

## **USING DECISION TREES TO IDENTIFY TOURISM STAKEHOLDERS BASED ON LEVEL OF PARTICIPATION IN TOURISM AND COMMUNITY POLITICAL ACTIVITIES**

Erick T. Byrd PhD

*Department of Recreation, Tourism, and Hospitality Management, The University of North Carolina at Greensboro*

Larry Gustke PhD

*Department of Recreation, Tourism, and Hospitality Management, The University of North Carolina at Greensboro*

Follow this and additional works at: <https://scholarworks.umass.edu/ttra>

---

Byrd, Erick T. PhD and Gustke, Larry PhD, "USING DECISION TREES TO IDENTIFY TOURISM STAKEHOLDERS BASED ON LEVEL OF PARTICIPATION IN TOURISM AND COMMUNITY POLITICAL ACTIVITIES" (2016). *Travel and Tourism Research Association: Advancing Tourism Research Globally*. 42. [https://scholarworks.umass.edu/ttra/2007/Presented\\_Papers/42](https://scholarworks.umass.edu/ttra/2007/Presented_Papers/42)

This is brought to you for free and open access by ScholarWorks@UMass Amherst. It has been accepted for inclusion in Travel and Tourism Research Association: Advancing Tourism Research Globally by an authorized administrator of ScholarWorks@UMass Amherst. For more information, please contact [scholarworks@library.umass.edu](mailto:scholarworks@library.umass.edu).

# Using Decision Trees to Identify Tourism Stakeholders Based on Level of Participation in Tourism and Community Political Activities

Erick T. Byrd, PhD

Department of Recreation, Tourism, and Hospitality Management  
The University of North Carolina at Greensboro

Larry Gustke, PhD

Department of Parks, Recreation and Tourism Management  
North Carolina State University

## Abstract

The purpose of this paper is to investigate the use of decision tree analysis in the identification of stakeholders who participate in and who do not participate in tourism and political activities in a community. Decision tree analysis is a tool for partitioning a data set based on the relationships between a set of independent variables and a dependent variable. The research reported here tests the application of, decision tree analysis, an analytical technique that is not traditionally used to segment stakeholders in tourism. Based on the results of the decision tree analysis four groups were identified: high participants, high-moderate participants, low-moderate participants, and low participants.

## Introduction

A stakeholder can be defined as “any group or individual who can affect or is affected by the achievement of the organizations objectives” (Freeman 1984:46), and that group or individual has a legitimate interest in the organization (Donaldson and Preston 1995). Many authors have argued for the need to include stakeholders in tourism development (Cottrell 2001; Davis and Morais 2004; De Lopes 2001; Hassan 2000; Long and Allen 1990; Murphy 1983; 1985). To include stakeholders in tourism development planning they must be identified and their interests and needs understood.

A way to accomplish stakeholder group identification is to apply the concept of segmentation. Simply, segmentation is the dividing of groups (markets, stakeholders) into sub-groups based on specific characteristics (Kotler, Bowen and Makens 2003). Segmentation assumes people are different and that the differences are related to a specific behavior or attitude. Based on the differences people can be grouped into segments (MacKay, Andereck and Vogt 2002).

Traditionally in the tourism industry stakeholder groups have been segmented in a community based on geographic (Cottrell 2001; Davis and Morais 2004; Murphy 1983), demographic (De Lopes 2001; Hassan 2000; Ryan 2002), and socio-graphic characteristics (Murphy 1985). Planners identify groups based on characteristics and then develop and implement strategies to include the different groups in the planning process. There are inherent biases to these approaches that limit their usefulness. The major bias is that planners make subjective judgments on who and what groups are included to represent stakeholders.

Owen (1998) in his review of the market segmentation research and literature concluded that there is a need for new approaches to segmentation. Arimond and Elfessi (2001) echoed this need. This need also holds true for stakeholder identification in a community. Owen states that the new techniques should be conducted using computer based software to lessen the “bias of

human judgment”. Arimond et al (2001) indicated new segmentation approaches should allow for simpler data collection and survey methods. Also, these new approaches should provide graphic displays that reveal the relationships inherent in the data.

One approach that addresses Owen’s (1998) and Arimond et al (2001) thoughts for new segmentation is the use of decision tree analysis. A decision tree is a model that can be used to classify or predict variables (Kass 1980; Koskela 2003). It is a method used to study the relationship between a dependent variable and multiple independent variables (Huba 2003). Decision trees are designed to handle a large number of independent variables at differing levels of measurement. Decision trees identify groups by dividing the dependent variable by the independent variables. The results of the decision tree analysis indicate which independent variables are most strongly related to the dependent variable. Decision tree analysis has been used in financial settings for credit scoring, manufacturing settings for quality control, and healthcare settings for determining treatments (SPSS 2002).

The purpose of this research paper is to use decision tree analysis in the identification of stakeholders who participate in and who do not participate in tourism and political activities in a community. Decision tree analysis is a tool for partitioning a data set based on the relationships between a set of independent variables and a dependent variable. The research reported here tests the application of, decision tree analysis, an analytical technique that is not traditionally used to segment stakeholders, but has been used in health care, market analysis, credit scoring and policy studies.

## **Methods**

The study described in this paper was conducted in the spring of 2003 in two rural North Carolina counties, Johnston County and Martin County. Stakeholders, which included the county residents and visitors, were mailed a questionnaire inquiring about their attitudes and perceptions of tourism development, their participation in local political and tourism activities, and their view of the environmental impacts of tourism. The responses were collected and analyzed to develop an understanding of what variables influence participation in tourism and community political activities and to identify stakeholder groups based on this participation.

To accomplish this, an index for participation in tourism and community political activities was developed from seven variables: visited the local visitor’s center, attendance at local festivals, visiting local attractions, voting in last general election, attendance at a city or county council meeting, ran for an elected office, and membership in the local chamber of commerce. An additive index was created from these variables to reflect an individual’s level of participation in tourism and community activities.

The initial analysis of the data produced descriptive statistics. Further analysis of the responses was conducted using SPSS Answer Tree. SPSS Answer Tree 3.1 is a statistical application that uses algorithms to develop decision trees. A decision tree is a model that can be used to classify or predict variables (Kass, 1980; Koskela, 2003). It is an exploratory method used to study the relationship between a dependent variable and independent variables (Huba, 2003). Decision trees indicate how groups develop by dividing the dependent variable by the independent variables. The specific statistics that are used in a decision tree are based on the algorithm that is employed. This study used the Exhaustive Chi-square Automatic Interaction Detection (CHAID).

The Exhaustive CHAID is a second-generation CHAID algorithm used in SPSS Answer Tree (Huba, 2003). The Exhaustive CHAID splits the data into subsets that best describe the

dependent variable (Kass, 1980). SPSS extended the capabilities of CHAID so that it can include nominal, ordinal, interval, and ratio variables (Huba 2003). CHIAD splits the data into subsets that best describe the dependent variable (Kass 1980). For nominal and ordinal variables, chi-square analyses are used, and for interval and ratio variables an analysis of variance is used (Huba 2003). The split is based on which variable has the lowest p value. If a tie occurs between two or more variables the variable that has the highest F value is selected as the predictor.

The strength of the CHAID method is found when attempting to detect patterns in large datasets that are comprised of differing levels of measurement. CHIAD can handle independent and dependent variables at all levels of measurement. Also, not all independent variables have to be at the same level of measurement. In other words if a dataset was comprised of independent variables that were nominal, categorical, interval and ratio the CHIAD decision tree analysis can include all the variables in the analysis (Huba, 2003).

## Findings

The Participation in Tourism and Political Activities Decision Tree (see Figure 1) was developed using the Exhaustive CHAID method in SPSS Answer Tree. The resulting tree can be divided into four main branches and 9 levels. Branch 1 consists of nodes 1, 6, 7, 14, 15, 16, and 17. Branch 2 consists of nodes 2, 8, 9, 10, 18, 19, 20, 21, 24, 25, 26, 27, 28, 29, 30, 33, 34, 35, 36, 37, 38, 39, 40, 41, and 42. Branch 3 consists of node 3. Branch 4 consisted of nodes 4, 11, 12, 13, 22, 23, 31, and 32. Node 5 consists of respondents that could not be identified as a specific perspective due to missing data.

The decision tree indicated that the following variables influence a stakeholder's participation in tourism and political activities: respondent's role in tourism (governmental official, resident, business owner, visitor), respondents' perception of impacts, length of residency, general demographics (age and gender) and their participation or lack of participation in recreational activities (bird watching, hiking, fishing, and attending festivals).

Four groups can be identified on Participation in Tourism and Political Activities Decision Tree. The four groups identified on the Participation in Tourism and Political Activities Decision Tree, were described as high participants, high-moderate participants, low-moderate participants, and low participants. Each group indicated some level of participation in tourism and political activities.

Branch 1 (see Figure 1) can be classified as *high participants* based on a mean score of 10.56 on the participation index at the main split (level 2). Mean participation index scores for the high participants ranged from 8.24 to 12.56. High participants account for 9.07% of the respondents. Variables that had an influence on this group were: their perspective, how long they had lived in the community, gender and activities they participated in.

Branch 2 (see Figure 1) can be classified as *low moderate participants* based on a mean score of 3.18 at the main split (level 2). Mean participation index scores for the low moderate participants ranged from a 0.40 to a 2.98. Low moderate participants account for 60.08% of the respondents. Variables that had an influence on this group were: their perspective, their perceived impact of tourism, how long they had lived in the community and activities they participated in

Branch 3 (see Figure 1) can be classified as *low participants* based on a mean score of 2.30 at the main split (level 2). Low participants account for 11.49% of the respondents. There are no identifiable variables that had an influence on this group.

Branch 4 (see Figure 1) can be classified as *high moderate participants* based on a mean score of 5.03 at the main split (level 2). Mean participation index scores for the low moderate participants ranged from a 3.22 to a 6.84. High moderate participants account for 18.15% of the respondents. Variables that had an influence on this group were: their perspective, their perceived impact of tourism, age, and their

These four groups can be separated into subgroups and those subgroups could be further separated resulting in a total of 42 groups. Each group is differentiated by multiple variables. For example a group that can be sub divided from the high participants would be stakeholders who are government officials, have lived in the area for over 50 years and fish.

### **Application of Results**

The results from this study can assist tourism planners in identifying stakeholder groups that will participate in tourism and political activities. With this knowledge tourism planners can identify which stakeholder groups will be the most influential and vocal in a community in regards to tourism development. This study identified four stakeholder groups in relation to participation in tourism and political activities: high participants, high-moderate participants, low-moderate participants, and low participants. As previously stated each group indicated some level of participation in tourism and political activities.

High participants are generalized as stakeholders who were government officials or were in governmental positions. It is intuitive in that the officials would be high participants in political activities. Of interest is that officials are active in tourism related activities. It can be inferred by this information that government officials are familiar at some level with tourism opportunities and issues in their community. Government officials will be involved in some way in most major decision in the county that is related to sustainable tourism development. It also indicates the importance of governmental officials in the tourism development process.

Low Moderate participants are generalized as stakeholders who were residents that did not participate in many natural resource based activities in 2002, and perceive that tourism's impact on the community is positive. Even though this group is low moderate participators they are extremely important to the sustainable tourism process. This is due to their size and their interaction with visitors. Therefore, this group should be encouraged to participate more in both tourism activities and political activities. This group will have little influence on what type of tourism development policies are developed, but will influence the success of these policies and will be affected by these policies.

Low participants can be generalized as stakeholders who are visitors to the local community. It is important to understand that even though the visitors are considered low participators and will have limited to no influence on policy development, they will as the residents be impacted by the policy and will be extremely influential in the success of any sustainable tourism process. Therefore, it is important to obtain their concerns and wants and account for these in the tourism development process.

High Moderate participants can be generalized as stakeholders who were business owners and perceived tourism has a positive impact on the community. It can be inferred that this group sees many of the economic benefits of tourism and therefore has positive perceptions of tourism. As high moderates this group of stakeholders will be influential in many of the policies that are developed related to tourism. By being able to identify these groups, tourism professionals can be better prepared to work with these groups and verify that those who are active in the community are involved in the tourism development of the community.

## **Conclusions**

The use of decision trees in tourism is a new way to approach many concepts that are integral to tourism such as marketing, planning and creating and maintaining quality of life. Decision trees have been used in the past with patient research such as, *Quality of Life in HIV/AIDS Patients* and *Patient Satisfaction with Management Care* (The Measurement Group, 2003), and marketing research such as consumer profiling (Eherler D. and T. Lehmann, n.d.). Decision tree analysis can be used to identify stakeholder groups in a community. Unlike the traditional method of stakeholder mapping that relies simply on the researchers or planners' knowledge and experiences, the decision tree analysis is grounded in statistical analysis and is a more robust tool for stakeholder identification and management.

Decision trees are a parsimonious technique that can facilitate interpretation and understanding for tourism professionals about many tourism concepts and behaviors. The parsimonious nature of decision trees is based on the descriptive and visual nature of the output. Variable interaction can be visually shown to tourism professionals such as DMO executives, state travel directors and community planners which facilitate the explanation and application of the results of research.

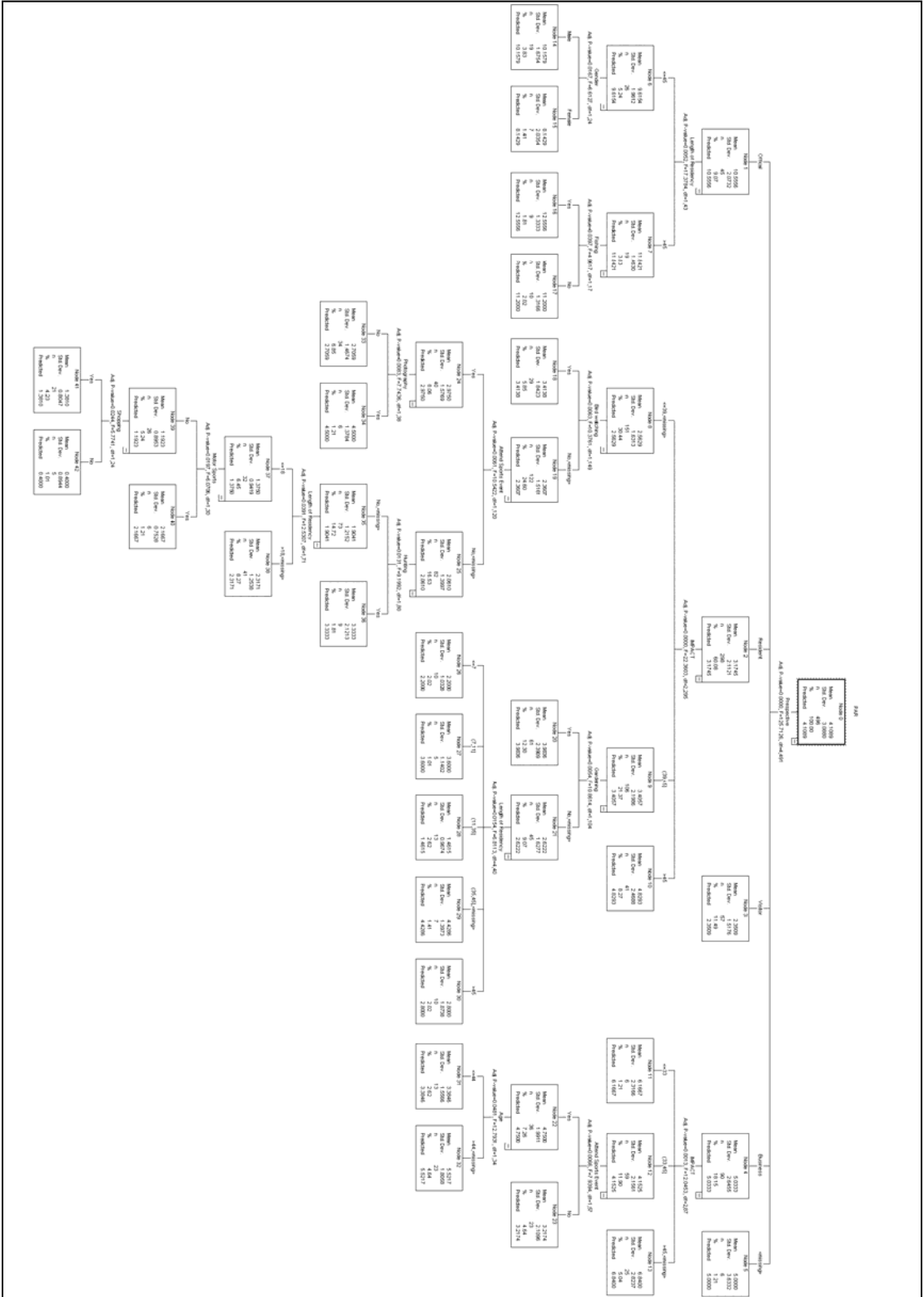


Figure 1: Participation in Tourism and Political Activities Decision Tree

## References

- Arimond, G. and Elfessi, A. (2001) A clustering method for categorical data in tourism market segmentation research, *Journal of Travel Research*, 39, 391-397.
- Cottrell, S.P. (2001) A Dutch international development approach: Sustainable tourism development, *Parks and Recreation*, 36, 9, 86-93.
- Davis, J.S. and Morais, D.B. (2004) Factions and enclaves: Small towns and socially unsustainable tourism development, *Journal of Travel Research*, 43, 3-10.
- De Lopez, T.T (2001) Stakeholder management for conservation projects: A case study of Ream National Park, Cambodia, *Environmental Management*, 28, 1, 47-60.
- Donaldson, T. and Preston, L.E. (1995) The stakeholder theory of the corporation: concepts, evidence, and implications, *The Academy of Management Review*, 20, 1, 65-91.
- Eherler, D. and Lehmann, T.(n.d.) Responder Profiling With CHIAD and Dependency Analysis. Institute of Applied Economic Research, Limburgs Universitair Centrum, Retrieved May 30, 2003 from <http://www.luc.ac.be/iteo/articles/lehmann.pdf>
- Freeman, R.E. (1984) *Strategic Management: A Stakeholder Approach*, Boston: Pitman.
- Hassan, S.S. (2000) Determinants of market competitiveness in an environmentally sustainable tourism industry, *Journal of Travel Research* 38, 3, 239-245.
- Huba, G.J. (2003) *CHAID*. The Measurement Group, Retrieved May 30, 2003 from <http://www.themeasurementgroup.com/Definitions/CHAID.htm>
- Kass, G.V. (1980) An exploratory technique for investigating large quantities of categorical data, *Applied Statistics*, 29, 2, 119-127.
- Koskela, T. (1999) Decision Trees, Retrieved May 30, 2003 from <http://www.hut.fi/u/timoko/treeprogs.html> (accessed May 30, 2003)
- Kotler, P., Bowen, J. and Makens, J. (2003) *Marketing for Hospitality and Tourism* (3rd ed.), Upper Saddle River, NJ: Prentice-Hall.
- Long, P., Perdue, R. and Allen, L. (1990) Rural residents perceptions and attitudes by community level of tourism, *Journal of Travel Research*, 28, 3, 3-9.
- MacKay, K.J., Andereck, K.L., and Vogt, C.A. (2002) Understanding vacationing motorist niche markets, *Journal of Travel Research*, 40, 356-363.
- The Measurement Group (2003) *Data Mining and Statistical Modeling*, Retrieved June 15, 2003 from <http://www.themeasurementgroup.com/datamining.htm>



Murphy, P.E. (1983) Perceptions and attitudes of decision making groups in tourism centers, *Journal of Travel Research*, 21, 3, 8-12.

Murphy, P.E. (1985) *Tourism: A Community Approach*, London: Methuen.

Owen, J.T. (1998) Market segmentation in hospitality research: No longer a sequential process, *International Journal of Contemporary Hospitality Management*, 10, 7, 289-296.

Ryan, C. (2002) Equity, management, power sharing and sustainability – Issues of the ‘New Tourism’, *Tourism Management*, 23, 17-26

SPSS (2002). *Introduction to Answer Tree*, Chicago, IL: SPSS.