

APPLICATION OF CLUSTER ANALYSIS WITH REGARDS TO THE PROFILE OF TOURISTS IN THE RURAL AREA OF GJIROKASTRA, ALBANIA

Irena Boboli 

“Eqrem Çabej” University, Gjirokastra, Albania

irena_boboli@yahoo.com

Etleva Dashi Muça

Agricultural University of Tirana, Tirana, Albania

evadashi@ubt.edu.al

Anila Resulaj

University of Tirana, Tirana, Albania

ana.ktona@fshn.edu.al

Abstract

Gjirokastra area is one of the most preferred destinations in Albania that attracts a considerable number of domestic and foreign visitors, because the tourist offer is diversified. Under these conditions, the creation of visitor – consumer profile and the determination of preferences and tourist demand represent an important factor. Information regarding the visitor's preferences and the focus in specific markets will ensure success and development of the destination. Tourist product is a set of goods and services. The demand for this product depends mostly on economic and socio-demographic factors, but also on tourist preferences. In this paper, our goal is to analyze the profile and consumer's behavior of different groups through the cluster analysis of data mining. Several characteristics of the visitors were taken into consideration such as: age, gender, profession, education and preferences for local products. From cluster analysis we identified 6 groups of tourists who were specified according to the aim of the visit

Keywords: *Tourism, rural areas, visitor profile, data mining, cluster analysis*

INTRODUCTION

Gjirokastra is one of the most preferred destinations in Albania and in 2005, became part of the World Cultural Heritage under the protection of UNESCO. It lies in the southern part of Albania, and is bordered by the districts of Tepelene, Permet, Vlora, Saranda and in the south east with Greece with a total area of 1,137.34 km². The space owns a favorable geographical position where the most important area is Drino Valley. This area attracts a considerable number of domestic and foreign visitors, because has many potentials such as historic, natural, cultural, archeological, gastronomic, religious, artisanal, etc. So, the tourism sector is considered as one of the most important potentials to foster economic growth in the long run (Bejko & Bobrati, 2011).

But, to evaluate the tourism effects in rural areas first we have to identify the groups of tourists interested to visit Gjirokastra villages.

Through finding out the category of different tourists, we will classify them in clusters in order to identify future market potentials. This enables them to define common reasons in buying a product within each market sector, same consumer behavior and same reactions in marketing activities. The «categories» of tourists should dispose internal homogeneity relative towards economic factors, whereas it should exist a sensitive diversification among categories towards economic factors also (Bull 2002).

Seen from economic aspect, one of the characteristics of tourist product is the emergence of new and unique kinds of tourism that has led to the segmentation of the market in order to attract tourist / consumer and fulfill his certain needs and wishes (Patsouratis, 2002). Heterogeneity is one of the most important characteristics of the tourist market. However, this is problematic for creating an experience. There is therefore a need to divide the market by means of identifiable variables to segment it more adequately. It is important to understand the characteristics and preferences of visitors, which can be done through market segmentation. The latter holds various benefits for tourism products, as it can lead to an increase in tourist numbers and better use and management of scarce resources (Slabbert & Laurens, 2011). Distinct cultures exist in different societies, therefore tourists are increasingly motivated to consume these local culture differences that are based on the cultural heritage of a certain country, region or city. Inside the destination, the tourist gaze is focused on several attributes, such as countryside and city landscapes that differ from the experience of their ordinary daily life (Marujo, Serra, & Borges, 2012).

Tourism in Gjirokaster should contribute towards a sustainable development of rural communities because the area has numerous natural and cultural wealth that deserve to

valorize. In Gjirokastra District have been evidenced some forms of tourism where the most important are mentioned below: tourism of culture and heritage, rural tourism, nature tourism.

LITERATURE REVIEW

Cluster analysis, is a technique or a group of techniques that are used for grouping different objects (goods, individuals, businesses, farms, etc) in groups in which they either share or do not share any clearly specified similarities (Osmani, 2006).

Cluster analysis can be widely used in the field of economy, finance, marketing. In any case the aim is to simplify the analysis, to draft programmes and development policies, or adequate decisions to be taken which can even be differentiated in accordance with the groups, so that the sources can be used more effectively (Osmani, 2006). Cluster analysis is a tool used in studying analysis data, which aims in classifying different objects in groups so that the degree of relatedness between two objects can be maximal if they belong in the same group and minimum if they belong in different groups. (Karagewrga, 2012).

A cluster analysis is a multivariate technique aimed at identifying market segments. This approach to market segmentation offers several advantages to researchers, such as usefulness in determining market niches based on similarities among tourists, classifying tourist groups, generating hypotheses about these tourist groups and testing a concept to determine whether specific types of customers are present in the dataset, and it is possible to analyse a vast number of respondents effectively (Jurowski & Reich, 2000). Identifying segments has been a strong marketing management tool that caught attention of many practitioners and researchers. Using and benefiting from results of various segmentation studies, destination managers, officers and firm owners are able to better position their destinations which in turn enables them to attract larger numbers of tourists to their destinations and establishments (Birdir, 2015). Finding out why people decide to visit different places and what influences their choice of destination are very important insights for developing suitable marketing strategies and managing market segments (Heung, Qu, & Chu, 2001).

So, to classify the groups of tourists we decided to develop them through data mining model.

Data mining is the efficient discovery of valuable, non obvious information from a large collection of data. (Bigus, 1996). Data mining is the exploration and analysis of large quantities of data in order to discover meaningful patterns and rules (Berry & Linoff, 2004). Data mining serves to discover (hidden, non-trivial) patterns in large amounts of data records in order to be used very effectively for (ex post) analysis and (ex ante) forecasting.

The objective of the data mining process is to discover new, interesting and useful knowledge using a variety of techniques such as prediction, classification, clustering,

association rule mining and sequential pattern discovery (Romero, Ventura, & de Bra, 2009). Data mining techniques such as association rules, clustering, decision trees etc. have been widely used for successfully segmenting and targeting customers across various industries. It provides an effective approach to discover and understand patterns in customer behavior thereby helping the decision maker to better group customers (Juwattanasamran. P., 2013).

The three main uses of data mining techniques in the tourism industry are: (1) forecasting expenditures of tourists, (2) analyzing profiles of tourists, and (3) forecasting number of tourist arrivals (Bose, 2009).

STUDY METHOD

In this paper we will analyze the profile of the visitors-consumers in the rural area of Gjirokastra. In function of this study a questionnaire was developed that helped in gathering the needed information. Through questions addressed at the visitors of the area of study, information was taken related with socio-demographic data, opinions as well as their preferences regarding the tourism development in Gjirokastra. The questioning period was May-September 2015 during which 300 questionnaires were completed. The samples were selected at random and the method used for the realization of the selection was cluster selection in two phases. In the first phase, the sites where the questioning would take place were chosen, and in the second phase a random selection of the visitors in these sites occurred.

The aim of our study was discovering possible groupings and specifying their profiles. For this reason, in order to analyze the gathered data we used data mining as well as the cluster technique.

The algorithm used is kmeans and the measuring distance is Euclidian. Euclidian distance calculates numerical values better and nominal values according to mode.

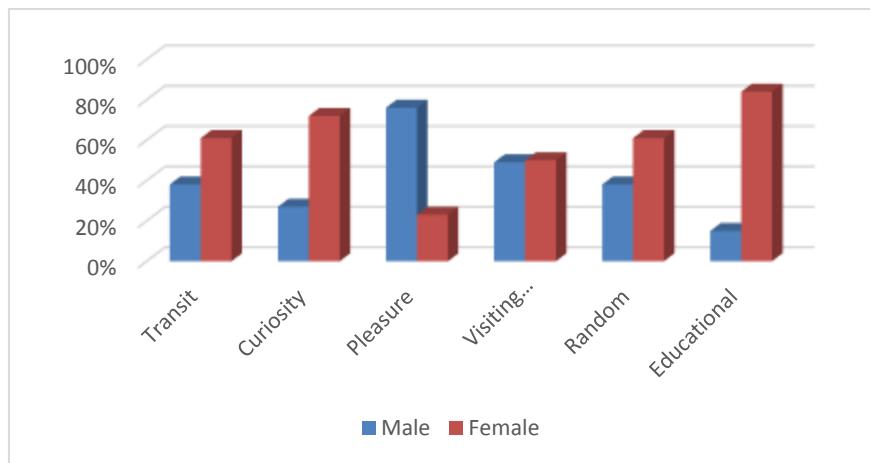
ANALYSIS AND RESULTS

Several characteristics of the visitors were taken into consideration such as: age, gender, profession, education and preferences for local products. The aim is to specify the visitor profile so that the visitor group with the biggest potential in the area could be specified.

From cluster analysis we identified 6 groups of tourists who were specified according to the aim of the visit: Transit, curiosity, pleasure, visiting friends and relatives, random and educational.

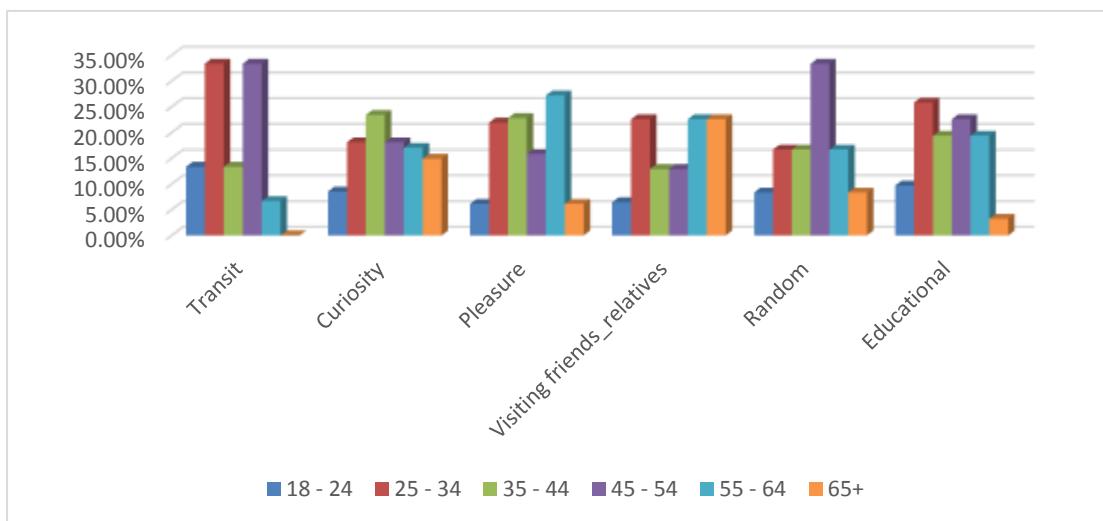
Through cluster group analyze we will identify a visitor profile of Gjirokastra region in order to understand the significant elements between them. At the same time we will know the most profitable tourists for rural communities.

Graph 1: Gender



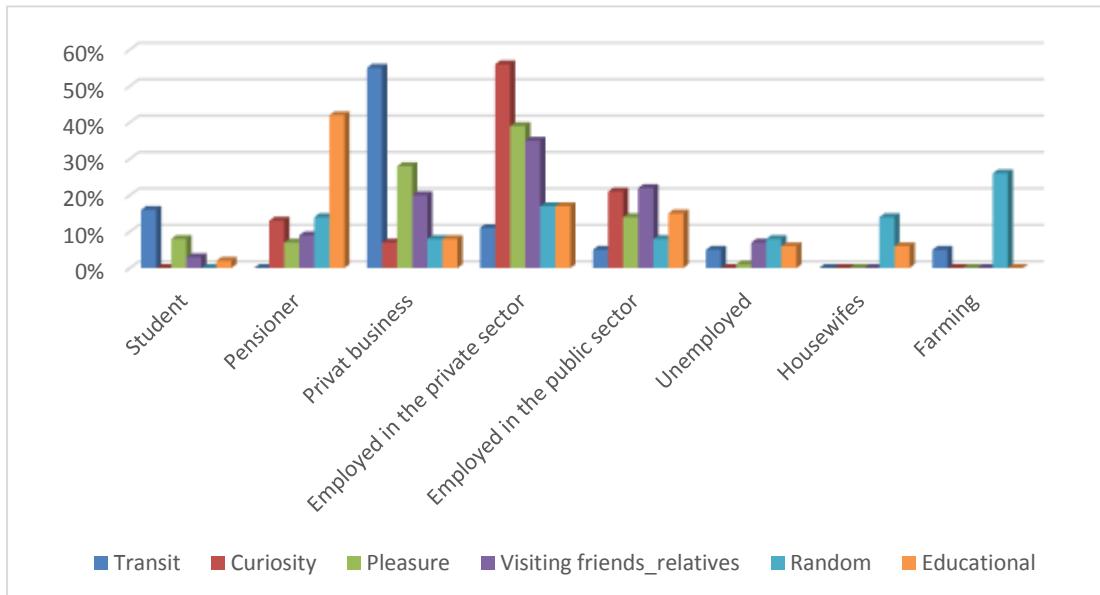
From the graphic 1 we can observe that educational cluster is dominated from women and pleasure cluster is dominated from men.

Graph 2: Age group



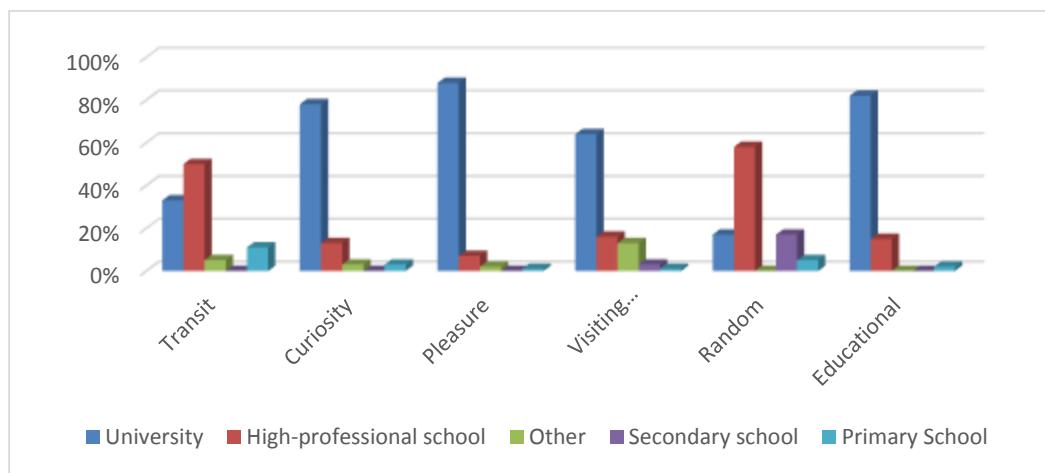
Gjirokastra is an attractive destination for different types of tourism and for this reason is preferred from all groups of ages.

Graph 3: Actual work



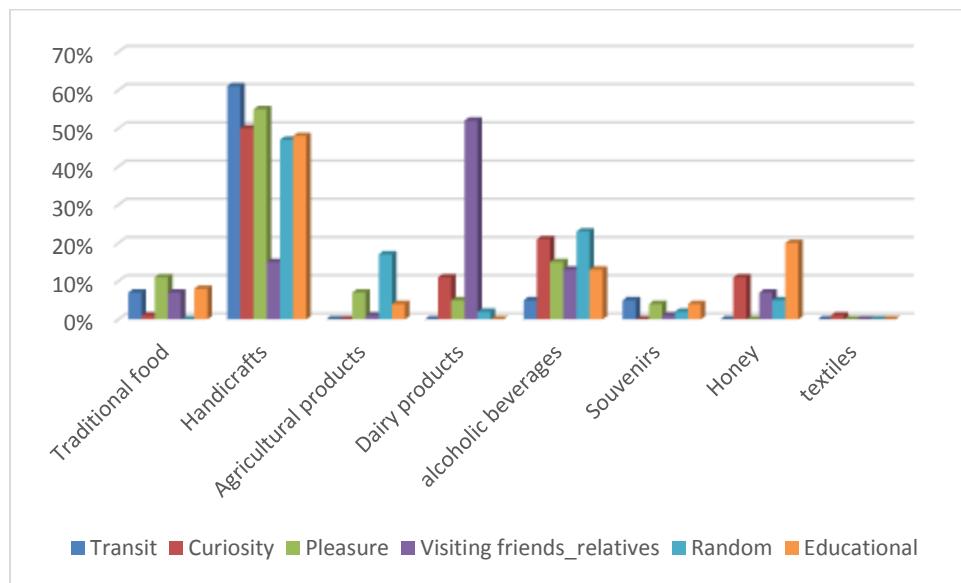
Employment is another characteristic we can mention in cluster analysis. Retired tourists prefer more to visit for educational reasons. Usually they are organized in big groups and make tours in different districts.

Graph 4: Educational level



From graphic 4 we can observe that the major parts of visitors are well educated. Education level is a key factor in consumer behavior.

Graph 5: Demand for local products



Transit cluster characteristics: more females, age: varied from 24 to 54, with secondary education, with private businesses and a preference for handicraft products.

Curiosity cluster characteristics: more females, age: varied from 24 to 64, with university education, employed in the public sector, and a preference for handicraft products and alcoholic beverages.

Pleasure cluster characteristics: more males, age: varied from 24 to 64, with university education, employed in the public section and with private businesses, with a preference for handicraft products.

Visiting friends/relatives cluster: males and females, age: varied from 24 to 64, with university education, employed in both public and private sector, with a preference for dairy products.

Random cluster characteristics: more females, age: varied from 24 to 64, with secondary education, employed in agriculture and public sector, with a preference for handicraft products.

Educational cluster characteristics: more females, age: varied from 24 to 64, with university education, pensioners, with a preference in handicraft products and honey.

DISCUSSIONS AND CONCLUSIONS

Gjirokastra zone is one of the most preferred destinations in Albania which attracts a considerable number of domestic and foreign visitors and where touristic supply is diversified. In the meantime knowing the characteristics and the preferences of the visitors is important. Knowing the profiles of the visitors-consumers will fulfill their needs and their desires better, will

raise the number of visitors and simultaneously will use the touristic potentials of the destination better. Tourists visit this district since (i) it is considered part of the world heritage by UNESCO, (ii) for its suitable geographical position (iii) cultural and natural monuments (iv) historic and cultural reasons, and (v) traditional food.

REFERENCES

- Ministria e Zhvillimit Urban dhe Turizmit . (n.d.). Draftstrategjia Kombëtare për Turizmin 2014-2020.
- Bejko, A., & Bobrati, M. (2011). The impact of cultural tourism in the development of Gjirokastra-the UNESCO town. European Scientific Journal , Vol 13 .
- Berry, M. J., & Linoff, G. S. (2004). Data Mining Techniques For Marketing, Sales, and Customer Relationship Management (Second Edition ed.). Indianapolis, Indiana: Wiley Publishing.
- Bigus, J. (1996). Data Mining With Neural Networks: Solving Business Problems from Application Development to Decision Support. N.York, Usa: McGraw-Hill.
- Birdir, S. (2015). Segmentation of Tourist Using Demographic and Travel. International Review of Management and Marketing, 5(4), 221-229.
- Bose, I. (2009). Data Mining in Tourism. The University of Hong Kong, Hong Kong, IRMA , pp. pp 936-940.
- Bull, A. (2002). Τουριστική Οικονομία (2nd edition ed.). Αθήνα: Κλειδάριθμος.
- Dudumi, A. (2001). Dinamika e sistemit të hapësirës rurale të rrëthit të Gjirokastrës .
- Heung, V., Qu, H., & Chu, R. (2001). The relationship between vacation factors and socio-demographic and travelling characteristics: The case of Japanese leisure travellers. Tourism Management, 22 , pp. 259-269.
- Jurowski, C., & Reich, A. (2000). An Explanation and Illustration of Cluster Analysis for Identifying Hospitality Market Segments . Journal of Hospitality and Tourism Research, 24 (1), 67-91 .
- Juwattanasamran. P., S. S. (2013). Applying Data Mining to Analyze Travel Pattern in Searching Travel Destination Choices. The International Journal Of Engineering And Science (IJES), Vol. 2 , Issue 4,ISSN(e): 2319 – 1813 ISSN(p): 2319 – 1805 , pp. Pages 38-44.
- Karagewrga, I. (2012). Ανάλυση Συστάδων (Cluster Analysis).
- Marujo, N., Serra, J. M., & Borges, M. d. (2012). VISITORS TO THE CITY OF ÉVORA: WHO ARE THEY? European Journal of Tourism, Hospitality and Recreation, Vol. 3, Issue 2 , pp. pp. 91-108.
- Niarou, K.-M. P. (2011). Ανάλυση Γονιδιακών Εκφράσεων μέσω της Μεθόδου Real-Time RT PCR.
- Osmani, M. (2015). Metodat e Statistikës per Ekonomistët. Tiranë: Pegi.
- Osmani, M. (2006). Metodat Statistikore Multivariate. Tiranë: Blendi.
- Rogo, N., Qiriazi, P., & Ruspi, G. (2008). Monumentet e natyres se rrëthit Gjirokaster. Gjirokastër.
- Romero, C., Ventura, S., & de Bra, P. (2009). Applying Web usage mining for personalizing hyperlinks in Web-based adaptive educational systems. Computers & Education. 53(3) pp. pp. 828–840.
- Sinani, A., & Duri, V. (2014). Oferta kulturore turistike në Qarkun e Gjirokastrës. Shkenca Shoqërore, Ekonomike dhe të Edukimit, Nr 36 .
- Slabbert, E., & Laurens, M. (2011). Motivation-Based Clustering:An Empirical Study of Vizitors to Kruger National Park in South Afrika. Book of Proceedings Vol.II- International Conference on Tourism & Management Studies-Algarve 2011 , pp. 1122-1126.
- Πατσουράτης, Β. Α. (2002). Η Ανταγωνιστικότητα του Ελληνικού Τουριστικού Τομέα. Ινστιτούτο Τουριστικών Ερευνών και Προβλέψεων (ΙΤΕΠ), Μελέτες: 12, , pp. σελ. 35-36.