

An Application of the CHAID Algorithm to Study the Environmental Impact of Visitors to the Teide National Park in Tenerife, Spain

Flora M^aD íz-P érez¹ & M^aBethencourt-Cejas²

¹Universidad de La Laguna, La Laguna, Tenerife, Canary Islands, Spain

²Universidad de La Laguna, La Laguna, Tenerife, Canary Islands, Spain

Correspondence: Flora M^aD íz-P érez, Universidad de La Laguna, 38071 La Laguna, Tenerife, Canary Islands, Spain.

Received: May 8, 2017

Accepted: June 19, 2017

Online Published: June 26, 2017

doi:10.5539/ibr.v10n7p168

URL: <https://doi.org/10.5539/ibr.v10n7p168>

Abstract

The significant and complex relationship between visitor numbers to a national park and the environment calls for appropriate policies to be adopted. This paper analyzes the relationship from the perspective of visitors to the Teide National Park (TNP) in Tenerife, aiming to establish strategies to reduce visitors' environmental impacts. This is particularly important as the TNP, with over 3,000,000 visitors in 2015, is the most visited park in Spain and one of the most visited in Europe. An empirical study was conducted during 2016 resulting in 805 valid questionnaires. A CHAID algorithm was then applied to segment visitors according to criterion variables. Findings show the first segmenting variable is transport type, with the car being the most frequently used by visitors. Specifically, the visitor segment coming by car is also associated with the longest stays in the TNP. Regarding the practical and social implications, it is assumed the longer the stay, the greater the environmental impact. These results highlight the need for new transport strategies for the park with improved, less polluting vehicles.

Keywords: transport and environmental policies, CHAID algorithm, market segmentation, socioeconomic characteristics

1. Introduction

Economic policies in the tourism sector have to face the contradiction between the search for greater economic profitability and, on the other hand, the conservation of natural heritage and the sustainability of natural resources associated with tourist destinations. An example of a natural resource that needs adequate protection for its use and enjoyment, not only by the present generation, but also by future ones, is that of a National Park (NP).

In Spain, there are a total of 15 National Parks (Ministry of Agriculture, Food and Environment, 2016). These 15 parks received 14,429,535 visitors (INE) in 2015: an increase of 50.1% compared to 2010 (in the midst of the economic crisis). However, the real increase is lower considering that in 2015 there was a new NP created. This new park is the Sierra de Guadarrama (Castilla Le ón and Madrid), and in 2015, it received 2,989,556 visitors. Adjusting for the new NP, the real increase stands at 19%, which is still a significant number from the perspective of the possible environmental impact. The increase in the number of visitors may also be linked to the gradual but constant economic recovery, after many years of recession.

Problems arising from the negative impacts of the volume of visitors on parks' natural resources are undoubtedly very diverse, given that not all territories have the same environmental load capacity. Island territories, for example, are more fragile and suffer greater deterioration (Ramdas & Mohamed, 2014; Zubair, Bowen & Elwin, 2011). The Canary Islands Autonomous Region, for example, consists of seven islands and contains four of the fifteen National Parks (Caldera de Taburiente, Garajonay, Teide and Timanfaya) located on different islands: La Palma, La Gomera, Tenerife and Lanzarote, respectively. The total number of visitors received by the four Canary Island parks amounts to 6,219,058: nearly half the total number of visitors to all Spanish NPs. As the total number of visitors is not an homogeneous group, the first step in any analysis is to segment the complete group of visitors with the aim of determining which of these segments generate the greatest environmental impact.

The study of market segmentation has been traditionally undertaken by resorting to regression methods. Nevertheless, the need for a significant number of segments and qualifying variables has led to the use of other procedures of multivariate analysis (D'Áz-Pérez & Bethencourt-Cejas, 2016; Legohérel, Hsu & Daucé 2015; Nicholson & Pearce, 2000). One such segmentation technique is CHAID (Chi-square Automatic Interaction Detection). This market segmentation technique is more sophisticated than other multivariate analysis techniques (McCarty & Hastak, 2007) but has rarely been used in this field. This is despite CHAID having very important advantages, since it does not involve the restrictive principle of parametric tests for predictive variables.

Considering the previous arguments, the originality and value of this paper is the study of the environmental impact of visitors based on the segmenting power of socioeconomic characteristics. This study experiments with the technique of analysis based on algorithms: CHAID (Chi Square Automatic Interaction Detection). This technique is novel both in the area of tourism market segmentation and for its use to analyze the demand associated with NPs in island territories. In this case, total number of visitors was classified using the CHAID algorithm and the criterion variable selected was duration of stay in the park, under the assumption that the longer the stay, the greater the environmental impact (Jim, 2000).

2. Background

2.1 Length of the Stay and the Environmental Impact

2.1.1 Length of Stay and the Environment

Many studies have demonstrated the importance of the length of stay on income generated by tourists in destination (Alegre, Mateo & Pou, 2011); however, less attention has been paid to the environmental impact of a longer stay. This is especially relevant for small island destinations as resources are limited, and the natural environment is fragile. Therefore increased numbers of tourist arrivals can put pressure on limited resources such as water and land availability, beyond the carrying capacity threshold limits of islands, thereby jeopardizing sustainability (Holden, 2000; UNEP, 1999). Moreover, on small islands there is often a high degree of endemism and biodiversity but, conversely, relatively small species numbers, and consequently a high risk of extinction among flora and fauna (Zubair, Bowen & Elwin, 2011).

The information obtained from the analysis of destination demand is a very important issue when the researcher is interested in determining which market segment causes the least stress on local resources. In this respect, a core issue in this context is carrying capacity analysis (Saarinen, 2006). Carrying capacity has been generally defined as *the maximum number of people who can use a site without any unacceptable alteration in the physical environment and without any unacceptable decline in the quality of the experience gained by tourists* (Mathieson and Wall 1982, 2). The relevance of carrying capacity should, however, be considered in a relative way. In fact, this concept of carrying capacity is based on different perspectives and opinions concerning nature and culture and their uses as resources, thus resulting in many different definitions. As established in the literature, the number of tourists or the time they spend in a particular destination reach the threshold level when human values and (changing) perceptions concerning the resources indicate that the maximum level has been reached (Hughes and Furley 1996; Odell, 1975).

2.1.2 Length of Stay and Destination Demand Determinants

In the literature, a renewed interest in studying the role of length of stay as a determinant of tourism destination demand has appeared (Alegre & Pou, 2007; Gokovali, Bahar & Kozak, 2007; Gomes de Menezes, Moniz & Cabral, 2008; Pestana Barros & Correia, 2007; Alegre, Mateo & Pou, 2011; Pestana Barros & Pinto Machado, 2010; Thrane, 2012). However, it has been many years since the initial studies were conducted (Mak, Moncur, and Yonamine, 1977; Mk & Nishimura, 1979; Bell & Leeworthy, 1990) and Fleischer & Pizam, 2002).

Economic theory generally treats the length of stay as a constraint on demand imposed by available time, but not by destination demand. However, other studies such as those carried out by Thrane, (2012) show that nationality, age, spending patterns and other trip-related characteristics are associated with length of stay. Moreover, no previous research has used a decision tree model to segment a NP natural resource, taking advantage of a criterion method from a destination management perspective.

2.2 The Tourism Market Segmentation

Several authors have emphasized the desirability of combining different market strategies to capture different segments; (Cook & Mindak 1984, Kardes 2002, Mok & Iverson 2000, Rhim & Cooper 2005, Solomon, Bamossy & Askegaard, 2002). These authors confirm the need to advance in the study of the most suitable segmentation techniques. The ultimate objective of previous studies has been to find as many market segments as possible, statistically speaking. This was, for example, the intention of authors researching tourism market

segmentation that have used tourist expenditure as a segmentation variable. Some of the first papers were written by LaPage (1969) and Styne & Mahoney (1980); however, they did not have much success from the perspective of the clear identification of different groups of tourists based on expenditure. However, more recent studies (Díaz-Pérez & Bethencourt-Cejas 2016; Díaz-Pérez, Bethencourt-Cejas & Álvarez-González 2005; Legohérel 1998; Legohérel & Wong 2006; Spotts & Mahoney 1991) have obtained precise information on the composition and characteristics of homogeneous groups of tourists according to their level of expenditure.

2.2.1 Chi-square Automatic Interaction Detection (CHAID) or Genetic Algorithm

Although the CHAID algorithm segmentation procedure, itself, was first introduced by Kass in 1975, in general, it has been rarely used in market segmentation. In addition, in relation to tourism market segmentation, researchers have used two types of analysis: *a priori* (in origin) or *post hoc* (when leaving destination). Often both *a priori* and *post hoc* analysis have had a descriptive nature. CHAID algorithms, on the other hand, are based on a criterion variable with two or more categories, which allow researchers to determine segmentation with respect to that variable and according to a combination of independent predictors (Chen 2003; Díaz-Pérez, Bethencourt-Cejas & Álvarez-González 2005; Legohérel, Hsu & Daucé 2015). However, the lack of knowledge and complexity of models (McCarty & Hastak 2007) has limited their use to relatively few cases of tourism market segmentation (Chen, 2003; Chung, Oh, Kim & Han 2004; Diaz-Perez, Bethencourt-Cejas & Alvarez-Gonzalez 2005; Diaz-Perez & Bethencourt-Cejas, 2016).

It is important to highlight some of the strengths of CHAID analysis as a method of tourism market segmentation. These strengths can be summarized based on the following characteristics: 1) Chi-square is a nonparametric statistic, thus any form of variable distribution is accepted; 2) both nominal and interval variables can be included in the model as independent variables (predictors); 3) continuous variables can be chosen as criterion variables, since they can be dichotomized and 4) the criterion variable will be selected according to the objectives of destination operators, a characteristic that increases the model's efficiency.

In addition, when comparing CHAID with other non-criterion methods, such as cluster analysis, we observe greater efficiency in terms of the number of variables and the amount of data contained in the former. In this sense, we appreciate, for example, how CHAID algorithms allow the classification of new cases observed in mutually selective, i.e. non-overlapping segments, which means that each element is contained in a single segment (Kass, 1980).

2.2.2 Choice of the Criterion Variable

CHAID analysis has been used in the study of tourism markets since 2000 with the goal of obtaining diverse results including: the identification of preferences when choosing hotel establishments using demographic variables (Chung, Oh, Kim & Han 2004), or clarification of preferences in hotel and restaurant choices (Legohérel, Hsu & Daucé 2015). It has also been used to obtain information on expenditure levels using both demographic variables and those related to the trip's characteristics (Díaz-Pérez, Bethencourt-Cejas & Álvarez-González, 2005) and to describe spending habits (Legohérel & Wong, 2006). Some studies have used it to identify future recommendations using product satisfaction, price and quality of service as independent variables (Chen, 2003); to know the likelihood of returning (Assaker & Hallak, 2012; Hsu & Kang, 2007) and to obtain information on intentions to recommend and visit the destination in the future (Vassiliadis, 2008).

However, we did not find in the literature any study that segmented the tourist markets by duration of stay, in this case on a visit to the TNP. What is more, if we look at the use of CHAID algorithms as a segmentation method, it is even more difficult to find studies applied to the whole industry and, of course, we do not find any referring to the segmentation of national parks.

2.4 Contribution of the Study Corresponding to This Research

In this paper, the most relevant contribution is summarized by the following aspects: 1) tourism markets associated with national park visits are segmented for the first time in Spain; 2), the "duration of the visit" is also established for the first time in the process of using it as a criterion variable in a CHAID analysis. This comparison will allow tourism destination operators and NP management authorities to make more informed decisions regarding the economic and environmental management of tourist destinations; 3) the application, for the first time, of CHAID to measure the environmental impact of different means of transport segmented according to socioeconomic variables. In fact, this study analyzed the socioeconomic characteristics of visitors to a national park and used them as criterion variables to segment the duration of the visit.

3. Objectives and Research Hypotheses

Within the framework of the Autonomous Region of the Canary Islands, the Teide National Park stands out with

a total of 3,289,444 visits, as the most visited NP, not only in the Canary Islands, but also in the entire territory of Spain (INE, 2015). Such an important place, the most popular in terms of visitor numbers, makes us realize how inappropriate the high frequency of visitors is for the conservation of an essential natural resource in the tourist offer of the island where it is based: Tenerife. In fact, the island of Tenerife receives annually five times the number of tourists as its total population.

The data provided by FRONTUR on tourist arrivals on international direct flights for the year 2015 show a total of 5,195,209 tourists arriving in Tenerife, 66.7% more than in 2010 (ISTAC 2002). Visitors to the Teide National Park grew in the same period by 36.6%, giving a 63.3% the ratio of visitors to TNP per number of tourist arrivals. This notable number of tourist visits to the Teide National Park (two out of three tourists visit) coupled with the booming tourism sector in the last two years (2015 and 2016) clearly requires the development of the necessary control measures within the framework of the island's tourism policy and, above all, for the conservation of an essential resource in the island's tourist offer.

3.1 Objectives

The approaches set out above lead to the following objectives for this study:

- 1) To experiment with the technique based on CHAID algorithms (Chi Square Automatic Interaction Detection) in the context of tourist market segmentation for the specific case of the demand associated with NPs, since these are considered important sustainable tourist products in island territories.
- 2) To incorporate, for the first time in the use of this technique by the scientific community, the segmentation based on the "duration of the visit" as the criterion variable.
- 3) Finally, the application of this technique will make it possible to better understand the different segments of the tourism markets currently visiting the Canary Island NPs and, therefore, lead to the development of competitiveness plans aimed at improving both productivity and the conservation of NPs as heritage resources.

3.2 Hypothesis

Based on the above objectives, this research sets out to demonstrate the following hypotheses:

H1. CHAID algorithms allow the construction of a decision tree, relevant for the management of Canary Island NP, using as criterion variable "the duration of the visit".

H2. The decision tree resulting from the application of the CHAID algorithms as a classification tool shows as its first classification variable in the hierarchy, "the country of origin".

Testing the above hypotheses will favour the adoption of precise and better defined tourism policies aimed at improving the competitiveness of TNP tourism and also of tourism environmental policy, which will result in greater and better insular tourist development.

4. Method

4.1 Variables

The variables considered in the segmentation will be of several types: demographic (age, educational level, gender, marital status), economic (household income level, current occupation, type of housing, high/low season and daily expenditure in destination by components) and geographical (country or region of origin of the tourist).

4.2 Characteristics of Information Collection

This study is a quantitative one, and in order to have a precise knowledge of visitors to the TNP, a structured questionnaire divided into two blocks was developed. A first block with questions aimed at obtaining information that determines the most outstanding characteristics of the population under study, namely: nationality or place of habitual residence; frequency of visits to the TNP; with whom the visit was made; sex, age, marital status, current occupation, studies performed, income level of the family unit, type of accommodation and municipality. The second block includes all those questions aimed at collecting information related to the specific objectives of this research.

An ad-hoc survey was developed allowing a much more precise knowledge of the visitors of the park.

4.3 Sampling Technique

Simple random probabilistic sampling was used with 805 effective surveys being carried out and fieldwork organized in two phases. The first phase of interviews was made in the high tourist season on the island including the Easter period. The second phase in low season corresponds to late spring and summer. A reliability of $\pm 2\delta$ was achieved, which is detailed below (Table 1), as a function of the sample size, $n = 805$. The

estimation error was $\pm 3\%$ of the proportions calculated on the total of variables, under the assumption of dichotomy.

Table 1. Maximum estimation error by variable

Variables	Maximum estimation error
Country or region of origin	$\pm 2.88\%$
Gender	$\pm 3.52\%$
Age	$\pm 3.06\%$
Marital status	$\pm 3.48\%$
Current occupation	$\pm 3.54\%$
Educational level	$\pm 3.45\%$
Household income level	$\pm 2.99\%$
Type of accommodation	$\pm 3.25\%$
Season High/Low	$\pm 3.49\%$

Note. Source: data from the questionnaires

4.4 Statistical Analysis Technique

The technique of analysis applied has been the CHAID algorithm, a novel technique from the point of view of its application to tourist market segmentation (Diaz-Perez, & Bethencourt Cejas, 2016).

The Chi-square Automatic Interaction Detection (CHAID) or Genetic Algorithm model generates a decision tree from a significant Chi-square. It is a technique of predictive analysis based on the choice of a criterion variable associated with the rest of the variables that configure the segments through a dependence relationship. Tourism destination operators could use this technique to conduct an analysis depending on previously established objectives, since it allows the *a priori* choice of a criterion variable. Chi-square is the basic statistic collected in CHAID analysis providing a choice between segments in a qualitative and more natural way and with greater explanatory power than other techniques. In addition, Chi-square is also designed for different variables: discrete and freely distributed variables, among others. In reaching such a great amount of advantages, automation has played a fundamental role; in fact, without the use of suitable software it would be difficult to calculate the complex CHAID algorithms.

With respect to the different forms of regression analysis, CHAID analysis is presented as a more rigorous technique insofar as the researcher does not incorporate any value judgment when selecting the independent variables. Such selection is executed by an automated statistical procedure, depending on the classification power of the significant variables. Although the selection of independent variables is also performed *a priori*, the procedure is designed to consider a sufficiently large number of possible exogenous variables, whose association or not with the criterion variable is in the hands of CHAID. It is therefore determined automatically by the software that supports it. To the extent that a non-criteria analysis employs a batch of variables that may not be significant predictors in explaining the configuration of the segments, the results may not achieve optimal homogeneity from the viewpoint of classifiers; which is not the case when CHAID is used as the classification technique. The number of categories of independent variables depends on whether the results of applying the Chi-square test are significant or not. In the resulting decision tree, the most significant variables appear on the first node of the segmentation. The process of node formation and segment configuration ends when the independent variables and the dependent variables no longer have a significant relationship between them.

Finally, the hierarchy obtained for a set of significant variables provides extremely useful information for a destination operator, since this technique identifies which variable segments the most and shows a ranking based on a decreasing segmentation power. Thus, this becomes a very useful tool when the policy maker has in mind the promotion of those forms of tourism with the greatest positive effect on the environment or local economy.

5. Results

5.1 Countries of Origin of Visitors

The information obtained from the questionnaires, with a procedure of purely descriptive statistics, characterizes visitors according to the country of origin in three groups: tourists from abroad (29 countries), tourists from Spain and Balearic Islands and from Tenerife and rest of Canary Islands.

The group of tourist from abroad is classified in the following: 1) visitors from 13 countries which are part of what is traditionally called Western Europe; 2) 10 can be considered as Central and Eastern Europe countries; 3) four American countries; and 5) others, mainly Israel and Australia.

Considering the percentage of visitors by country of origin, we find the largest share (14.5%) are Germans, followed closely by those coming from the United Kingdom (12%) and at a greater distance, France (5.9%),

Italy (5.2%) and Sweden (2.7%). In the bloc of countries, we have called Central and Eastern Europe, Russia stands out with 4.1%, followed by Poland with 1.8%, Czech Republic, 1.8%, Latvia, 0.9%, Estonia and Romania with 0.7% each and a number of countries with 0.5% (Lithuania, Croatia, Ukraine). Out of the four American countries the highest percentage corresponds to USA with 0.7%, Argentina and Brazil with 0.5. Finally, the percentage of visitors from Israel stands at 1.4%. The percentage of visitors from mainland Spain and Balearic Islands reaches the important amount of 21%. And finally, there is 16% of the visitors to the TNP that are resident on the island of Tenerife, to which must be added 4.6% of visitors from other islands of the archipelago.

The above data show the main countries of origin of visitors to the TNP, which we can complete by establishing a more precise profile of these visitors using CHAID algorithm.

5.2 Results of Applying the CHAID Algorithm

The results obtained from applying the CHAID algorithm segmentation technique to the total of the visitors to the TNP, considering “the duration of the visit” as a criterion variable and the rest of variables above as predictors, are those that appear in the following figure. All the values included in the decision tree are significant (Chi-square 141.662), and well classified in a nearly 90% (upper to 11.0705) of the total visitors for five degrees of freedom. As we can see, the first node of the classification is not the country of origin, but the mean of transport used.

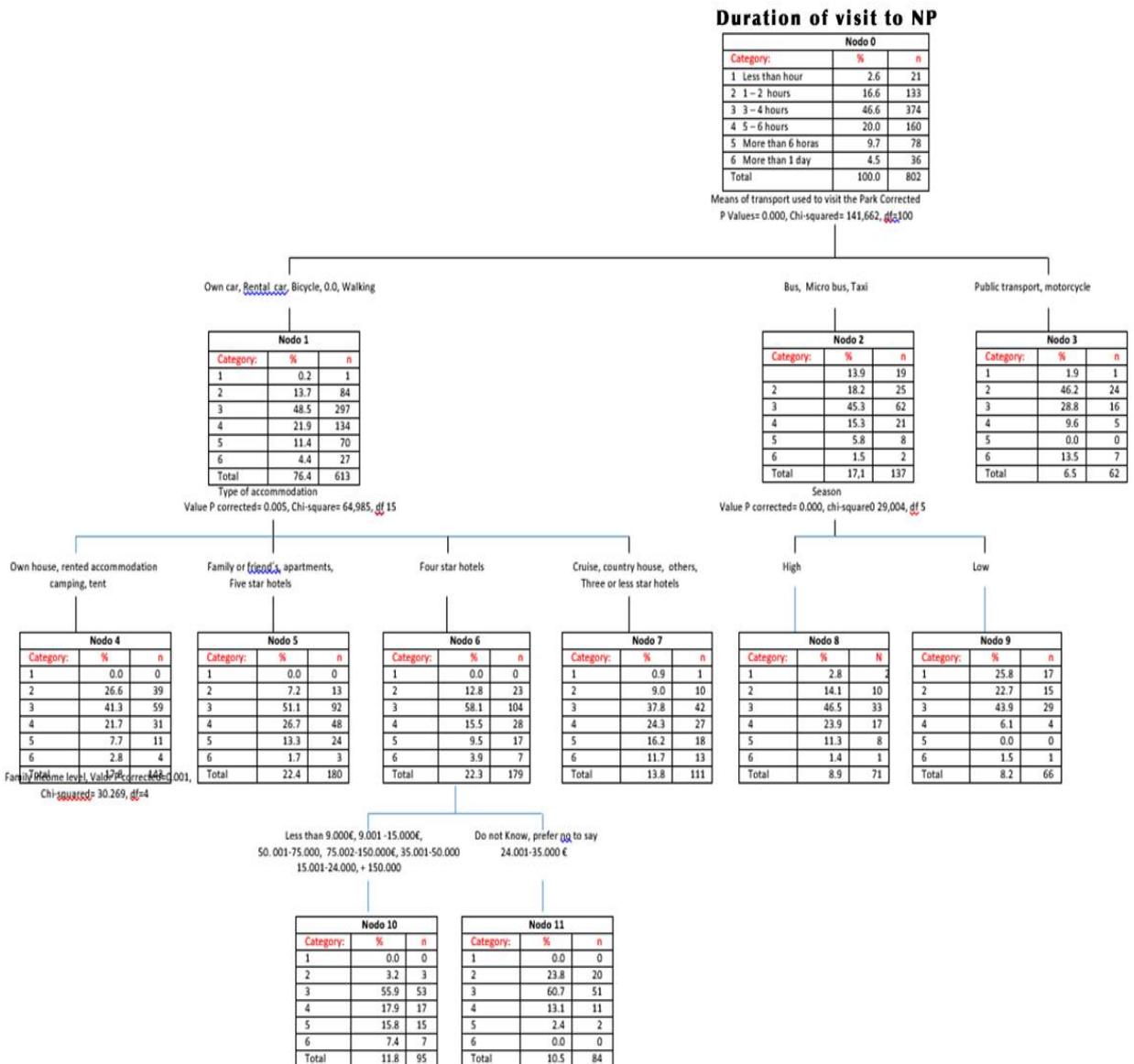


Figure 1. Results of applying CHAID algorithm

Chi-squares and P values relative to the second and third node can be checked in the figure and are all significant, as well. The tree represents segmentation levels and, most importantly shows that the most relevant variable when segmenting is the means of transport used to visit TNP. The second level is determined by those who use an individual means of transport, and within this, the third level is given by the type of accommodation, which includes those who stay in 4-star hotels depending on their income level.

Returning to the second level, some relevant results were found:

First, for those who visit using collective transport, the segmentation variable that discriminates best is the season (high / low).

Second, visitors who pollute the most, those who use cars to visit, are the ones that spend the most time in the park and, therefore, those that most deteriorate it. It is worth remembering that over 3 million visitors a year enter the TNP. Everything seems to indicate that this segment is configured by local people and tourists who rent cars.

Third, the bus, minibus and taxi segment is possibly segmented in high and low season because it corresponds, in its entirety, to visitors to the TNP who are non-local tourists. Interestingly, those who come in high season spend more time inside the park.

Finally, it appears that those who use public transport correspond to those who spend more than one day in the National Park and reach a percentage of 13.5% of the total frequency of visitors of this segment. In the other hand, those who visit the TNP by motorcycle are the ones that spend the least time in the park, usually between one and two hours.

6. Conclusions & Discussion

As can be seen from the results of the analysis, the application of the CHAID algorithm segmentation technique makes it possible to conclude the following:

- 1) The first of the hypotheses is fulfilled, that is to say, CHAID algorithms allow the construction of a decision tree relevant to the needs of the management authorities of the Canary Island NPs, using as a criterion variable "the duration of the visit".
- 2) The decision tree resulting from the application of the CHAID algorithms as a classification tool reveals as the first classification variable in the hierarchy "the means of transport used" and not "the country of origin of the tourist". Thus, the second hypothesis is rejected. In fact, the geographic origin of the tourist does not appear clearly as a predictive variable in the model, although as we said in the results, in the background of some segments we can sense a relevant role for visitors' country of origin, with a bivalent character: local / not local.

All in all, the findings show that the first segmenting variable is transport type and, that the means most frequently used by visitors is the car. Specifically, the segment of visitors coming by car is associated with the longest stay in the TNP.

Regarding the practical and social implications, it is assumed the longer the stay, the greater the environmental impact. These results, therefore, show the need for new transport strategies with less polluting vehicles. In this sense there should be strategies such as: 1) the greater use of public transport or, 2) the use of buses for trips within the TNP and on the way to the park or, perhaps also 3) the use of electric vehicles to move around within the park.

In short, the results testing the previous hypotheses could help in the adoption of precise and defined tourism policies aimed at improving the competitiveness of TNP tourism production as well as a tourism environmental policy. This would also result in enhanced and more sustainable island tourist development.

6.1 Discussion

The basic premises of this paper are as follows. First, the idea that the longer the visit, the greater the environmental deterioration. Second, it is assumed that the use of the car as a means of transport, compared to other transport, such as bus, minibus ... causes greater pollution and environmental deterioration, since the car is highly polluting due to its high fossil fuel consumption per traveller.

At this point, and in view of the statistical data collected above, it does not seem an easy task to resolve the contradiction between objectives to improve the competitiveness and sustainability of the NP resource. In fact, the competent local authorities will have to deal with the dialectical relationship between competitiveness (measured by the number of visits per year) and sustainability (measured by the state of conservation of the natural heritage). In addition, all this is in the context of an island territory with a high population density and a

number of tourists a year that quintuples the total population of the island.

Bearing in mind the above premises, and considering that the total number of visitors is not a homogeneous group, the first step in this research has been to segment the whole group of visitors with the aim of determining which segments generate the greatest environmental impact.

Nevertheless, the starting point for future research should be segmenting the tourism market of the Teide NP based on different perspectives of how nature should be used as a tourism resource. Additionally, as it is possible to find many different definitions of carrying capacity, future studies should be focused on what is considered to be the threshold limit in carrying capacity of tourists on the one hand and, at the same time, by local people, on the other.

Acknowledgment

This work has been supported by funding from the National Park of El Teide, the Canarian Department of Education, Culture and Sport, Government of the Canary Islands, Spain.

References

- Alegre, J. Mateo, S., & Pou, L. (2011). A latent class approach to tourists' length of stay. *Tourism Management*, 32, 555-563. <https://doi.org/10.1016/j.tourman.2010.05.003>
- Assaker, G., & Hallak, R. (2012). European travelers' return likelihood and satisfaction with mediterranean sun-and-sand destinations: a chi-square automatic identification detector e based segmentation approach. *Journal of Vacation Marketing*, 18(2), 105-120. <https://doi.org/10.1177/1356766711435977>
- Bell, F. W., & Leeworthy, V. R. (1990). Recreational demand by tourists for saltwater beach days. *Journal of Environmental and Management*, 18(3), 189-205. [https://doi.org/10.1016/0095-0696\(90\)90001-f](https://doi.org/10.1016/0095-0696(90)90001-f)
- Chen, J. S. (2003). Market segmentation by tourists' sentiments. *Annals of Tourism Research*, 30(1), 178-193. [https://doi.org/10.1016/S0160-7383\(02\)00046-4](https://doi.org/10.1016/S0160-7383(02)00046-4)
- Chung, K. Y., Oh, S. Y., Kim, S. S., & Han, S. Y. (2004). Three representative market segmentation methodologies for hotel guest room customers. *Tourism Management*, 25(4), 429-444. [https://doi.org/10.1016/S0261-5177\(03\)00115-8](https://doi.org/10.1016/S0261-5177(03)00115-8)
- Cook, V. J., & Mindak, W. A. (1984). A search for constants: The heavy user revisited. *Journal of Consumer Marketing*, Spring, 79-81. <https://doi.org/10.1108/eb008118>
- D íz-P érez, F. M., & Bethencourt-Cejas, M. (2016). CHAID algorithm as an appropriate analytical method for tourism market segmentation. *Journal of Destination Marketing & Management*, 5(3), 275-282. <https://doi.org/10.1016/j.jdmm.2016.01.006>
- D íz-P érez, F. M., Bethencourt-Cejas, M., & Álvarez-Gonz ález, J. A. (2005). The segmentation of Canary island tourism markets by expenditure: implication for tourism policy. *Tourism Management*, 26(6), 961-964. <https://doi.org/10.1016/j.tourman.2004.06.009>
- Fleischer, A., & Pizam, A. (2002). Tourism constraints among Israeli seniors. *Annals of Tourism Research*, 29(1), 106-123. [https://doi.org/10.1016/S0160-7383\(01\)00026-3](https://doi.org/10.1016/S0160-7383(01)00026-3)
- Gokovali, U., Bahar, O., & Kozak, M. (2007). Determinant of length of stay: A practical use of survival analysis. *Tourism Management*, 28(3), 736-746. <https://doi.org/10.1016/j.tourman.2006.05.004>
- Gomes de Menezes, M., & Cabral, V. (2008). The determinants of length of stay of tourists in the Azores. *Tourism Economics*, 14(1), 205-222. <https://doi.org/10.5367/000000008783554866>
- Holden, A. (2000). *Environment and tourism*. London: Routledge.
- Hsu, H. C., & Kang, S. K. (2007). CHAID- based Segmentation: International Visitors' Trip Characteristics and Perceptions. *Journal of Travel Research*, 46, 207-216. <https://doi.org/10.1177/0047287507299571>
- Hughes, G., & Furley, P. (1996). Threshold, Carrying Capacity and the Sustainability of Tourism: A Case Study of Belize. *Caribbean Geography*, 7(1), 36-51.
- INE. (2015). Instituto Nacional de Estad ítica. <http://www.ine.es/>
- ISTAC (Instituto Canario de Estad ítica). (2002). Demanda turística: turistas y pasajeros. Gobierno de Canarias. <http://www.gobiernodecanarias.org/8/03/2015>
- Jim, C. Y. (2000). Environmental changes associated with mass urban tourism and nature tourism development in Hong Kong. *The Environmentalist*, 20(3), 233-247. <https://doi.org/10.1023/A:1006708031927>

- Kardes, F. R. (2002). *Consumer behavior and managerial decision making*. Upper Saddle River, NJ: Prentice-Hall.
- Kass, G. V. (1975). Significance testing in automatic interaction detection. *Journal of the Royal Statistical Society. Series C (Applied Statistics)*, 1 January, 24(2), 178-189. <https://doi.org/10.2307/2346565>
- Kass, G. V. (1980). An Exploratory Technique for Investigating Large Quantities of Categorical Data. *Applied Statistics*, 29(2), 119-127. <https://doi.org/10.2307/2986296>
- LaPage, W. F. (1969). *Campground marketing: The heavy-half strategy*. USDA forest service research note NE_93. Upper Darby, PA: Northeastern Forest Experiment Station.
- Legohérel, P. (1998). Toward a market segmentation of the tourism trade: Expenditure levels and consumer behavior instability. *Journal of Travel Research*, 29(4), 12-18. https://doi.org/10.1300/j073v07n03_02
- Legohérel, P., & Wong, K. F. (2006). Market segmentation in the tourism industry and consumers` spending: what about direct expenditure? *Journal of Travel and Tourism Marketing*, 20(2), 15-30. https://doi.org/10.1300/J073v20n02_02
- Legohérel, P., Hsu, C. H. C., & Daucé B. (2015). Variety-seeking: Using the CHAID segmentation approach in analyzing the international traveler market. *Tourism Management*, 46, 359-366. <https://doi.org/10.1016/j.tourman.2014.07.011>
- Mak, J., & Nishimura, E. (1979). The economics of a hotel room tax. *Journal of Travel Research*, 17, 2-6. <https://doi.org/10.1177/004728757901700401>
- Mak, J., Moncur, J., & Yonamine, D. (1977). Determinants of visitor expenditure and visitor length of stay: a cross-section analysis of U.S. visitors to Hawaii. *Journal of Travel Research*, 15, 5-8. <https://doi.org/10.1177/004728757701500302>
- Mathieson, A., & Wall, G. (1982). *Tourism: Economic, Physical and Social Impacts*. Longman: New York.
- McCarty, J. A., & Hastak, M. (2007). Segmentation approaches in data-mining: A Comparison of RFM, CHAID, and logistic regression. *Journal of Business Research*, 60, 656-662. <https://doi.org/10.1016/j.jbusres.2006.06.015>
- Mok, C., & Iverson, L. G. (2000). Expenditure-based segmentation: Taiwanese tourists to Guam. *Tourism Management*, 21(3), 299-305. [https://doi.org/10.1016/S0261-5177\(99\)00060-6](https://doi.org/10.1016/S0261-5177(99)00060-6)
- Nicholson, R., & Pearce, D. (2000). Who goes to events: A comparative analysis of the profile characteristics of visitors to four south island events in New Zealand. *Journal of Vacation Marketing*, 6, 236-253. <https://doi.org/10.1177/135676670000600304>
- Odell, R. (1975). Carrying Capacity Analysis: Useful but Limited. R. Scott, D.J. Brower, D. Miner (Eds.), *Management and Control of Growth*, vol. III, The Urban Land Institute: Washington DC, 22-28.
- Pestana, B. C., & Correia, A. (2007). *Survival analysis in tourism demand: The length of stay in Latin American destinations*. Working paper 03/2007/DE. Departamento de Economia, Instituto Superior de Economia e Gesta, Universidade Técnica de Lisboa.
- Pestana, B. C., & Pinto, M. L. (2010). The length of stay in tourism. *Annals of Tourism Research*, 7(3), 692-706. <https://doi.org/10.1016/j.annals.2009.12.005>
- Ramdas, M., & Mohamed, B. (2014). Visitor Perceptions on the Impacts of Tourism Activities, Development and Infrastructure on the Environment of Perhentian Islands. *SHS Web of Conferences*, 12 (0108), 4th International Conference on Tourism Research (4ICTR). <https://doi.org/10.1051/shsconf/20141201081>
- Rhim, H., & Cooper, L. G. (2005). Assessing potential threats to incumbent brands: New product positioning under price competition in a multisegmented market. *International Journal of Research in Marketing*, 22(2), 159-182. <https://doi.org/10.1016/j.ijresmar.2004.06.005>
- Saarinen, J. (2006). Traditions of sustainability in tourism studies. *Annals of Tourism Research*, 33(4), 1121-1140. <https://doi.org/10.1016/j.annals.2006.06.007>
- Solomon, M., Bamossy, G., & Askegaard, S. (2002). *Consumer behaviour: A European perspective*. Essex: Pears.
- Spotts, D. M., & Mahoney, E. M. (1991). Segmenting visitors to a destination region based on the volume of their expenditure. *Journal of Travel Research*, 29(4), 24-31. <https://doi.org/10.1177/004728759102900405>

- Stynes, D. M., & Mahoney, E. M. (1980). *Michigan downhill ski market study: Segmenting active skiers*. Research report 391. East Lansing, MI: Michigan University Agricultural Experiment Station.
- Thrane, C. (2012). Analyzing tourists length of stay at destinations with survival models: A constructive critique base don a case study. *Tourism Management*, 33, 126-132. <https://doi.org/10.1016/j.tourman.2011.02.011>
- UNEP United Nations Environment Programme. (1999). Small islands developing states. Our planet. (Accessed 10.01.2017)
- Vassiliadis, C. A. (2008). Destination product characteristics as useful predictors for repeat visiting and recommendation segmentation variables in tourism: a CHAID exhaustive analysis. *International Journal of Tourism Research*, 10, 439-452. <https://doi.org/10.1002/jtr.678>
- Zubair, S., Bowen, D., & Elwin, J. (2011). Not quite paradise: Inadequacies of environmental impact assessment in the Maldives. *Tourism Management*, 32, 225-234. <https://doi.org/10.1016/j.tourman.2009.12.007>

Copyrights

Copyright for this article is retained by the author(s), with first publication rights granted to the journal.

This is an open-access article distributed under the terms and conditions of the Creative Commons Attribution license (<http://creativecommons.org/licenses/by/4.0/>).