than the threshold. Thus, it is considered neutral (=).
Table 6 lists all the differences.

Table 6. Differences and heuristics

2014	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
Tourism									371661	415007	437940	486401
d										43346	22933	48461
										+	+	+

2015	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
Tourism	366321	325447	478018	396161	380212	363657	349575	377133	384661	437939	475917	548006
d	-120080	-40874	152571	-81857	-15949	-16555	-14082	27558	7528	53278	37978	72089
Sign	-	-	+	-	-	-	-	+	=	+	+	+
Heuristic	Rise	0	0	0	0	0	Fall	Fall	0	0	0	0

2016	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
Tourism	415503	410711	519779	433110	440856	414718	409773	462919	449140	517136	574545	654830
d	-132503	-4792	109068	-86669	7746	-26138	-4945	53146	-13779	67996	57409	80285
Sign	-	=	+	-	=	-	=	+	-	+	+	+
Heuristic	Rise	0	0	0	0	0	0	0	0	0	0	0

Step 6. Determine the heuristic by consecutive differences. For example, the differences of October, November, and December in 2014 are all positive (+); hence, January 2015 is Rise. The differences of April, May, and June of 2015 are all negative (-). Hence, July 2015 is Fall.
Table 6 lists all the heuristics.

Step 7. Adjust the forecasts. Following the above examples, for January 2015, there is Rise as the heuristic. Hence, the forecast becomes:

$$\begin{aligned} F_{2015,1} &= google_{2015,1} \times R_{1,A} + \sigma_1 \\ &= 333{,}383.49 + 42{,}250.73 \\ &= 375{,}634.22 \end{aligned}$$

For July 2015, there is Fall as the heuristic. Hence, the forecast is adjusted as:

$$F_{2015,7} = google_{2015,7} \times R_{7,A} + \sigma_7$$
$$= 385,960.38 - 43,398.96$$
$$= 342,561.43$$

Table 7 lists all the adjusted forecasts in bold.

Step 8. Calculate the root mean squared errors (RMSEs) for the adjusted forecasts. The RMSEs of year 2015 and 2016 are:

RMSE (2015) = 37,323.24
RMSE (2016) = 27,793.01

Table 7. Forecasts with heuristics

2015	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
Forecast	**375,634.22**	375,948.07	433,410.55	424,411.75	404,409.43	417,869.69	**342,561.43**	**378,607.95**	444,992.09	477,352.99	518,822.42	574,105.60

2016	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
Forecast	**428,552.23**	462,231.23	521,561.84	473,618.90	440,092.62	448,150.11	419,042.71	491,812.12	474,658.23	531,139.24	562,057.63	612,896.52

5 Discussion

To use Google Trends to conduct tourism demand forecasting, several issues need to be handled, including search engine, language, keyword, etc. This study conducted a survey with two groups of students: one is 60 international students at the 2016 Hanyang International Summer School, Hanyang University, Seoul, South Korea; the other group is 22 international students from Ton Duc Thang University, Ho Chi Minh City, Vietnam visiting Feng Chia University, Taichung, Taiwan for a one-semester English program in 2017. The survey investigates what the students were planning to do when going abroad to visit Taiwan. This survey aims to resolve some of these issues of this study.

First, what is the search engine that most tourists use? Should Google be not the most popular search engine, then the volume collected by Google Trends may not be proper to reflect tourism demand. According to the survey, 77 of 82 students used Google as the search engine, which is 94%. Thus, it is proper to use Google Trends for tourism demand forecasting.

Second, what is the language used by tourists when they browse the Internet? If we intend to investigate a domestic topic, then the local language can be used. However, when dealing with an international issue, it may not be very straightforward to determine the language used by different nationalities. If the language cannot match what most tourists use, then the data volume obtained from Google Trends can be easily twisted. In this case, forecasting performance is greatly affected. According to the survey, 100% of international students used English for their search purposes. Hence, for tourism demand forecasting, this study uses English as the language for Google Trends to obtain the tourism data.

Third, what is the proper keyword for Google Trends to gather the relevant data for tourism demand? The keyword must be carefully chosen to represent the target subject. When asked what keywords tourists entered when planning to visit Taiwan, according to the survey, 52 of 82 students (63%) replied “Taiwan” as one of the keywords. Hence, this study chooses “Taiwan” as the keyword for Google Trends.

The next issue, which is also very difficult to resolve, is how much time ahead we should reserve for time series forecasting. For tourism demand forecasting, we need to determine how many days ahead tourists are likely to start their search for tourism information. This study takes one month ahead for tourism demand forecasting.

There are some other issues. It is very natural for a tourist to search for tourism information more than once for a single upcoming trip. If the occurrence is high, then the Google Trends data tend to be greater than real tourism demand.

Future studies may focus on some of the above issues for more concrete discussions and analyses. By doing so, forecasting performance should improve in accuracy.

6 Conclusion

It is widely acknowledged that the Internet is now extremely popular and common all over the world. With people using Internet applications in their daily lives, the Internet has generated a huge amount of data. To understand what people are interested on the Internet, one can use mega-data forecasting as a research tool.

Google is one of the top search engines. As a result, Google Trends is a good representation of certain consumer behaviors or Internet trends. The advantages of using Google Trends are that the data are easily accessible and ready for use. Hence, a systematic approach to analyze Google Trends and to conduct forecasts can be rather valuable. This study proposes algorithms to use Google Trends for forecasting tourism demand. To demonstrate our approach, we utilize tourism data from the Taiwan Tourism Bureau to test the forecasts.

Three algorithms serve various purposes, such as estimation, forecasting, and forecasting with a heuristic. With more information, the forecasting approach with a heuristic improves overall model performance. As issues still remain for further improving the forecasting technique by using Google Trends, future studies may endeavor to solve them.

Acknowledgments

The authors acknowledge financial support from the Ministry of Science and Technology, Taiwan, R.O.C. under grant MOST 105-2410-H-035 -029 -MY2.

References

[1] K.-H. Huarng, T. H.-K. Yu, Internet Software and Services: Past and Future, *The Service Industries Journal*, Vol. 31, No. 1, pp. 79-89, January, 2011.

[2] K.-H. Huarng, T. H.-K. Yu, H.-Y. Cheng, A Study of Online Auction in YAHOO! TAIWAN, *International Journal of Biomedical Soft Computing and Human Sciences*, Vol. 13, No. 1, pp. 27-31, January, 2008.

[3] K.-H. Huarng, T. H.-K. Yu, J. J. Huang, The Impacts of Instructional Video Advertising on Customers' Purchasing Intentions on the Internet, *Service Business*, Vol. 4, No. 1, pp. 27-36, March, 2010.

[4] C. Wu, K.-H. Huarng, S. Fiegantara, P. C. Wu, The Impact of Customer Online Satisfaction on Yahoo Auction in Taiwan, *Service Business*, Vol. 6, No. 4, pp. 473-487, December, 2012.

[5] V. D. D. Gunawan, K.-H. Huarng, Viral Effects of Social Network and Media on Consumers' Purchase Intention, *Journal of Business Research*, Vol. 68, No. 11, pp. 2237-2241, November, 2015.

[6] V. Mayer-Schonberger, K. Cukier, *Big Data: A Revolution that Will Transform how We Live, Work, and Think*, Eamon Dolan/Houghton Mifflin Harcourt, 2013.

[7] H. Sun, S. Hu, S. McIntosh, Y. Cao, Big Data Trip Classification on the New York City Taxi and Uber Sensor Network, *Journal of Internet Technology*, Vol. 19, No. 2, pp. 591-598, March, 2018.

[8] B. J. Chang, Agile Business Intelligence: Combining Big Data and Business Intelligence to Responsive Decision Model, *Journal of Internet Technology*, Vol. 19, No. 6, pp. 1699-1706, November, 2018.

[9] L. Vaughan, Y. Chen, Data Mining from Web Search Queries: a Comparison of Google Trends and Baidu Index, *Journal of the Association for Information Science and Technology*, Vol. 66, No. 1, pp. 13-22, January, 2014.

[10] C. Tisdell, Introduction: Basic Economics of Tourism: An Overview Mainly of Volume I, in: C. Tisdell (Ed.), *The Economics of Tourism*, Edward Elgar, 2000, pp. xi-xxxix.

[11] K.-H. Huarng, T. H.-K. Yu, L. Moutinho, Y.-C. Wang, Forecasting Tourism Demand by Fuzzy Time Series Models, *International Journal of Culture, Tourism, and Hospitality Research*, Vol. 6, No. 4, pp. 377-388, December, 2012.

[12] S. F. Witt, C. A. Witt, Forecasting Tourism Demand: A Review of Empirical Research, *International Journal of Forecasting*, Vol. 11, No. 3, pp. 447-475, September, 1995.

[13] X. Li, B. Pan, R. Law, X. Huang, Forecasting Tourism Demand with Composite Search Index, *Tourism Management*, Vol. 59, pp. 57-66, April, 2017.

[14] M. Ettredge, J. Gerdes, G. Karuga, Using Web-based Search Data to Predict Macroeconomic Statistics, *Communications of the ACM*, Vol. 48, No. 11, pp. 87-92, November, 2005.

[15] S. Baker, A. Fradkin, *What Drives Job Search? Evidence from Google Search Data*, http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1811247.

[16] G. Guzman, Internet Search Behavior As an Economic Forecasting Tool: The Case of Inflation Expectations, *The Journal of Economic and Social Measurement*, Vol. 36, No. 3, pp. 119-167, December, 2011.

[17] T. Preis, D. Reith, H. E. Stanley, Complex Dynamics of Our Economic Life on Different Scales: Insights from Search Engine Query Data, *Philosophical Transactions of the Royal Society A: Mathematical, Physical & Engineering*, Vol. 368, No. 1933, pp. 5707-5719, December, 2010.

[18] H. Choi, H. Varian, Predicting the Present with Google Trends, *The Economic Record*, Vol. 88, No. 280, pp. 2-9, March, 2012.

[19] Y. Shimshoni, N. Efron, Y. Matias, *On the Predictability of Search Trends*, Google Research Blog, January, 2009.

[20] U. Gunter, I. Önder, Forecasting International City Tourism Demand for Paris: Accuracy of Uni- and Multivariate Models Employing Monthly Data, *Tourism Management*, Vol. 46, 123-135, February, 2015.

[21] A. Guizzardi, A. Stacchini, Real-time Forecasting Regional Tourism with Business Sentiment Surveys, *Tourism Management*, Vol. 47, pp. 213-223, April, 2015.

[22] M. Akın, A Novel Approach to Model Selection in Tourism Demand Modeling, *Tourism Management*, Vol. 48, pp. 64-72, June, 2015.

[23] F. L. Chu, Forecasting Tourism Demand with ARMA-based Methods, *Tourism Management*, Vol. 30, No. 5, pp. 740-751, October, 2009.

[24] G. Athanasopoulos, R. J. Hyndman, Modelling and Forecasting Australian Domestic Tourism, *Tourism Management*, Vol. 29, No. 1, pp. 19-31, February, 2008.

[25] H. Song, S. F. Witt, K. F. Wong, D. C. Wu, An Empirical Study of Forecast Combination in Tourism, *Journal of Hospitality & Tourism Research*, Vol. 33, No. 1, pp. 3-29, February, 2009.

[26] C. H. Aladag, E. Egrioglu, U. Yolcu, Forecast Combination by Using Artificial Neural Networks, *Neural Processing Letters*, Vol. 32, No. 3, pp. 269-276, December, 2010.

[27] K. Lahiri, L. Yang, A Non-Linear Forecast Combination Procedure for Binary Outcomes, *Studies in Nonlinear Dynamics & Econometrics*, Vol. 20, No. 4, pp. 421-440, September, 2016.

[28] R. R. Andrawis, A. F., Atiya, H. El-Shishiny, Combination of Long Term and Short Term Forecasts, with Application to Tourism Demand Forecasting, *International Journal of Forecasting*, Vol. 27, No. 3, pp. 870-886, July-September,

2011.
[29] K.-H. Huarng, L. Moutinho, T. H.-K. Yu, An Advanced Approach to Forecasting Tourism Demand in Taiwan, *Journal of Travel and Tourism Marketing*, Vol. 21, No. 4, pp.15-24, October, 2007.
[30] L. Moutinho, K.-H. Huarng, T. H.-K. Yu, C.-Y. Chen, Modeling and Forecasting Tourism Demand: The Case of Flows from Mainland China to Taiwan, *Service Business*, Vol. 2, No. 3, pp. 219-232, September, 2008.
[31] B. Pan, Z. Xiang, R. Law, D. R. Fesenmaier, The Dynamics of Search Engine Marketing for Tourist Destinations, *Journal of Travel Research*, Vol. 50, No. 4, pp. 365-377, July, 2011.
[32] B. Pan, D. C. Wu, H. Song, Forecasting Hotel Room Demand Using Search Engine Data, *Journal of Hospitality and Tourism Technology*, Vol. 3, No. 3, pp. 196–210, December, 2012.
[33] P. Singer, M. A. Ferri, L. Aiello, C. Cacia, Internet As a Point of Synergy Between Communication and Distribution: Hypothesis of Model Applied to Tourism, *International Journal of Digital Content Technology and its Applications*, Vol. 4, No. 7, pp. 23-37, October, 2010.
[34] C. Goh, Exploring Impact of Climate on Tourism Demand, *Annals of Tourism Research*, Vol. 39, No. 4, pp. 1859-1883, October, 2012.
[35] T. Baum, S. Lundtorp, *Seasonality in Tourism*, Pergamon, 2001.
[36] N. Koenig-Lewis, E. E. Bischoff, Seasonality Research: The State of The Art, *International Journal of Tourism Research*, Vol. 7, No. 4/5, pp. 201-219, July-October, 2005.

Biographies

Kun-Huang Huarng received his PhD from Texas A&M University, USA. He is a Professor of Product Innovation and Entrepreneurship at National Taipei University of Business, Taiwan, where he also serves as the Vice President. He is the Editor-in-Chief of International Journal of Business and Economics.

Tiffany Hui-Kuang Yu received her doctoral degree from Economics, Texas A&M University, USA. She just completed the terms as the director of the Science and Technology Division at Taipei Economic and Cultural Office in Houston, USA. She is currently a Professor of Public Finance, Feng Chia University, Taiwan.