Analysis Of Cross Tabulation Through Chi-Squared Test And Pareto Analysis On Malaysian International Tourism Data

SuliadiSufahani*,*, Marinah Muhammad**, Zuhaim Ismail****

* Department of Mathematics, Faculty of Science, UniversitiTeknologi Malaysia, Skudai, 81310 Johor Bahru, Johor, Malaysia
** Department of Earth Science, Faculty Agro Industry and Natural Resources, Universiti Malaysia Kelantan, 16100 PengkalanChepa, Kelantan, Malaysia
*** Department of Mathematics and Statisitcs, Faculty of Science Technology and Human Development, UniversitiTun Hussein Onn Malaysia, Parit Raja, 86400 BatuPahat, Johor, Malaysia

Abstract- The objectives of this study to presents an analysis of the data, retrieved from the guest book at the Kelantan Tourist Information Centre Malaysia, using descriptive cross tabulationwith Chi Square test and Pareto analysis. The analysis outcome was discussed and further suggestions were proposed based on the chi-squared test and Pareto analysis results.

Index Terms- Chi square method, Pareto Analysis, Quantitative method.

I. INTRODUCTION

Since April 1994, the Kelantan's Tourist Information Centre, TIC has been given the responsibility to transform Kelantan as a famous tourist destination in the country. On January 2007, TIC has taken an initiative steps by collecting the tourist profile data through guestbook. However, there was no further analysis is made of the data. Therefore, this study provides the genuine results. During the year of 2014, a total number of836international tourists visited TIC in order to find information on Kelantan. 73% of them are European. They were mostly from France and followed by Netherland, United Kingdom and Germany. Another 10% are from Oceania countries namely Australia and New Zealand and the rest are from various parts of the world.

Forecasting models must be developed with a clear understanding in both the nature of the situation for which a forecast is desired and the resources available for making the forecast [6]. It is important to ensure that the variable selected relates directly to the forecast data needed [5]. This does not mean that forecasts are useless, but that those who use them should be constantly monitoring their operating environment to detect any factors that indicate any inconsistencies or irregular patterns [2].

This paper begins with a brief description, the follows by data source of the study and the proposed analysis. The research methodology is in the third section. The results and analysis are in the fourth section and finally ended with the conclusions.

II. DATA

This paper presents an analysis of the secondary data which obtained from the guestbook at the Kelantan's Tourist Information Centre, TIC. Tourists are required to fill in some information such their name, age, country of origin, comment about Kelantan, length of stay and next destination. Recent research found that, tourism demand modelling and forecasting research relies heavily on secondary data in term of model construction and estimation [7]. This study analysed the tourist profile data that was collected in year 2014 and shown in Table 1.

Table1: International Tourist Arrivals in Kelantan by Country for year 2014

Continental	Country	Total	Percentage %
Asia	Thailand	549153	95.7
	Indonesia	2765	1.2
	Singapore	1345	
	Philipine	1654	
	China	137	
	Middle East	106	
	Taiwan	33	
	Hong Kong	5	
	Pakistan	238	
	Japan	450	
	South Korea	68	
Oceania	Australia	334	0.1
	New Zealand	117	
Europe	U.Kingdom/ Ireland	1657	1.1
	Germany	774	
	France	813	
	Norway /Sweden /Denmark /Finland	733	
	Belgium /Luxumberg /Netherland	411	
	Russia	164	

	Others West Europe	940	
	East Europe	298	
North America	United State Of America	568	
	Canada	545	0.2
South America	Latin America	172	
Others		9835	1.7
	Grand Total	**573649**	**100.0**

III. RESEARCH METHODOLOGY

In this analysis, cross tabulations was used to develop the contingency table for two categorical variables. Through this analysis, Chi Square test for independence was used to determine whether there is a significant relationship between two categorical variables or not:

H_o:There is no relationship between the variables.
H_1:The exits a relationship between the variables.

The significance level chosen for two tailed Chi Square test was 0.05. The cross tabulation analysis also included the correlation value; to measure the level of correlation between two categorical variables. For example, chi-Square and contingency table were used to study the differential response of patient with different psychiatric diagnoses to a given rehabilitation programme [1]. It also used to examine the world's image from an African perspective, exploring those countries in the world that are best known to Africans and the factors responsible for the pattern of recall and representation in the maps [3]. While Summers [8], used chi-square to identify factors that are associated with forecast management and tree planting practices of small farmers in the Brazilian Amazon.

This study also considered the Pareto analysis for the further investigation of the data. The technique helps the decision makers to select few good options from a long list of options. The Pareto's approach is called the "80/20 Rule."

IV. RESULT AND ANALYSIS

A. Chi-Squared Test

The objective of this study is to give an overview of the international tourist behaviour towards tourism sector in Kelantan, Malaysia. It can be done by using cross tabulation analysis and Chi-Square test. Table 2 shows the contingency table from cross tabulation analysis between variable "tourists' countries of origin" and "tourists' next destination". Based on the results, 22% of the international tourist preferred Kelantan as a transit point to the "Perhentian Islands" which in the neighboring state, while another 14% intending to visit the "Kuala Koh National Park".This clearly shows that "Kuala Koh National Park" is the most preferable destination for international tourists in Kelantan and has the potential to become a world-class tourism attraction. If the authorities decided to develop "Kuala Koh National Park", then it will solve 50% of the major problems in Kelantan tourism sector. A cross tabulation analysis

was also done between variable 'tourists' comment' with three other variables namely "tourists' countries of origin", "tourists' age" and "number of nights spent by tourist". The analysis yielded a chi square of 42.064, 72.230 and 61.473 for respective variables (refer Table 3). It means that there is a significant relationship between tourists' comment on Kelantan with the countries of their origin and age but not for the number of night spent by them in Kelantan. The level of significant is 0.05. Therefore we can conclude that, the main international tourist's behaviour in Kelantan is to enjoy and explore the natural beauty of tropical rain forest that full with uniqueness and distinctiveness of natural treasures. It is not impossible if one day Kelantan will become one of the famous places for eco-tourism in the world.

Table 2: Distribution of tourist by country and next destination

Next Destination	Country					Total
	North America	Europe	Asia	Africa	Oceania	
Perhentian Island	6	65	4	0	6	81
Other island in Malaysia	1	8	1	1	1	12
Kuala Koh National Park	5	39	0	1	5	50
Highland in Malaysia	0	13	0	0	1	14
Kuala Lumpur	1	36	8	0	7	52
Other states in Malaysia	2	39	4	0	4	49
Thailand	2	36	4	2	9	53
Singapore	0	22	0	0	2	24
Other Asia Countries	1	5	2	0	2	10
Going Home	1	16	1	0	2	20
Total	19	279	24	4	39	365

Table 3: Chi-Square Tests on Tourists' Comment

Variable	Value	df	Asymp. Sig. (2-sided)
Countries of origin	42.064	28	.043
Age	72.230	42	.003
Number of night spent	61.473	49	0.109

B. Pareto Analysis

The second objective of this study is to analyse the best way to attract and increase international tourist to Kelantan. By using the Pareto analysis on the variable, "tourists' comment", the result is shown in Figure 1. Based on the Pareto chart, it clearly shows that the tourism sector in Kelantan can be improve if the authorities take part in locating and develop the natural asset that has the potential to become a beautiful tourism sites. Kelantan naturally has its own beautiful natural assets and potentially to be developed as a source of attraction. Although Kelantan has a lot of sources that can be exploited to become a good tourism attractions, but the growth of its tourism sector is still not satisfied. A rapid development should have been taken by responsible parties in order to optimize the output of this sector.

Figure 1: Pareto Chart based on Tourists' Comment
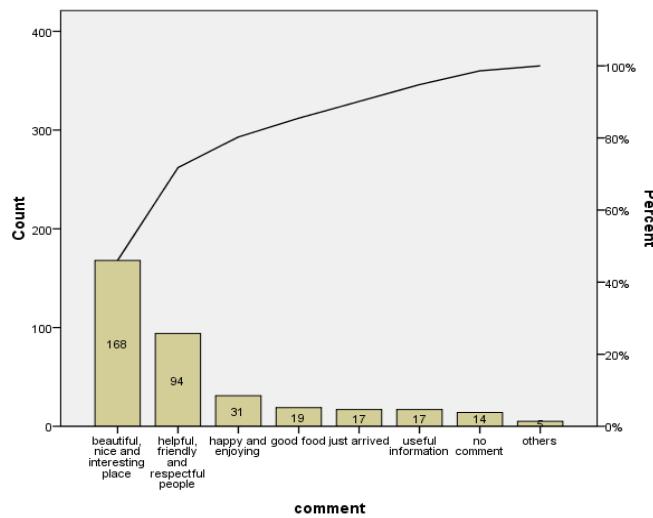
By considering these comments, decision maker in tourism sector could generate 80% of the benefit by doing the whole job. Or in terms of quality improvement, a large majority of problems (80%) are produced by a few key of causes (20%). This is also known as the vital few and the trivial too many. Through this finding, we can conclude that, the best way that to optimize the output of tourism sector in Kelantan is by doing a rapid development in three categories of attractions; (i) natural attractions, (ii) culture, craft and heritage attractions and (iii) shopping arcade attractions. There is an undeniably great synergy between tourism and shopping [4, 9]. Promotion through attractive advertisements regarding the shopping arcade can also help to ensure it popularity and attraction. By using this kind of practice, Kelantan will be better known hence could attract more tourist and obvious increase economic gains. This will also create job opportunities for the local people and increase foreign exchange.

V. CONCLUSION

The objective of this study is to integrate the previous process of the information-transmission and information-diffusion process which categorized as traditional tourism demand models. The study was carried out by applying the cross tabulation, chi-squared test and Pareto analysis on the secondary data. Even though it gives statistical significant results, the model is still has some deficiencies. The estimates parameters identified that the diffusion phenomenon occurs, however the original parameters are not an exchangeable estimators.

REFERENCES

[1] Bhat M., 1996. Outcome of Rehabilitation Programme: Is there a differential response? European Psychiatry, vol 11, supplement 4, 281 – 295.

[2] Burger, C.J.S.C., Dohnal, M., Kathrada, M., & Law, R. (2001). A practitioners guide to time-series methods for tourism demand forecasting a case study of Durban, South Africa. Tourism Management 22(4), 403 – 409.

[3] Chokor B.A., 2003. Pattern of Representation of Countries in Cognitive Sketch Maps of the World with Special Reference to Africa. Journal of Enviromental Psycology, vol 23, pp. 427 – 437.

[4] Jansen-Verbeke, M., 1994. The synergy between shopping and tourism. In: Theobald, W.F. (Ed.), Global Tourism: The Next Decade. Butterworth-Heinemann, Oxford, pp. 428 –446.

[5] Sheldon, P. J. (1993). Forecasting tourism: expenditure versus arrivals. Journal of Travel Research, 32, 13 – 20.

[6] Smith, S. L. J. (1995). Tourism analysis: A handbook. London: Longman

[7] Song, H. & Li, G., (2008). Tourism demand and forecasting – A review of recent research. Tourism Management 29, 203 -220.

[8] Summers P.M. et al, 2004. Tropical Forest Management and Silvi-cultural Practice by Small Farmers in the Brazilian Amazon: Recent Farm –Level Evidence from Randonia. Forest Ecology and Management, vol 19, Nos. 2-3: 161 -177.

[9] Timothy, D.J., 2005. Shopping Tourism, Retailing and Leisure. Channel View, Clevedon, UK.

AUTHORS

First Author – SuliadiSufahani, Senior Lecturer, UniversitiTun Hussein Onn Malaysia, sfsufahani@gmail.com.
Second Author – MarinahMuhammade, Senior Lecture, Universiti Malaysia Kelantan, marinah@umk.edu.my.
Third Author – Zuhaimy Ismail, Professor, UniversitiTeknologi Malaysia, zuhaimyi@yahoo.com.

Correspondence Author –
[a]SuliadiSufahani, sfsufahani@gmail.com
[b]Marinah Muhammad, marinah@umk.edu.my