Title: Socio-Economic Household Data Analysis Using the Clustering and Association Technique for Data Mining

Researcher: Chadarat Phipathananunth
Faculty/Department: School of Science and Technology Department of Information and Communication Technology / The University of the Thai Chamber of Commerce

Sririthorn Jalearnrat
Faculty/Department: School of Science and Technology Department of Computer Science / The University of the Thai Chamber of Commerce

Sasithorn Mongkolsripattana
The University of Chicago - University of the Thai Chamber of Commerce Research Center

Year of Accomplishment: 2013
No. of Pages: 48 Pages

Key words: data mining, K-means algorithm, association rule

Abstract*

In this research, we studies and analyses the data from the Household Socio-Economic Survey 2009 of the Office of National Statistics, Thailand. We use K-means algorithm to cluster the expenditure of the population. By using DB Index and SD Validity, Index, we found that the appropriate number of clusters is three clusters. Then we use association rule technique to determine the relationship between variables. The results show that the association rules are similar among clusters. For example, the average monthly household income is related to the average monthly household cost. Besides, household size is related to number of earners per household. Furthermore, the monthly
tobacco cost for a household is associated with a number of members who are entitled to reimbursement for medical expenses.

*The research was Financially Supported by the University of the Thai Chamber of Commerce