教育資料與圖書館學

Journal of Educational Media & Library Sciences

http://joemls.tku.edu.tw

Vol. 54 , no. 2 (2017) : 161-184

台灣出版產業銷售分析：資料庫知識發現觀點

A Revenue Analysis on Taiwan's Publishing Industries from the Prospective of Knowledge Discovery Using Government's Financial Database

徐 明 珠* Ming-Ju Hsu*

Associate Professor

E-mail：mjhsu@mail.fgu.edu.tw

Journal of Educational Media & Library Sciences 54: 2 (2017) : 161-184
DOI:10.6120/JoEMLS.2017.542/0054.RS.BE

Research Article

A Revenue Analysis on Taiwan's Publishing Industries from the Prospective of Knowledge Discovery Using Government's Financial Database

Ming-Ju Hsu*

Abstract

This research focused on retrieving and analyzing data and information from the "Financial Database" established by the Ministry of Finance of Republic of China (Taiwan), carrying out a Knowledge Discovery from the Financial Database (KDFD) 2013~2015, primarily on Taiwan's Nine(9)-subclass Publishing Industries. The results of the research showed that: a) the sales revenues of Taiwan's Publishing Industries have declined year after year from 2013 to 2015; b) within the year there was a wave of steep drop in sales from May to June, then the sales revenues gradually recovered and reached the peak in November to December and c) Newspapers, magazines and books publishing were still the dominate part of the sales for the Publishing Industries (82.1%). While the Digital Publishing Industries accounted for an average of 16% of total sales revenues from 2013 to 2015, the growth spurs from 11% in 2013 to 16% in 2014, then to 20% in 2015 were quite impressive, indicating a potential growth for Digital Publishing Industries. The definition of the publishing industries categorized by the government of Taiwan included nine subclasses in the whole division, further study could be conducted for each subclass of the publishing division to obtain its actual sales revenues for a more realistic comparison with surveyed data.

Keywords: *Publishing Industry, Digital publishing, Data mining, Knowledge discovery in database, Decision tree analysis, Sales revenue*

Introduction

Research background

Based on Wischenbart & Krenn's Publishing Industry revenue statistical views in 2014,[1] the accounting and statistics of sales revenues on Taiwan's Publishing Industry faced several layers of complexities due to a lack of standard or regulation; even simple parameters such as what types of publishing industries should be included, or which sectors could be excluded, what kind of statistics should be collected for comparison, what were the definitions of Consumer Price at Market Value or Total Production Output Value, etc., all were confusing and lacked of clarifications.

* Associate Professor, Department of Communication, Fo Guang University, Yilan, Taiwan
E-mail: mjhsu@mail.fgu.edu.tw

2016/12/20 received; 2017/04/12 revised; 2017/05/23 accepted

The official surveys on Taiwan's Publishing Industry initially were conducted by the Council for Cultural Affairs, Executive Yuan for nearly the last two decades; the survey responsibility was then transferred to Government Information Office, and then further taken over by the Ministry of Culture due to repeated administrative reconstructions. Publishing Industry Surveys were mainly focused on book publishing, while magazine publishing surveys were scarcely conducted. The targets of the surveys had always been focused on the "Class 5811-5820" of the Standard Industrial Classification System. After the recent adjustment of the supervision responsibility for the Publishing Industry in Taiwan, three used-to-be independent surveys for the book, magazine and digital publishing, respectively, were integrated into one comprehensive survey for synergy and efficiency in 2015. Moreover, the survey frequency has been changed from once every other year to annual, to better grasp the context and fast changing trends of the Publishing Industry.[2] However, the survey related investigations and preparations as well as the subsequent data analysis for Taiwan's Publishing Industry were very time consuming, labor-intensive and expensive; the numerous data were somewhat scattered and hard to theorize for trend or pattern determinations. The resulted analysis tended to be misleading and shown significant gaps between expected and supplied information.[3] In view of these problems, this research thus collected the actual Publishing Industry's sales revenues data to analyze its overall market performances and future development trends. Hopefully the study results and conclusions could provide a meaningful guideline for the government setting better policies and Publishing Industries designing more productive, forward-looking marketing strategies.

Research purposes

Sales revenue is the lifeline of any industry's development; sales revenues data also reveals the market strengths or weaknesses of a given industry. The purposes of this research study were: a) Use Literature Review to collect the official survey reports on publishing industries' sales revenues, comparing them with the actual sales data from the publishing industries; b) Use Data Mining and Decision Tree analysis to uncover any hidden patterns, trends and relationships on Publishing Industries' sales revenues versus parameters such as fiscal years, bi-monthly period and types of enterprises from the Financial Database.

Standard Industrial Classification and Sales Survey for Taiwan's Publishing Industry

Category and classification of Taiwan's publishing industry

In January of 1967, the government in Taiwan announced the "Standard Industrial Classification", the Publishing Industry was one of the Division

in the Standard Industrial Classification, under Section J "Information and Communication". In the last fifty years, the classifications of "Publishing industries" have gone through ten times of amendments. In the latest version released in January 2016, the Publishing Industry in conjunction with "Audio/Video Production, Dissemination and ICT Services" were jointly subordinated into "Section J",[4] becoming a more important and significant section by themselves.[5]

Based on the "Standard Industrial Classification", Publishing Industries were categorized into "Division 58", which included Newspapers, Magazines, Journals/Periodicals, Book, Software and Other Types of Publishing Industries, etc. "Division 58" was further divided into "Group 581" and "Group 582" two subcategories, respectively. "Group 581" included Newspapers, Magazines, Journals and Periodicals, Books and Other Publishing Activities, while "Group 582" included the Software Publishing Industries, which further subdivided into "Game Software Publishing" and "Other Software Publishing". Table 1 listed a detail classification for Taiwan's Publishing Industries.

Table 1 Standard Industrial Classification for Taiwan's Publishing Industries

Division 58 publishing activities	Group	Class	Subclass	Publishing type (a sub-category classified)	Description	Form of publication
58	581	5811	5811-11	Publishing of newspapers	This subclass includes the activities of publishing of printed newspapers, including advertising newspapers.	Printed Form
58	581	5811	5811-12	Digital news publishing	This subclass includes the activities of publishing of newspapers, including advertising newspapers in electronic forms, including on the Internet.	Digital Form
58	581	5812	5812-11	Publishing of magazines and periodicals	This subclass includes the activities of publishing of periodicals and other journals, which can be done in print.	Printed Form
58	581	5812	5812-12	Digital publishing of magazines and periodicals	This subclass includes the activities of publishing of periodicals and other journals in electronic forms, including on the Internet.	Digital Form
58	581	5813	5813-11	Book publishing	This subclass includes the activities of publishing books in print (Reference books, dictionaries, encyclopedias, atlases, maps and technical manuals).	Printed Form

58	581	5813	5813-12	Digital book publishing	This subclass includes the activities of audio books or web publishing books and other forms (Reference books, dictionaries, encyclopedias, atlases, maps and technical manuals).	Digital Form
58	581	5819	5819-00	Other publishing activities	This subclass includes the activities of catalog, brochures, photos, postcards, greeting cards, reproductions of works of art, telephone books, advertising material, directory and mailing list.	Printed Form
58	582	5820	5820-11	Game software publishing	This subclass includes the activities of publishing of computer games for all platforms.	Digital Form
58	582	5820	5820-99	Other software publishing	This subclass includes the activities of publishing of ready-made (non-customized) software: operating systems, business and other applications.	Digital Form

Source: Directorate-General of Budget, Accounting and Statistics, *Standard Industrial Classification System of the Republic of China (Rev.10)* (Taipei: Executive Yuan, 2016), 9-10, 98-99.

Sales survey for Taiwan's Publishing Industries

There were three key elements in conducting a survey for Taiwan's Publishing Industry, namely the total volume of publications, the total number of publishing houses and the total production of publishing output. These three elements were essential in reflecting the strength and magnitude of the publishing industry of any country.[6] Note the term "Total Production of Publishing Output" was not clearly defined and needed to be further specified; in the meantime, the most important Publishing Output value of printed book publishing in Taiwan was its actual total sales revenues.[7]

Taiwan's survey study on Publishing Industries went through three different periods in the last 20 years. The Council for Cultural Affairs started to issue a "Market Survey for Taiwan's Book Publishing" in 1998, this annually issued survey ended in 2001. Then the Government Information Office took over the governing responsibility and conducted surveys for Book and Magazine Publishing Industries, respectively (from 2003 to 2011). In 2012, the Ministry of Culture further took over the survey responsibilities from the Government Information Office and in 2015 the surveys for Book, Magazine and Other Types of Publishing Industries were integrated into one consolidated survey.

There were several intertwined factors affecting the accuracy of the survey results on Taiwan's Publishing Industry in the past: a) A lack of objective, competent and reputable survey institutions in Taiwan; b) Many publishing companies treated some business data as vital trade secret and either were reluctant to provide, or giving plausible data, making the year-by-year comparison difficult and unreliable[8]; c) Survey agencies could have withheld some data due to unspecified reasons, the transparency and credibility of the survey became questionable[9]; d) Governmentally published Production Output Values were estimations only[10]; and e) All other factors such as the gaps between the non-book type sales, number of copies printed, quantity of books ordered versus the actual number of sales achieved. Moreover, other complexities such as various discounts through different sales channels, book return rates, damages and loss, unfavorable payment terms, profit sharing, delinquent accounts and revenue turnover issues, etc., these factors and complexities could have made accurate surveys and reliable total sales revenue of Publishing Industry in Taiwan very difficult.[11]

Publishing revenue includes publishers' net revenue (from book sales, exports, rights and licenses for print and digital, including educational and professional publishing) and market value at consumer/retail prices. The difficulty in gathering different data sources and gaps between reality and statistics may cause the accuracy of the data.[12] The total sales revenue of Taiwan's Publishing Industry was a critical indicator of the competency and overall operation efficiency of the industry. Take the book publishing industry as an example, its annual sales revenue was a symbol of the industry's strength and magnitude, its year-by-year sales revenue comparison could indicate the industry's growth potential or decline concern, etc.[13] Surveys on Taiwan's Publishing Industry have been focused on books and magazines revenues for a long time. In recent years, the Ministry of Culture (MOC) started to promote the growth and development of the Digital Content Industry and thus began to issue Output Values (revenues) surveys on Digital Publishing Industry and e-Magazine Industry. The Ministry of Economic Affairs also published some revenue data on the Digital Content Industry, see Table 2.

The total book sales calculated by MOC was based on the total sale volumes (survey result from publishing entities including new, reprint, and revised sale volumes) multiplied by the average price of each category. The unaccounted sales value was an average number from those publishing entities who did specify such values. Therefore, MOC published total book sales could deviate from the actual book sales significantly.

The Digital Content Industry were divided into eight different divisions by the Industrial Development Bureau of Ministry of Economic Affairs, three

Table 2 Survey Report of the Output Value (Revenue) on Taiwan's Publishing Industries from 2013 to 2014 by Governmental Agency

	2013 NT$ billion	2014 NT$ billion	Source
Magazine Publishing Industry	16.5	15.8	Ministry of Culture
Digital Magazine Publishing Industry	0.13	0.28	Ministry of Culture
Book Publishing Industry	28.1	24.3	Ministry of Culture
Digital Book Publishing Industry	0.51	0.53	Ministry of Culture
Digital Game Industry	45.32	50.6	The Industrial Development Bureau of MOEA
Content Software Industry	202.5	210	The Industrial Development Bureau of MOEA
Digital Publishing and Archive (Digital Magazines, Books etc., Digital Archive)	52.8	49.2	The Industrial Development Bureau of MOEA

Source: Ministry of Culture, *Survey of Taiwan's Publishing Industry for 2013-2014* (Taipei: Ministry of Culture, 2015), II-VII; Industrial Development Bureau, Ministry of Economic Affairs, *Digital Content Industry Almanac 2014* (Taipei: Ministry of Economic Affairs, 2015), 126-28, 162-63, 182-83.
Note: Data was not available on year 2015; Ministry of Culture did not conduct survey on Newspaper Publishing Industry.

divisions related with the Publishing Industries were the Digital Publishing and Digital Archives, the Digital Games and the Content Software Industry, respectively. These classifications were somewhat inconsistent with the Publishing Industry's definitions of Digital Newspapers Publishing, Digital Magazines (Periodicals), Digital Book Publishing, Games Publishing and Other Software Publishing based on the "Standard Industrial Classification System of the Republic of China" on Publishing Industry (Rev.10, 2016).

On the other hand, the output values surveyed by the Ministry of Economic Affairs were both a qualitative research and a quantitative study. The qualitative research used interviews with publishers and experts of publishing industries, combined with the opinions from the industries' experts, manufacturing associations and major publishing firms to reach a consensus on the total output value. The quantitative study collected revenue data from the questionnaire on digital content related publishers, weighted in accordance with their respective sizes, further supplemented with the published annual survey report and industry assets report to reach an estimation of the total sales revenues,[14] the estimated revenue was not the actual sales revenue, either.

Research Design and Implementation

In the past, official survey reports on Publishing Industry were relied on the questionnaire inquiring the publishing firms to estimate their approximate annual sales revenues. This research, on the contrary, used Data Mining technology to retrieve the actual sales revenues from the bimonthly business taxes reported by the publishing companies residing in the Financial Database of the Ministry of Finance. Note enterprises with monthly sales less than NT$200,000 were deemed as small scale enterprises for business tax purposes, and for those with monthly sales less than NT$80,000, they were exempted from the business tax, with some exceptions applied to certain business.[15]

Research methodology

As per the study of Chiu,[16] several methods commonly used to understand and predict the sales market of publishing industries included, but not limited to the followings: inquiring experts' opinions, focus group interviews, industry census or questionnaire survey and data comparing and analysis. All of them were approximate, estimations of sales for the publishing industries. Contrary to all previous research, this study applied Data Mining technique to retrieve the actual sales revenue data from the financial database resided in the Ministry of Finance. Noted that the data obtained were Publishing Industries' sales tax data, which were required to be reported to the Finance Ministry by law, thus, these were actual sales revenue data from the Publishing Industries.

Research implementation

a. Steps of the Knowledge Discovery in Database

This research applied data mining technique directly into the Financial Database of the Ministry of Finance to retrieve the business tax information, then backtracking and transforming the tax data into the actual sales revenues of the Publishing Industries. Since all the data were based on business tax reported under government regulation, the sampling errors were very low and the accuracy and confidence level of the data analysis very high. The steps of Knowledge Discovery[17] in Financial Database (KDFD) for the sales revenues of the Publishing Industries were shown in the schematic diagram of KDFD process, see Figure 1.

b. Data Parameters in Financial Database

Revenue data in the Financial Database were screened and retrieved under different parameters such as year, bi-monthly sales, and types of enterprises to study the relationship between the sales revenue and the designate parameter, see Table 3.

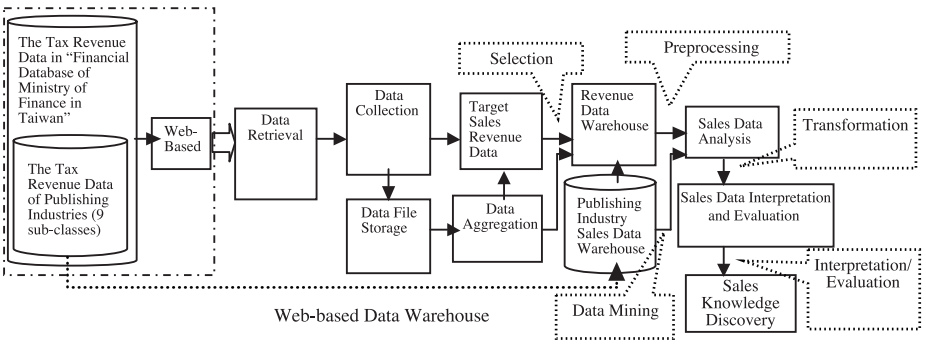

Figure 1 Knowledge Discovery in Financial Database (KDFD)

Source: Ming-Ju Hsu, "Jiao yu bu da xue xiao yuan ke cheng zi yuan wang zhi yuan ke cheng fa zhan mo shi zhi jian gou chuan bo xue men xi suo wei shi zheng [A Study on a Supporting Model for Curriculum Development Using Data Warehouse Technology]," *Zhong hua yin shua ke ji nian bao* [Journal of CAGST] 2012, (June 2012): 580.

Table 3 Descriptions of the Data Variables in the Financial Database to be Studied

Parameters	Descriptions	Data presentation
Years	Divided into year 2013, 2014, 2015	Year 2013 to 2015
Bi-monthly	Since the business taxes were declared bi-monthly, revenues within a year were grouped into 6 periods, coded from 1 to 6. Jan.-Feb. (1), Mar.-Apr. (2), May-June (3), July-Aug. (4), Sept.-Oct. (5), Nov.-Dec. (6)	Presented in 6 bi-monthly periods within a year, coded 1 to 6.
Types of publishing enterprises (Publishing sub-class)	Types of publishing enterprise were divided into 9 sub-classes: (1) 5811-11 Newspapers Publishing Industry, (2) 5811-12 Digital Newspapers Publishing Industry, (3) 5812-11 Magazines (periodicals) Publishing Industry, (4) 5812-12 Digital Magazines (periodicals) Publishing Industry, (5) 5813-11 Book Publishing Industry, (6) 5813-12 Digital Book Publishing Industry, (7) 5819-00 Other Publishing Activities, (8) 5820-11 Game Software Publishing Industry, (9) 5820-99 Other Software Publishing Industry.	Presented in 9 different types of Publishing sub-classes, coded as 1 to 9
Revenues	Revenues classified by year, bi-monthly and types of enterprise	Revenue under each category

Statistics and Analysis

Descriptive statistics

This research applied Data Mining technique to the Financial Database of Ministry of Finance to discover the actual sales revenues of Publishing Industries under 9 different types from fiscal year 2013 to 2015. There were significant gaps in Publishing Industries' output values estimated by the government's official surveys and the actual sales revenues obtained from the Financial Database by this study. For example, the actual sales revenues data mined from the Financial Database were greater than the sales output values estimated by government surveys in Magazine, Digital Magazine (2014) Publishing Industries, respectively.

While the actual sales revenues were less than the sales output values estimated by the surveys in Book, Digital Book, Game Software, and Other Software Publishing Industries. Note the classifications and scopes of Game Software Publishing Industry and Other Software Publishing Industry by the surveys were different from the Financial Database of the Ministry of Finance, which inevitably would cause some inherent difference between the surveyed sales output values and the actual sales revenues. In addition, the surveys conducted by the Industrial Development Bureau of the Ministry of Economic Affairs have included the total assets of Digital Archives into the sales output values shown in Table 2, thus, likes apples and oranges, creating significant gaps, making comparisons unrealistic. Table 4 showed Publishing Industries' sales revenues comparisons between surveys by the Ministry of Culture, Ministry of Economic Affairs and those retrieved from the Financial Database of the Ministry of Finance.

Table 4 Sales Revenues of Publishing Industries: Surveyed Results by Ministry of Culture & Ministry of Economic Affairs vs. Data Mined Results from Financial Database of the Ministry of Finance

Enterprise types and publishing subclasses	Gov.'s official surveys Ministry of Culture Min. of Econ. Affairs			Financial Database of the Ministry of Finance Unit: Thousand NTD			Compared Results
	2013	2014	2015	2013	2014	2015	
5811-11 Newspapers publishing industry	—	—	—	14,513,158	12,886,395	12,579,813	2013 - 2015 No surveyed data from Ministry of Culture
5811-12 Digital newspapers publishing industry	—	—	—	0	9,412	16,289	2013 - 2015 No surveyed data from Ministry of Culture
5812-11 Magazines (periodicals) publishing industry	16.5 billion NTD	15.8 billion NTD	—	17,216,463	16,736,813	16,419,750	1. 2013 actual sales> surveyed sales by 716 million NTD 2. 2014 actual sales > surveyed sales by 936 million NTD 3. 2015 no surveyed data from MOC
5812-12 Digital magazines (periodicals) publishing industry	130 million NTD	280 million NTD	—	0	1,417,860	586,453	1. 2013 actual sales < surveyed sales by 130 million 2. 2014 actual sales > surveyed sales by 1.14 billion NTD 3. 2015 no surveyed data from Ministry of Culture
5813-11 Book publishing industry	28.1 billion NTD	24.3 billion NTD	—	26,995,481	22,692,177	19,048,408	1. 2013 actual sales < surveyed sales by 1.10 billion NTD 2. 2014 actual sales < surveyed sales by 1.61 billion NTD 3. 2015 no surveyed data from Ministry of Culture

5813-12 Digital book publishing industry	510 million NTD	530 million NTD	—	105,439	140,261	201,758	1. 2013 actual sales < surveyed sales by 404 million NTD 2. 2014 actual sales < surveyed sales by 389 million NTD 3. 2015 no surveyed data from Ministry of Culture
5819-00 Other publishing activities	—	—	—	1,567,059	1,569,597	1,358,785	2013 - 2015 no surveyed data from MOC
5820-11 Game software publishing industry	45.3 billion NTD (Digital game industry)	50.6 billion NTD (Digital game industry)	—	6,664,011	7,529,996	9,080,080	1. 2013 actual sales < surveyed sales by 38.64 billion NTD 2. 2014 actual sales< surveyed sales by 43.07 billion NTD 3. 2015 no surveyed data from Industrial Development Bureau of MOEA
5820-99 Other software publishing industry	202.5 billion NTD (Content software industry)	210 billion NTD (Content Software Industry)	—	714,522	1,438,809	2,179,380	1. 2013 actual sales < surveyed sales by 201.79 billion NTD 2. 2014 actual sales < surveyed sales by 208.56 billion NTD 3. 2015 no surveyed data from Industrial Development Bureau of MOEA

Note: The surveyed sales data for fiscal year 2015 were not yet available from the Ministries of Culture and Economic Affairs.

Data mining

a. Decision Tree Analysis

This research used "Decision Tree" to explore the actual sales revenues of Publishing Industries against parameters such as years, bi-monthly sales and enterprise types from 2013 to 2015, respectively.

Algorithms of the Decision Trees included: ID3 (Iterative Dichotomiser 3); C4.5; C5.0; Exhaustive CHAID (Chi Square Automatic Interaction Detector); QUEST (Quick Unbiased Efficient Statistical Tree); and CRT (Classification and Regression Tree, also known as C&RT, CART), etc. The criteria used to evaluate attributes for splitting node of branching in Decision Tree included: Information Gain for ID3; C4.5; C5.0[18]; Gini index, impurity measure for CRT[19] and Chi-Square test of independence for CHAID.[20]

The research used CRT method to explore the scaled dependent variables, all splits are binary recursive partitioning and each parent node was split into only two child-nodes, thus maximizing the within-the-node homogeneity. All observation values in the terminal node possessed the same value of the dependent variable, thereby producing accurate, homogeneous subsets. The impurity measurement of the scaled dependent variables used the least-squared deviation (LSD), to compute it as the within-the-node variance, which can be adjusted for

any frequency weights or influence values.[21]

CRT used the impurity function to find the best location from all the independent variables as a splitting node, to divide the sample into two subgroups with highest homogeneity, whose evaluations depended upon the attributes of variables. For example, categorical (nominal, ordinal) dependent variables could use Gini, Twoing or Ordered Twoing methods; and LSD method could be used to evaluate continuous variables.[22] Each node represented a dependent variable on the CRT map. Each node on the tree displayed mean, standard deviation, observed value and predicted value for scaled dependent variables. The predicted value was the mean value for the dependent variable at the specified node.[23]

The findings were presented in the followings:

Table 5 listed the sales revenues vs. fiscal years, bi-monthly and enterprise types, respectively, obtained by the Decision Tree analysis. Figure 2 displayed the Decision Tree for sales revenue vs. fiscal year relations with 5 nodes. Figure 3 displayed the Decision Tree for sales revenue vs. bi-monthly relation with 3 nodes. Figure 4 displayed the Decision Tree for annual sales revenue vs. the enterprise type relation with 3 nodes, respectively.

Table 5 Sales Revenues vs. Fiscal-Year, Bi-Monthly and Enterprise Type, Obtained by the Decision Tree Analysis, respectively

A. Sales Revenues vs. Fiscal-Year

Nodes	Descriptions	The number of observations
Node 0	From 2013 to 2015, the average bi-monthly sales per publishing sub-class was $1.20 billion NTD	All numbers of observations, 162 (100%)
Node 1	The average bi-monthly sales per publishing sub-class was 1.26 billion NTD in 2013	54 items of observations (33.3%)
Node 2	The average bi-monthly sales per publishing sub-class was $1.17 billion NTD in 2014 and 2015 combined	108 items of observations (66.7%)
Node 3	The average bi-monthly sales per publishing sub-class was 1.19 billion NTD in 2014	54 items of observations (33.3%)
Node 4	The average bi-monthly sales per publishing sub-class was 1.14 billion NTD in 2015	54 items of observations (33.3%)

B. Sales Revenue vs. Bi-Monthly

Nodes	Descriptions	The number of observations
Node 0	From 2013 to 2015, the average bi-monthly sales per publishing sub-class was 1.20 billion NTD	All numbers of observations, 162 (100%)
Node 1	The average bi- monthly sales per publishing sub-class was 1.09 billion NTD in Jan.-Feb. (1), May-June (3), and July-Aug. (4) combined	81 items of observations (50.0%)
Node 2	The average bi- monthly sales per publishing sub-class was 1.30 billion in Mar.-Apr. (2), Sept.-Oct. (5), and Nov.-Dec. (6) combined	81 items of observations (50.0%)

C. Sales Revenue vs. Enterprise Type

Nodes	Descriptions	Total number of observations
Node 0	From 2013 to 2015, the average bi-monthly sales per publishing sub-class was 1.20 billion NTD	All items of observations, 162 (100%)
Node 1	The average bi-monthly sales per publishing sub-class of the Newspapers, Magazines (Periodicals) and Books Publishing Industries combined was 2.95 billion NTD	54 items of observations, 33.3% (sales revenue were 159,088 million, accounted for 82.1% of the total sales revenue)
Node 2	The average bi-monthly sales per publishing sub-class of Digital Newspapers, Digital Magazines (Periodicals), Digital Book, Other Publishing Activities, Games and Other Software Publishing combined was 320.2 million NTD	108 items of observations, 66.7% (sales revenue were 34,579 million, accounted for 17.9% of the total sales revenue)

Note: Numeric codes 1 -6 were a bi-monthly coding, 1 represented the first two months of the year, and so forth.

b.Sales Revenues vs. Three Different Parameters

1. Sales Revenue vs. Fiscal-Year

Figure 2 displayed the Decision Tree for sales revenue vs. fiscal year relations with 5 nodes.

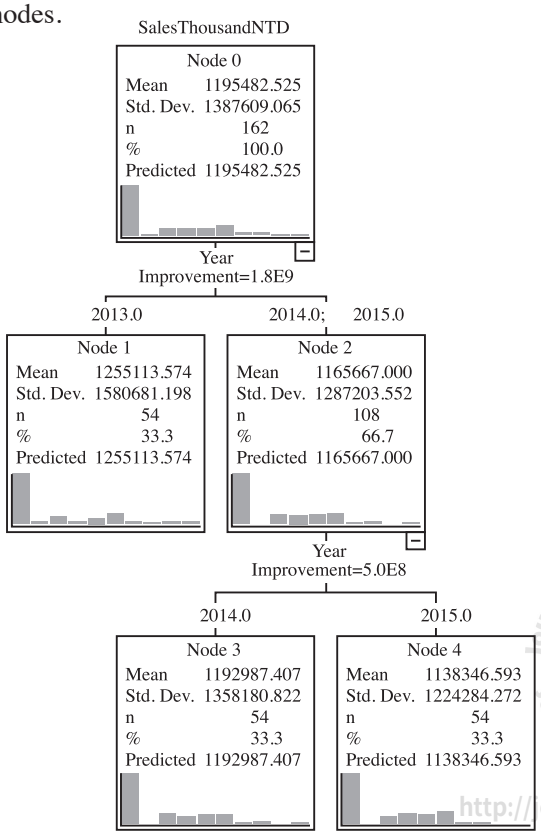

Figure 2 Decision Tree for Annual Sales Revenue vs. Fiscal-Year

In this study, the decision tree nodes were divided into fiscal-year 2013, 2014 and 2015 three groups, and the total revenue was the dependent variable (Table 6). Levene's Test of Homogeneity of Variances[24] showed F=1.729, p=.181>0.05, p-value was greater than the level of significance, so it agreed with the Test of Homogeneity of Variances. The results showed that there was no significant difference in variances among the three fiscal-year groups and the degree of dispersion was similar. Further, the result of the one-way variance analysis showed F=0.095, p=0.910> 0.05, indicating that there was no significant difference in revenues for each fiscal year statistically. However, comparing the means of the three fiscal years indicated that the average sales revenues of each publishing subclass were decreasing year by year, from 1.26 to 1.19 then to 1.14 billion NTD continuously.

2. Sales Revenue vs. Bi-Monthly Period

Figure 3 displayed the Decision Tree for sales revenue vs. Bi-Monthly Period relation with 3 nodes.

In this study, the decision tree nodes were divided into two groups, with bi-month sales period of January-February, May-June and July-August as one group, and March-April, September-October and November-December as another group, with the total revenues as the dependent variable (Table 6). Levene's Test showed F=3.258, p=.073>0.05 and it met the test of homogeneity variance. The result indicated no significant difference in variances between the two groups and the degree of dispersion was similar. Assuming an equal variance, t = -.943, p = 0.347> 0.05, indicating that there was no significant difference between the two groups.

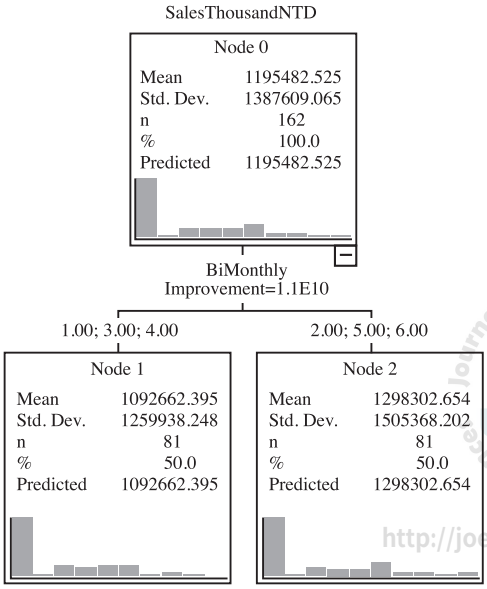

Figure 3 Decision Tree for Sales Revenue vs. Bi-Monthly

However, the means indicated the sales revenue of the second half of the year was better that that of the first half of the year.

In general, the average bi-monthly sales revenues per publishing sub-class were low in the first 2 months of the year, then increased some, but experienced a steep drop in May/June, July/August months, then gradually picked up in September/October and reached the peak sales months of November/December (Table 6). Factors affecting this cyclical sales performance were not clearly known, but seasonal buying and consumer sentiments at certain occasions probably accounted for significant swings in buying mood at a given time.

3. Sales Revenues vs. Enterprise Type

Figure 4 displayed the Decision Tree sales revenue vs. enterprise type relation with 3 nodes.

In this study, the decision tree nodes were divided into two groups, the Newspapers, Magazines and Books Publishing Industries as one group, and the Digital Newspapers, Digital Magazines (Periodicals), Digital Book, Other Publishing Activities, Games and Other Software Publishing as another group (Table 6). Levene's Test showed F=18.747, p=0.000 <0.05, indicating variance was not homogeneous. In this case, t=-21.133, p=0.000 <0.05, the average value of the revenue of each subclass of Newspaper, Magazine and Book publishing

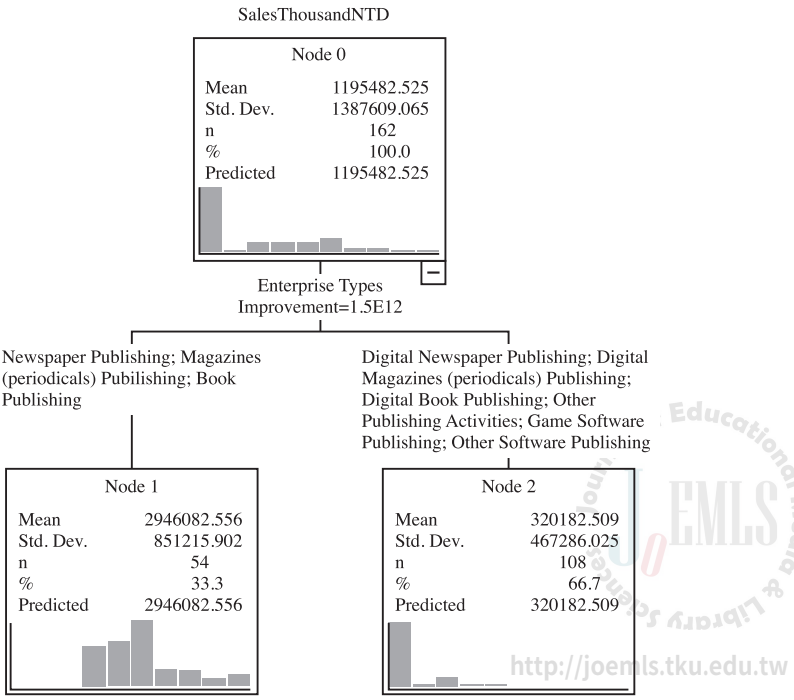

Figure 4 Decision Tree for Sales Revenue vs. Enterprise Type

Table 6 Summary Table of Mean, Standard Deviation, Number of Nodes on Fiscal-Year, Bi-Monthly, Enterprises Type by Decision Tree

Fiscal-Year				
	Gain for nodes	Mean	Std. deviation	N
Sales (Thousand NTD)	2013	1255113.574	1580681.198	54
	2014	1192987.407	1358180.822	54
	2015	1138346.593	1224284.272	54
	Total	1195482.525	1387609.065	162
Bi-Monthly				
	Gain for nodes	Mean	Std. deviation	N
Sales (Thousand NTD)	Jan.-Feb. (1), May-June (3), July-Aug. (4)	1092662.395	1259938.248	81
	Mar.-Apr. (2), Sep.-Oct. (5), Nov.-Dec. (6)	1298302.654	1505368.202	81
	Total	1195482.525	1387609.065	162
	Means of bi-monthly: (1) 1078536.296; (2) 1253976.407; (3) 1086841.519; (4) 1112609.370; (5) 1232572.148; (6) 1408359.407			
Enterprises Type (publishing sub-class)				
	Gain for nodes	Mean	Std. deviation	N
Sales (Thousand NTD)	Newspapers, magazines (Periodicals) and book publishing	2946082.556	851215.902	54
	5 Digital Related Publishing sub-classes and Other Publishing Activities	320182.509	467286.025	108
	Total	1195482.525	1387609.065	162

was significantly higher than that of other six publishing subclasses combined, among them, 5 were Digital Publishing related sub-classes. This finding clearly demonstrated that the printed form publishing was still the dominate part of the sales for the publishing industries.

The average bi-monthly sales revenues per publishing sub-class were 1.20 billion NTD from 2013 to 2015. While the average bi-monthly sales revenues per publishing sub-class for Newspapers, Magazines and Book publishing were 2.95 billion NTD, this was 9.20 times of the average bi-monthly sales revenues per publishing sub-class for the other six (6) sub-classes sales of 320.2 million NTD. These 3 Publishing Industries accounted for 82.1% of the total sales revenue. Among the 6 sub-classes, including 5 Digital Publishing related sub-classes, Digital Newspapers, Digital Magazines (Periodicals), Digital Books, Games Software, Other Software Publishing and Other Publishing Activities had an average of 17.9% of market shares during the 2013-2015 periods.

c. Market Shares of Digital Publishing Sales Revenues

Digital Publishing Industries (5 sub-classes) accounted for an average of 16% of the total sales revenues during the 2013 to 2015 time period. A year-to-

year comparison indicated that the sales revenues of Digital Publishing industries had grown significantly, from 11% of market shares in 2013 to 16% in 2014, then to 20% in 2015, respectively, please see Table 7.

Digital publishing clearly will continue to gain in popularity and enjoys a double-digit growth rate for the foreseeable future. Mobil devices simply create a powerful platform, allowing instant communication and access to digital news, texts, photos, books, music, movies, etc. anytime anywhere with no or little cost. It's considered a disruptive force by the traditional publishing industries. For example, self-publishing has changed the interaction between authors and publishers, enabling a faster, more cost-effective method of publishing impossible

Table 7 2013-2015 Sales Revenues vs. Enterprise Types and Publication Types

Unit: Thousand NT

Enterprise types and publishing subclasses	Enterprise types	2013 Thousand NT (%)		2014 Thousand NT (%)		2015 Thousand NT (%)		Sub-total Thousand NT (%)	
5811-12 Digital Newspapers Publishing Industry	Digital	0 (0.00)	11%	9,412 (0.01)	16%	16,289 (0.03)	20%	25,701 (0.01)	16%
5812-12 Digital Magazines (periodicals) Publishing Industry	Digital	0 (0.00)		1,417,860 (2.20)		586,453 (0.95)		2,004,313 (1.03)	
5813-12 Digital Book Publishing Industry	Digital	105,439 (0.16)		140,261 (0.22)		201,758 (0.33)		447,458 (0.23)	
5820-11 Game Software Publishing Industry	Digital	6,664,011 (9.83)		7,529,996 (11.69)		9,080,080 (14.77)		23,274,087 (12.02)	
5820-99 Other Software Publishing Industry	Digital	714,522 (1.05)		1,438,809 (2.23)		2,179,380 (3.55)		4,332,711 (2.24)	
5811-11 Newspapers Publishing Industry	Printed	14,513,158 (21.41)	89%	12,886,395 (20.00)	84%	12,579,813 (20.46)	80%	39,979,366 (20.64)	84%
5812-11 Magazines (periodicals) Publishing Industry,	Printed	17,216,463 (25.40)		16,736,813 (25.98)		16,419,750 (26.71)		50,373,026 (26.01)	
5813-11 Book Publishing Industry	Printed	26,995,481 (39.83)		22,692,177 (35.22)		19,048,408 (30.99)		68,736,066 (35.49)	
5819-00 Other Publishing Activities	Printed	1,567,059 (2.31)		1,569,597 (2.44)		1,358,785 (2.21)		4,495,441 (2,32)	
Sub-total		67,776,133 (100.00)		64,421,320 (100.00)		61,470,716 (100.00)		193,668,169 (100.00)	

to imagine just a few years ago; another example is Kindle Unlimited (Amazon's Kindle Unlimited e-book subscription service), offering unlimited books to read at discounted fees, which totally disrupted the pricing structures of traditional publishing houses.[25]

According to PwC (Pricewaterhouse Coopers International),[26] although revenues of printed and audio publishing will continue to fall, it would still account for 70% of the total publishing revenues. The combined revenues generated by global print/audio publishing would decrease to $92.39 billion by 2019, an annual decline rate of 1.9% from $101.63 billion of 2014. Moreover, educational and professional books will take in $72 dollars of every $100 dollars spent in printed books by 2019.

In Taiwan, although sales revenues of printed newspaper, magazine/periodical and book publishing decreased every year from 2013 to 2015, while their digital counterpart of news, magazine and book publishing increased every year at the same time, the magnitudes of the former were simply too huge for the later to be presented with any significance in comparison. More specifically, printed newspaper publishing revenues vs. digital news publishing was 99.94% vs. 0.06%; magazine publishing vs. digital magazine publishing 96.2% vs. 3.8%; and books publishing vs. digital books publishing 99.4% vs. 0.6%. In short, of the three types of publishing enterprises mentioned, paper-media sales accounted for more than 96% of the sales revenues in each subclass (Table 7).

Analysis of the bi-monthly sales revenues indicated that publishing sales were cyclical, January / February sales were typically the lowest, mostly because of shutdown for the Chinese New Year holidays, sales would pick up in March to April, went down continually in the summer season, then went up again in September/October months and reached the peak sales season of November and December. Various cyclical factors should be studied to introduce promotional activities to stimulate the sales revenues.

A major concern was the highest revenue creator in traditional publishing, the Book publishing, had experienced significant sales revenue drops in all 3 years studied. The magnitudes of reductions were far greater than the revenues generated by digital newspaper, magazines and books combined, leaving a huge revenue hole to be filled. A bright side was the emerging Game Software Publishing and Other Software Publishing, which continued to grow from 2013 to 2015 and reached a total 3-year sales revenue of 27,606,798 thousand NTD, which has already exceeded the 2013 Books publishing sales revenue of 26,995,481 thousand NTD.

"Digital Disruption" has caused a shift of publishing paradigm in Taiwan, although the printed publications still hold the majority share of the market at this moment. The followings were several suggestions to help improve its position:

1. In the field of printed publishing, maintain and enhance the paper-media reading experience; create a buying atmosphere at the bookstores; make purchasing more convenient by expanding into new sales channels; diversify marketing strategies and emphasize on consumer interactions; conduct focused promotions and increase the content-value and application of the publication, etc.

2. In response to the impacts of digital publishing, improve efficiency and streamline operation are essential for the traditional publishing industry. Reform and transform, seek synergies and opportunities created by digital publishing can also increase its sales revenues. For example, develop game related publishing and software, IT related publishing, etc.

3. Get involved and mobilize. In other words, traditional publishing industry should also break the barrier by itself and seeks opportunities created by digital disruption. Self-Publishing is a good example, after all, the author's digital creation, also needs to be printed. Editing support, remote consultation and micro-publishing are other opportunities that can enhance the value and position of traditional publishing industry.

Conclusions and Suggestions

Revenues Comparison between this Research and Government's Survey Results

This research compared the publishing industries' sales revenues estimated by surveys by government agencies with the actual sales data obtained from the Financial Database of the Ministry of Finance. The research indicated that the actual sales revenues for the Magazines Publishing Industry were greater than the surveyed estimation by the government agencies, while the actual sales revenues were less than the sales output values estimated by the surveys in Books, Digital Books, Game Software, and Other Software Publishing Industries. The comparison has shown a significant gap between the two methods, most of the surveyed sales output values were greater than the actual sales revenues. In addition, a major factor contributed to this huge difference was a lack of clear and consistent definition for the classification and scope for Digital Publishing Industries, for example, Other Software and Game Software Publishing Industries have shown significant gaps between the surveyed sales output values and the actual sales revenues.

Knowledge Discovery in Financial Database on Publishing Industries

This research used "Decision Tree" to explore parameters such as fiscal year, bi-monthly, and enterprise type on sales revenues from 2013 to 2015. The research found that the average bi-monthly sales per publishing sub-class was around 1.20 billion NTD, and the average bi-monthly sales per publishing sub-

class has shown a continual decline from 2013 to 2015. Within the year, January/February had the lowest sales revenues because of Chinese New Year holidays, the revenues picked up in March/April, then showed a seasonal decline in the summer months, gradually picked up in September/October and reached the sales peak by Novermber/December period. In term of enterprise types, newspapers printing, magazines (periodicals) publishing and hard copy books publishing were still accounted for the majority of the sales for the Publishing Industries in Taiwan. While the Digital Publishing Industries (5 sub-classes) accounted for an average of 16% of the total sales revenues during the 2013 to 2015 time period, the growth spurs from 11% in 2013 to 16% in 2014, then to 20% in 2015 were quite impressive, indicating a potential growth for Digital Publishing Industries in Taiwan.

Although digital publishing in Taiwan has shown impressive growth rates during the studied 3 years (2013-2015), it's still at its initial, developing stage. Sales revenues of traditional publishing still dominate and accounts for 84% of the total revenues generated. Nevertheless, Digital Disruption is for real and should be welcomed, for it signified a new era and a more readily connected world.

Recommendations Derived from the Results of the Study

Promote the advantages of Digital Publishing to further drive up its sales revenues. In each subclass of newspaper publishing, magazine publishing and book publishing, paper-media sales accounted for more than 96% of the sales revenues (Table 7). In 2016, the printed to digital book publishing sales ratio in the United States was projected to be $30.52 billion dollars to $7.59 billion dollars,[27] or 75.13% to 24.87%. In comparison, the averaged 99.4% to 0.6% ratio from 2013-2015 in Taiwan indicated its digital book publishing needs to be heavily promoted to reach a more balanced ratio for a healthier competition with traditional publishing.

Traditional publishing industry in Taiwan may be impacted negatively in the future, it can survive and even strive in digital publishing era, as described by the suggestions mentioned in the previous section.

Finally, future government survey of the sales revenues of the Publishing Industries in Taiwan is recommended to adopt this research method, i.e., data mining the Financial Database maintained by the Financial Ministry, to obtain the actual sales revenues, instead continue using estimated revenue numbers.

Recommendation for Future Research

The definition of the publishing industries categorized by the government of Taiwan included nine subclasses in the whole division, future study could be conducted for each subclass to obtain its actual revenue for a more realistic comparison with surveyed data.

Acknowledgements

This work was supported by Mr. S. N. Chen of the Ministry of Finance, he can be reached at snchen@mail.mof.gov.tw

Notes

1. Rudiger Wischenbart and Jennifer Krenn, "Global Publishing Monitor 2014," in *IPA Annual Report October 2013-October 2014* (Geneva: International Publishers Association, 2014), 13, accessed May 28, 2016, http://www.internationalpublishers.org/images/reports/2014/IPA-annual-report-2014.pdf.

2. 文化部，102年暨103年臺灣出版產業調查報告（台北市：文化部，2015），1。【Ministry of Culture, *Survey of Taiwan's Publishing Industry for 2013-2014* (Taipei: Ministry of Culture, 2015), 1. (in Chinese)】

3. 邱炯友，「論圖書出版產業調查之設計與執行」，資訊傳播與圖書館學，10卷1-4期（2004年6月）：15-16；李琴，「書業調查：讓人感覺有點亂」，光明日報，2002年4月4日，檢索於2016年5月14日，http://www.china.com.cn/chinese/RS/127908.htm。【Jeong-Yeou Chiu, "The Planning and Implementation of Book Publishing Industry Survey," *Journal of Information, Communication, and Library Science* 10, no. 1-4 (June 2004): 15-16 (in Chinese); Qin Li, "Shuye Diaocha: Rang Ren Ganjue Youdian Luan," *Guangming Daily*, April 4, 2002, accessed May 14, 2016, http://www.china.com.cn/chinese/RS/127908.htm. (in Chinese)】

4. 行政院主計總處，行業標準分類（第*10*次修訂）（台北市：行政院，2016），1-2，檢索於2016年5月25日，http://www.dgbas.gov.tw/public/Attachment/51230162221KI9NIEKP.pdf。【Directorate-General of Budget, Accounting and Statistics, *Hangye Biaozhun Fenlei (Di 10 Ci Xiuding)* (Taipei: Executive Yuan, 2016), 1-2, accessed May 25, 2016, http://www.dgbas.gov.tw/public/Attachment/51230162221KI9NIEKP.pdf. (in Chinese)】

5. 徐明珠，「出版業行業標準分類之探討」，中華印刷科技學會年報，2016期（2016年4月）：210-11。【Ming-Ju Hsu, "Discussion of the Standard Industrial Classification of Publishing Industry in Republic of China," *Journal of CAGST*, no. 2016 (April 2016): 210-11. (in Chinese)】

6. 邱炯友，「論圖書出版產業調查之設計與執行」，16-17。【Chiu, "The Planning and Implementation of Book Publishing Industry Survey," 16-17. (in Chinese)】

7. Ibid., 17.

8. 邱炯友、花湘琪，「台灣圖書出版市場之背景與國際比較」，在中華民國89年臺灣圖書雜誌出版市場研究報告，行政院文化建設委員會編（台北市，行政院文化建設委員會，2001），1-2，檢索於2016年5月14日，http://www.magazine.org.tw/events/school/report/2000tpmr/A.PDF；王乾任，臺灣出版產業大未來（台北市：生活人文，2004），37-66。【Jeong-Yeou Chiu and Xiang-Qi Hua, "Taiwan Tushu Chuban Shichang zhi Beijing yu Guoji Bijiao," in *Taiwan Publishing Market Report (Book Publishing Industry and Magazine Publishing Industry 2000)*, ed. Council for Cultural Affairs (Taipei: Council for Cultural Affairs, Executive Yuan, 2001), 1-2, accessed May 14, 2016, http://www.magazine.org.tw/events/school/report/2000tpmr/A.PDF (in Chinese); Chien-Ren Wang, *Taiwan Chuban Chanye Daweilai* (Taipei: Living Humanities Press, 2004), 37-66. (in Chinese)】

9. 林麗如，「兩岸出版交流政策之研究──臺灣的觀點」（碩士論文，國立政治大學行政管理碩士學程，2011），6。【Li-Roo Lin, *Research on Cross-Strait Exchange Policies on the Printing Industry — Taiwan's Perspectives* (master's thesis, Master for Eminent Public Administration, National Chengchi University, 2011), 6. (in Chinese)】

10. 梁德馨，「從『2007圖書出版及行銷通路業經營概況調查』看臺灣圖書出版產業的未來」，全國新書資訊月刊，4月號（2008年4月）：8；行政院新聞局，*99*年圖書出版產業調查報告（台北市：行政院，2010），檢索於2016年5月26日，http://mocfile.moc.gov.tw/mochistory/images/Yearbook/2010survey/catalog3-1.html。【Dexin Liang, "Cong *2007 Tushu Chuban ji Xingxiao Tongluye Jingying Gaikuang Diaocha* Kan Taiwan Tushu Chuban Chanye de Weilai," *New Books: Recent Publications in Taiwan, ROC*, no. 4 (April 2008): 8 (in Chinese); Government Information Office, Executive Yuan, *2010 Taiwan Publishers Survey* (Taipei: Executive Yuan, 2010), accessed May 26, 2016, http://mocfile.moc.gov.tw/mochistory/images/Yearbook/2010survey/catalog3-1.html. (in Chinese)】

11. 邱炯友，「解讀臺灣圖書與雜誌期刊產業官方調查報告」，全國新書資訊月刊，168期（2012年12月）：11-12，檢索於2016年5月14日，http://isbn.ncl.edu.tw/NCL_ISBNNet/C00_index.php?PHPSESSID=mn86s82c14ncepbcg6ccc7fuu0&Pfile=2484&KeepThis=true&TB_iframe=true&width=900&height=650。【Jeong-Yeou Chiu, "Jiedu Taiwan Tushu yu Zazhi Qikan Chanye Guanfang Diaocha Baogao," *New Books: Recent Publications in Taiwan, ROC*, no. 168 (December 2012): 11-12, accessed May 14, 2016, http://isbn.ncl.edu.tw/NCL_ISBNNet/C00_index.php?PHPSESSID=mn86s82c14ncepbcg6ccc7fuu0&Pfile=2484&KeepThis=true&TB_iframe=true&width=900&height=650. (in Chinese)】

12. International Publishers Association, *Annual Report October 2014 - October 2015* (Geneva: International Publishers Association, 2015), 17, accessed June 6, 2017, https://www.internationalpublishers.org/images/annual-reports/ipa_ar_online.pdf.

13. 行政院文化建設委員會，中華民國八十九年臺灣圖書雜誌出版市場研究報告（台北市：行政院文化建設委員會，2001），7-8；邱炯友，「臺灣圖書出版統計數據─閱讀習慣與消費行為」，資訊傳播與圖書館學，8卷2期（2001年12月）：39-40。【Council for Cultural Affairs, Executive Yuan, Republic of China, *2000 Taiwan Tushu Zazhi Chuban Shichang Yanjiu Baogao* (Taipei: Executive Yuan, 2001), 7-8 (in Chinese); Jeong-Yeou Chiu, "Issues Affecting Book Publishing Statistics: With Particular Reference to Reading Habit and Consumer Behavior in Taiwan," *Journal of Information, Communication, and Library Science* 8, no. 2 (December 2001): 39-40. (in Chinese)】

14. 經濟部工業局，*2014 Taiwan*數位內容產業年鑑（台北市：經濟部，2015），126。【Industrial Development Bureau, Ministry of Economic Affairs, *2014 Digital Content Industry in Taiwan* (Taipei: Ministry of Economic Affairs, 2015), 126. (in Chinese)】

15. 財政部賦稅署、財政部臺北國稅局，「參加亞洲稅務行政及研究組織（SGATAR）第45屆年會會議報告」，（論文發表於亞洲稅務行政及研究組織（SGATAR）第45屆年會，新加坡，2015年11月24-26日），檢索於2016年7月4日，http://report.nat.gov.tw/ReportFront/report_detail.jspx?sysId=C10500017。【Taxation Administration, Ministry of Finance, R. O. C. and National Taxation Bureau of Taipei, Ministry of Finance, "Canjia Study Group on Asian Tax Administration and Research (SGATAR) Di 45 Jie Nianhui Huiyi Baogao," (paper presented at The 45th Study Group on Asian Tax Administration and Research (SGATAR)

Meeting, Singapore, November 24-26, 2015), accessed July 4, 2016, http://report.nat.gov.tw/ReportFront/report_detail.jspx?sysId=C10500017.】

16. 邱炯友，「臺灣圖書出版統計數據—閱讀習慣與消費行為」，39-40。【Chiu, "Issues Affecting Book Publishing Statistics: With Particular Reference to Reading Habit and Consumer Behavior in Taiwan," 39-40. (in Chinese)】

17. Jiawei Han and Micheline Kamber, *Data Mining: Concepts and Techniques*, 2nd ed. (Burlington: Morgan Kaufmann, 2006), 5-12；徐明珠，「教育部大學校院課程資源網支援課程發展模式之建構—傳播學門系所為實證」，中華印刷科技學會年報，2012期（2012年6月）：565-67。【Jiawei Han and Micheline Kamber, *Data Mining: Concepts and Techniques*, 2nd ed. (Burlington: Morgan Kaufmann, 2006), 5-12; Ming-Ju Hsu, "A Study on a Supporting Model for Curriculum Development Using Data Warehouse Technology," *Journal of CAGST*, no. 2012 (June 2012): 565-67. (in Chinese)】

18. J. R. Quinlan, "Induction of Decision Trees." *Machine Learning*, no. 1 (March 1986): 90-102; Rutvija Pandya and Jayati Pandya, "C5.0 Algorithm to Improved Decision Tree with Feature Selection and Reduced Error Pruning," *International Journal of Computer Applications* 117, no. 16 (May 2015): 18-20.

19. Leo Breiman, Jerome Friedman, Charles J. Stone, and R. A. Olshen, *Classification and Regression Trees* (Monterey: Wadsworth & Brooks/Cole Advanced Books & Software, 1984), 24-25.

20. 陳士杰，「決策樹學習」，檢索於2017年3月10日，http://sjchen.im.nuu.edu.tw/MachineLearning/final/CLS_DT.pdf。【Shi-Jie Chen, "Decision Tree Learning," accessed March 10, 2017, http://sjchen.im.nuu.edu.tw/MachineLearning/final/CLS_DT.pdf. (in Chinese)】

21. IBM, *IBM SPSS Decision Trees 20* (n.p., IBM Corporation: 2011), 12; IBM, *IBM SPSS Decision Trees: Easily Identify Groups and Predict Outcomes* (n.p., IBM Corporation Software Group: 2012), 1-5.

22. IBM, *IBM SPSS Decision Trees: Easily Identify Groups and Predict Outcomes*, 1-5; Quest Software, "Computational Formulas," Quest Software Inc., 2016, accessed March 12, 2017, http://documentation.statsoft.com/STATISTICAHelp.aspx?path=GXX/Gcrt/Overviews/ComputationalFormulas；長榮大學醫務管理學系，牙醫費用群組之發展及其檔案分析之應用研究（*CDA-9402*）（台北市：中華民國牙醫師公會全國聯合會，2006），12-13。【IBM, *IBM SPSS Decision Trees: Easily Identify Groups and Predict Outcomes*, 1-5; Quest Software, "Computational Formulas," Quest Software Inc., 2016, accessed March 12, 2017, http://documentation.statsoft.com/STATISTICAHelp.aspx?path=GXX/Gcrt/Overviews/ComputationalFormulas; Department of Health Care Administration, Chang Jung Christian University, *Yayi Feiyong Qunzu zhi Fazhan ji Qi Dangan Fenxi zhi Yingyong Yanjiu (CDA-9402)* (Taipei: Taiwan Dental Association, 2006), 12-13. (in Chinese)】

23. IBM, *IBM SPSS Decision Trees 20*, 51.

24. Howard Levene, "Robust Tests for Equality of Variances," in *Contributions to Probability and Statistics: Essays in Honor of Harold Hotelling*, ed. Ingram Olkin (Stanford University Press, 1960), 278.

25. 文化部，「美國電子出版業現況」，國內外文化產業訊息及趨勢分析雙月報，3

期（2015年6月）：13-16；Nick Statt, "Kindle Unlimited: Good for Customers, Not So Good for Authors?" CBS Interactive Inc., July 19, 2014, accessed March 12, 2017, https://www.cnet.com/news/amazon-kindle-unlimited-good-for-customers-not-so-good-for-author。【Ministry of Culture, "Meiguo Dianzi Chubanye Xiankuang," *Guoneiwai Wenhua Chanye Xunxi ji Qushi Fenxi Shuangyuebao*, no. 3 (June 2015): 13-16 (in Chinese); Nick Statt, "Kindle Unlimited: Good for Customers, Not So Good for Authors?" CBS Interactive Inc., July 19, 2014, accessed March 12, 2017, https://www.cnet.com/news/amazon-kindle-unlimited-good-for-customers-not-so-good-for-author.】

26. PwC, "Outlook Insights: An Analysis of the Global Entertainment and Media Outlook 2015-2019," PwC, 2015, accessed March 12, 2017, http://www.pwc.com /outlook.

27. Statista, "Revenue of Book Publishing (NAICS 51113) in United States from 2008 to 2020 (in Million U.S. Dollars)," Statista, 2017, accessed March 12, 2017, https://www.statista.com/forecasts/409695/united-states-book-publishing-revenue-forecast-naics-51113.

Ming-Ju Hsu ORCID 0000-0002-0447-3644

教育資料與圖書館學　54 : 2 (2017) : 161-184
DOI:10.6120/JoEMLS.2017.542/0054.RS.BE

研究論文

台灣出版產業銷售分析：資料庫知識發現觀點

徐明珠*

摘要

本研究運用財政部「財政統計資料庫」進行資料庫知識發現，以2013至2015年出版業為資料探勘目標，獲得台灣地區出版產業的銷售態樣。研究結果發現：a)我國出版業的銷售收入從2013年到2015年逐年下降；b)年度內銷售額從5月到6月急劇下降，然後每兩個月銷售收入逐漸恢復、增加，達到11月至12月高峰期，c)紙本新聞出版、雜誌（期刊）出版和書籍出版仍占出版產業營業銷售額大宗（82.1%）。而在數位出版占整體出版銷售額比例方面，2013至2015年三年整體比例為16%，2013年至2015年，各年度分別為11%、16%、20%，數位出版占整體出版銷售額比例逐年成長。未來宜針對細部子類分析，具體而微構建出版業銷售態樣，並結合商業模式、行銷創新論點，預應出版業趨勢與發展。

關鍵詞：出版業，數位出版，資料探勘，資料庫知識發現，決策樹分析，銷售額

* 佛光大學傳播學系副教授
E-mail: mjhsu@mail.fgu.edu.tw