

MULTIDIMENSIONAL DATA ANALYSIS AND KNOWLEDGE MANAGEMENT TECHNOLOGIES FOR TOURIST DECISION SUPPORT

George Stalidis

Alexander Technological Educational Institute
57400 Thessaloniki
stalidgi@mkt.teithe.gr

Dimitrios Karapistolis

Alexander Technological Educational Institute
57400 Thessaloniki
karap@mkt.teithe.gr

ABSTRACT

The aim of this paper is to match multidimensional data analysis methods with knowledge engineering technologies, to support marketing decisions in the area of tourism. Informed decisions in marketing planning should in general be based on market data available as statistics or on primary surveys launched for this purpose. The challenge addressed by the current work is to be able not only to reveal useful information from survey data but also to express the analysis results in the form of knowledge, so that they are maintainable, transferable and usable by non-analysts through computerized decision support tools. The analysis methods employed include Correspondence Analysis, multidimensional Hierarchical Clustering and Discriminant Analysis (Benzecri, 1973, Benzecri, 1992). The analysis results are then utilized, firstly by feeding them to a neural network (Simpson, 1990) which is trained to perform automatic classification and, secondly by extracting knowledge in the form of rules and expressing this knowledge using ontology-based models and rule-based systems (Schreiber, 2008).

The reported work includes the first results of the three-year project “DANKMAN” funded by the program Archimedes III and in particular, the conceptual design of an integrated data analysis and decision support platform for the tourist sector and work in progress on the redevelopment of data analysis software components as well as the first results of a pilot survey in the field of tourism planning. The starting point for the above project is the data analysis software MAD (Karapistolis, 2002), which is a product of the Data Analysis Laboratory of the Department of Marketing, ATEITH. This software implements a set of multidimensional statistical methods in the category of explorative factor analysis, which have been selected as particularly suited to primary questionnaire-based surveys, since they are effective in the analysis of data on qualitative characteristics (Karapistolis, 2008). Another important feature of methods in this family, in particular multiple correspondence analysis, is the multivariate treatment of the data through simultaneous consideration of multiple categorical variables. The multivariate nature of correspondence analysis can reveal relationships that would not be detected in a series of pair-wise comparisons of variables and through the graphical display of row and column points in factorial axes, it is possible to detect structural relationships among the variable categories and statistical units. It is also worth noting that Correspondence Analysis has highly flexible data requirements, as it can be used with frequency data, scales and ratings in heterogeneous datasets and is therefore able to consolidate non-uniform data from a wide range of sources (Greenacre, 2007).

The application area addressed is the tourism sector and in particular to develop decision support tools for destination management organizations of Northern Greece. The aim was to analyze survey data from multiple sources and formulate rules on trends, characteristics of market segments and competition in relation to the profile of each destination. The system would then be able to suggest suitable target market segments by matching the destination profile with visitor profiles and information on current trends. By selecting a specific target market, the marketer may also receive indications of the destination properties that are missing and are found to be related with better positioning in order to increase its competitiveness.

A pilot survey was conducted in order to show the abilities of the proposed approach and to test the initial implementation of the analysis components. The survey was targeted to visitors of Northern Greece and was aimed at identifying distinct visitor profiles, capturing the destination image and perceived value and to establishing a relation among types of visitors, satisfaction and intention of revisiting. Data were collected through a closed-type questionnaire and were coded as a set of categorical and ordinal variables. With the application of Multiple Correspondence Analysis it was possible to identify the main classes of visitors and to associate them with specific viewpoints on destination image. Using multidimensional cluster analysis on the population sample, visitors were clustered into representative groups. At

the next step, a knowledge model was developed, suitable for expressing the analysis results in the form of rules. The model consisted of an ontology that contained the necessary vocabulary and a rule-based component that was used to formulate sets of association rules expressing the relation among visitor classes and factors of destination image, perceived value and satisfaction. The results were reusable knowledge in electronic form consisting of a set of rules explaining the attitudes of several types of visitors towards the surveyed destination.

REFERENCES

- Benzécri, J.-P. (1973). *L'Analyse des Données. Volume II. L'Analyse des Correspondances*. Paris, France: Dunod.
- Benzecri, J.-P. (1992), *Correspondence Analysis Handbook*, New-York: Dekker, P. 1994, *L'analyse factorielle*, Paris : PUF
- Greenacre, M. (2007). *Correspondence Analysis in Practice*, Second Edition. London: Chapman & Hall/CRC.
- Karapistolis, D. (2002). "The MAD software", *Data Analysis Bulletin*, vol 2, pp 133 (in Greek).
- Karapistolis, D. (2008). *DATA ANALYSIS – Statistics without models*, Vol I, Anikoula Publications, Greece.
- Schreiber, G. (2008). "Knowledge Engineering", in: *Handbook of Knowledge Representation*, F. van Harmelen, V. Lifschitz, B. Porter (Eds.), Elsevier, 2008, pp. 929–946.
- Simpson, P.K.(1990). *Artificial neural systems: foundations, paradigms, applications, and implementations*. Pergamon Press, Elmsford, NY.