**ISCTE IUL**

**Instituto Universitário de Lisboa**

Science and Information Technology Department

# The loyalty of European tourists from a destination country perspective

Ricardo Jorge Maria Correia da Silva

Dissertation submitted as partial fulfilment of the requirement for the degree of Master in Information Systems Management.

Orientador
Doutor Raul Manuel Silva Laureano, Professor Auxiliar,
ISCTE-IUL

October, 2018

## Acknowledgments

First of all, I would like to thank Professor Raul Laureano for his attention, advice, motivation and persistence. Without your support it would not have been possible to submit this project.

To the family, for the motivation and support given not only during the realization of the dissertation but throughout the entire academic course.

To friends, for the friendship and support given, making it possible to overcome all difficulties during the entire masters.

At last, I would like to thank everyone who, in one way or another, have contributed to this thesis. My sincere thanks!

## Resumo

O Mercado do turismo representa em muitos paises um peso significativo na sua economia. Contudo, este mercado tende a ser volátil e, deste modo, de forma a garantir uma sustentabilidade económica, é necessário garantir a lealdade dos turistas. A lealdade é influenciada pela satisfação sendo que esta é resultado de diferentes determinantes ao longo da viagem.

A possibilidade de verificar quais turistas mais leais permite que as agências de turismo, assim como as organizações governamentais posicionarem e alterarem a sua estratégia de forma a potencializar e aumentar a lealdade dos turistas.

De forma a avaliar a lealdade dos turistas, foram realizadas diferentes análises e tecnicas de data mining de forma a verificar quais os turistas mais leais, e quais os tipos de lealdade existentes. Foi ainda criado um modelo preditivo de forma a avaliar a influência das variaveis sócio-demográficas na lealdade dos turistas.

Para tal foram utilizados dados de três anos concecutivos (2014, 2015 e 2016), referentes aos dados do Flash Eurobarometer, de um modo específico, utilizando o questionário "Preferences of European toward tourists", onde foram estudados dados referentes aos turistas da Europa 28.

Através das tecnicas mencionadas, foi possível verificar quais os turistas mais ou menos lais assim como prever o destino de férias dos turistas. Por fim, foi ainda possível ver o impacto das caracteristas socio-demográficas na lealdade dos turistas e qual o seu peso.

**Palavras-Chave:** Turismo; Lealdade; Data Mining; Análise de Dados; Europa.

## Abstract

Tourism market represents in many countries a significant weight in its economy. However, this market tends to be volatile and, in order to ensure economic sustainability, it is necessary to ensure the loyalty of tourists. Loyalty is influenced by satisfaction and it's a result of different determinants along the trip.

The ability to analise which tourists are more loyal allows tourism agencies as well as government organizations to position and change their strategy in order to boost and increase the loyalty of tourists.

To evaluate the loyalty of the tourists, different analyzes of data mining techniques were carried out to verify the most loyal tourists, and the type of loyalty. A predictive model was also created in order to evaluate the influence of socio-demographic variables on the loyalty of tourists.

In order to do so, it was used three-year data (2014, 2015 and 2016) for the Flash Eurobarometer, in a specific way, using the "Preferences of European toward tourists" questionnaire, where were studied data from Europe 28.

Through the techniques mentioned, it was possible to verify, which tourists are the most loyal and which are the least, as well predict the holiday destination of the tourists. Finally, it was still possible to see the impact of socio-demographic characteristics on the loyalty of tourists and their weight.

**Keywords:** Tourism; Loyalty; Data Mining; Data analytics; Europe.

# Index

## List of Tables

## List of Figures

## List of Abbreviations

AUC – Area Under The Curve

CART - Classification and Regression Trees

CHAID - Chi-square Automatic Interaction Detector

CRISP-DM - Cross Industry Standard Process for Data Mining

EU – European Union

GDP - Gross domestic product

LOS - Length of stay

ROC - Receiver Operating Characteristics

SPSS - Statistical Package for the Social Sciences

UNWTO - United Nation World Tourism Organization

# Chapter 1 - Introduction

## 1.1. Theoretical background and motivation

Many European countries are nowadays dependent on tourism, such as Portugal, Malta and Greece. For instance, in Portugal, tourism has been increasing sharply and in the first half of 2017 it was registered a growth of 20% over the previous year (Turismo de Portugal, 2017), with tourism weighing around 5.8% of GDP (World Travel & Tourism Council, 2017), numbers that show its importance.

The concept of tourism has evolved over the years. The first definition of tourism was made by Guye and Feuler in 1905. The authors defined Tourism as "*a phenomenon which is dependent on the people's increasing need for a change and relaxing, the wish of recognizing the beauties of nature and art and the belief that nature gives happiness to human beings and which helps nations and communities*" (Esen & Uyar, 2010). However, the definition has undergone significant transformations and the reason, the length and the satisfaction of the trip have become important factors to be considered as well. Nowadays the definition has been unified by the World Tourism Organization as *"a social, cultural and economic phenomenon which entails the movement of people to countries or places outside their usual environment for personal or business purposes"* (UNWTO, 1994:4-5).

Turism develops in three stages: i) discovery, ii) local response and initiative, and iii) institutionalized 'institutionalization. It is also explicit in Christaller's concept that types of tourists change with the tourist areas (Butler, 1980).

To explore the evolution of tourist areas, a recognizable cycle using a basic S curve is presented to illustrate their waving and waning popularity. There can be no doubt that tourist areas are dynamic, that they evolve and change over time. There are a variety of factors which are responsible for the initial popularity of the area. including changes in the preferences and needs of visitors, the gradual deterioration and possible replacement of physical plant and facilities, and the change (or even disappearance) of the original natural and cultural attractions. In some cases, whilst these attractions remain, they may be utilized for different purposes or come to be regarded as less significant in comparison with imported attractions (Butler, 1980).

According to Christaller (1963), the first people to discover a place are usually the painters. In search of art and inspiration, painters are the pioneers and, gradually, the place

develops as a so-called artist colony. Soon, a cluster of poets follows, kindred to the painters, then cinema people, gourmets, and the jeunesse dorde.

The place becomes fashionable and the entrepreneur takes notice. They capitalize on the good name of this former painter's corner and on the gullibility of tourists. Finally, tourist agencies come with their package rate travelling parties; Now, the indulged public avoids such places. At the same time, in other places the same cycle occurs again as more and more places become fashionable (Christaller, 1963).

Figure 1 represents an evolution based on the number of tourists over time. When some areas reach the Stagnation stage it can be ascertained that the peak numbers of visitors will have been reached. Capacity levels for many variables will have been reached or exceeded, with attendant environmental, social, and economic problems (Butler, 1980).

As we can see in Figure 1 there are different ways that the curve can follow up. To explain the different hypotheses, a phase of decline and one of rejuvenation will be exposed.
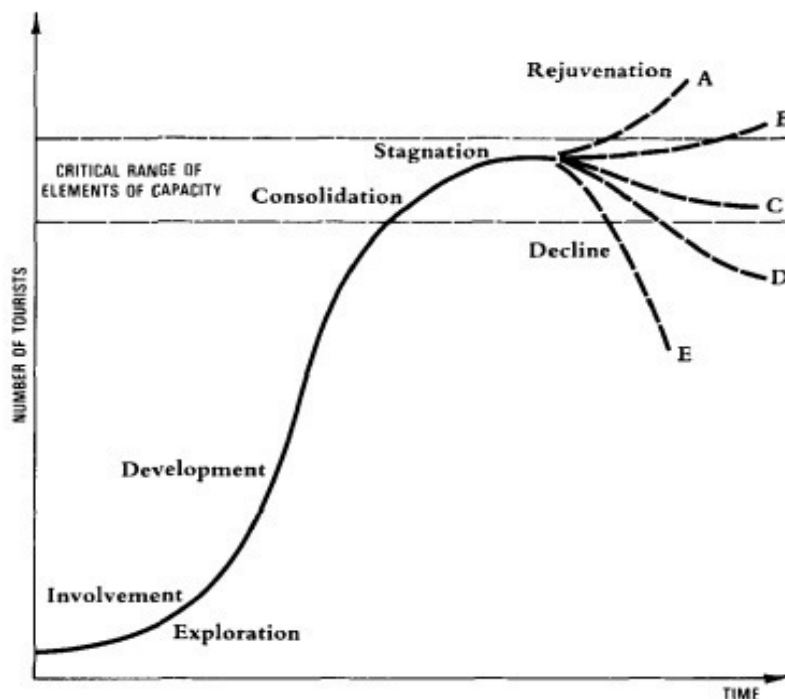


*Figure 1 - Hypotethital evolution of a tourist area. (Butler, 1980)*

In the rejuvenation phase the existence of a change in the tourism attractions that the place is based is implicit. In this case it is important to explore different areas that nowadays are not explored either by creating a new attraction or by exploiting natural resources that were not yet explored. Nonetheless, a new way of presenting to tourists is

necessary. If the country does not have the capacity for this transformation it will result in a phase of decline.

As mentioned, the alternative is the decline stage and that means that the country is not able to compete with the new attractions provided by other countries. That will result in a declining market, both spatially and numerically. As countries change many factors could affect the cycle of tourism area. So, either because there was a terrorist's attack that leads to a lack of tourists, a major event such as web summit that brings people all around the world to visit and explore the country or the creation of a resort that creates the facilities needed for the country's next phase, countries must have the ability to stand out from the others to ensure the loyalty of tourists.

Tourism loyalty is the key to understand what makes a tourist return to a destination that he likes and this concept is especially important to the companies that are responsible for tourism marketing and tourism management, so they can understand and go after the tourist's needs. To avoid entering a phase of decline, there are different determinants that can influence a person to return to their destination country.

The literature about loyalty and its determinants is focused on a micro level, such as loyalty to a hotel (Kandampully, 2000) or a cultural (Kim, Suh & Eves, 2010) or sporting event (Gedenk Neslin, 1999). Thus, in the current competitive context between countries, loyalty analysis should also focus on a macro level, that is, at the countries level or even at the continents level where studies are scarce (Talib, et al. 2015).

Morevoer, the existing studies tend to be based on small samples with data collected through a questionnaire (Kozak & Rimmington, 2014) and are often limited to verifying hypotheses stated by theories through statistical techniques in a new context (Kozak & Rimmington, 2014). By expanding the data universe to the European tourists, it becomes possible to recover different techniques of data mining and to extract useful and new knowledge for the different tourism stakeholders.

## 1.2. Problem and research objectives

Thus, it is essential for countries not only to gain new markets (tourists), but also to ensure the loyalty of their tourists to the country. Having established this, a new problem arises of how to increase the level of loyalty to the country in an extremely competitive European context.

Faced with this problem, the following objectives are defined,

O1. Identify pairs of countries visited in consecutive years.

O2. Evaluate loyalty to one's country, foreign country, or both.

O3. Identify socio-demographic characteristics that may explain loyalty, i.e., identify profiles of loyal tourists.

In order to achieve these objectives, this research aims at European tourist's behaviour (EU 28) in light of a certain experience in each country to ascertain whether they would return to the same country the following year. This is a theme to be associated with different socio-demographic characteristics to evaluate if tourists coming from a certain country have more tendency to return to the same destination if they have an enjoyable experience or not.

This study allows tourism professionals access to new knowledge to collect information in order to optimize their own advertising to reach the desired people. Additionally, in a scientific approach, the study, based on a sample of thousands of European tourists, will allow us to evaluate the data mining techniques that produce the best results at the level of a new context (loyalty to the destination country) and, in parallel, it allows analysis of loyalty in a broader European context which may or may not confirm some predictive models of loyalty (obtained at micro level).

Lastly, the study will identify differences between countries, both origin and destination, which will allow each to know the reality of the others and position itself in the best way to attract more people to visit the country.

As mentioned, the alternative is the decline stage, and which would mean that the country is not able to compete with the new attractions provided by other countries. That will result in a declining market, both spatially and numerically.

## 1.3. Methodology

This research used the CRISP-DM (CRoss-Industry Standard Process for Data Mining) methodology, which is composed of six phases: business understanding, data understanding, data preparation, modelling, evaluation, and deployment (Chapman et al., 2000).

In the business understanding phase, the goal is to evaluate what leads tourists to be loyal to the country of destination. This analysis has the aim of understanding what leads tourists to become more loyal. In particular, widentify profiles of loyal and non-loyal tourists

Given that the data from Eurobarometer (European Comission, 2015) is available and not appropriately studied and the expensive and time-consuming nature of the activity of data collection, this study uses the data collected by the Eurobarometer from tourists of 2014 to 2016 in the 28 countries in the European Union.

In the second phase of CRISP-DM – data comprehension – 630 attributes collected from 90.101 European tourists were analysed and included in the Excel database. Given the size of the data, it is also necessary to prepare it and transforme it to reach the outcome.

In the modeling phase different analyses were carried out using contingency tables, the apriori model and classification trees. During this phase, all the variables considered to be eligible were tested to assess the loyalty of European tourists. The models were evaluated in the fith phase of the methodology.

In the last phase, deployment, the intention is to present this dissertation and its results to the tourism loyalty by presenting it in conferences and also publishing it in scientific papers.

## 1.4. Structure

This dissertation follows a traditional structure. In addition to this introduction, it contains four other chapters. Chapter 2 – Literature Review – includes an overview of tourism, together with pre, during and post-travel tourist behaviour, loyalty with destination and with accommodation and a summary of the studies found. In chapter 3 – Methodology – the CRISP-DM methodology is described, with emphasis on the comprehension, preparation, modelling and evaluation phases. In chapter 4 – Results – the findings are presented and discussed. In chaptper 5 – Conclusions – the conclusions are described focusing on contributions and study limitations.

## Chapter 2 – Literature review

### 2.1. Tourism - Concept and relevance

The first step is to analyse the meaning of tourism and tourist to understand what leads a person to be loyal to a holiday destination. When reviewing the literature, it was verified the existence of a resemblance between trust, satisfaction and loyalty (Vieira, 2016), so it is necessary understand their differences and issues before the data analysis.

Tourism, according to the dictionary, is *"the commercial organization and operation of holidays and visits to places of interest"*. The word emerged late in the era of the grand tour, English wealthy classes whereby young men were dispatched on extensive circuits of continental Europe to finish their education. Its roots include the Greek term for a tool used to describe a circle, reflected in the essential feature of tourism, returning to the point of departure (Leiper, 1979).

Travelling has always been an intrinsic feature of the human being. The first civilizations would travel to seek food and shelter and, with the evolution of society, travelling started to have as objective prominent but became limited to trade, pilgrimage, studies, migration, royal affairs and exploration.

The first definition of tourism emerges in 1905 and referes to tourism as *"A phenomenon unique to modern time which is dependent on the people's increasing need for a change and relaxing, the wish of recognizing the beauties of nature and art and the belief that nature gives happiness to human beings and which helps nations and communities' approaching to each other thanks to the developments in commerce and industry and the communication and transportation tools' becoming excellent"* (Esen & Uyar, 2010).

The term tourism suffers a lot of changes with the passing of the years and in 2000, the World Tourism Organization, The Commission of the European Communities and the United Nations in a first attempt to unify the term tourism, defined this as "*an activity that comprises other activities of persons traveling to and staying in places outside their usual environment for not more than one consecutive year for leisure, business and other purposes"* (World Tourism Organization, 2001).

In 2005, the World Tourism Organization succeded in unifying the definition in order to prevent missunderstandings, so, nowadays the tourism is defined as "*a social, cultural*

*and economic phenomenom which entails the movement of people to countries or places outside their usual enviorment for personal or business purposes. These people are called tourists and tourism has to do with their activities and experiences*" (UNWTO, 2018).

**2.2. Loyalty**

2.2.1.     Concept and Evolution

The concept of loyalty has been evolving over the years. In general, loyalty is defined by the repetition of purchase, either at the frequency level or by the quantity of products of the same brand acquired. Loyalty is defined by Oliver in 1999 as a commitment to repurchase a product or service consistently in the future. This causes a repetitive set of purchases of the same brand and set of brands, regardless of all influencers (positives or negatives) and all marketing companies that should influence the consumer's buying decision. That definition leads to an existence of loyalty phases and different ways that loyalty could affect the behaviour of the people (tourists).

2.2.2.     Loyalty Phases

Oliver's framework (Oliver, 1999) follows a model divided in a cognition, affect, conation, nonetheless, it is different in measure, according to the point of view, however, in a different approach, according to their point of view, each tourist becomes "loyal" in a phased way and is influenced by different factors that directly change their behavior. In this study, tourists from different countries may be affected in different ways, based on their cultures and quality of life. According to Oliver, consumers are influenced to become loyal in a cognitive sense first, consecutive in an affective sense, afterward in a conative manner, and eventually in a behavioural manner, which is described as "action inertia."

The first phase, denominated as cognitive loyalty, is based on brand advertising (or referring to the case study, a country). This stage is called cognitive loyalty or loyalty based only on brand belief. Cognition can be based on information acquired either directly or indirectly including information based on recent experiences (Oliver, 1999). In this case loyalty is associated with a feeling of satisfaction and a good performance that leads to a habit.

Affective loyalty is the second phase of loyalty profess, in this case a person has been evolving a satisfaction sensation towards and a brand or a country due to a series of good

experiences. Loyalty at this stage is still very much conditioned by the number and type of alternatives available (Sambandam & Kenneth, 1995). Commitment at this stage is referred to as affective loyalty and this is seen in the minds of consumer consumers with a feeling of cognition and affection. Considereing that cognition is directly subject to counter argumentation, affect is not as easy to change the opinion.

Loyalty in this case is directly related to an existing appreciation of the brand or country. However, like cognitive loyalty, it is still very volatile and susceptible to competition and different alternatives. In fact, many of the people who end up leaving the mark at this stage, define their experience as positive and were even satisfied with the product. In this way a level of commitment would be preferable, thus guaranteeing a greater loyalty (Oliver, 1999).

Cognitive loyalty is related to desire and is the phase after affective loyalty. At this stage loyalty is influenced by episodes of satisfaction and affection regarding the brand. The concept of cognitive is related to a desire, in this case the purchase of the brand or return to the country that one liked. Cognitive loyalty is associated with a desire to purchase a product and this wish is stronger than in affective loyalty. However, despite its intent to purchase this may be unfulfilled, with room for a deeper level of commitment (Oliver, 1999).

Action loyalty is the study of the mechanism by which intentions are converted to actions and its is referred to as "action control" (Kuhl and Beckmann 1985). All previous loyalty phases are reflected in the fact that the consumer is ready to go back and bought the product (in this case study, that the tourist is willing to return to the country), however, if alternatives appear, they can cause the consumer loses brand loyalty. It is necessary that loyalty reaches a level where, regardless of the obstacle, the consumer re-acquires the product or the service. Action loyalty is perceived as a necessary result of the engagement of these three states. If this commitment is achieved, an inertia of action develops, which leads the consumer to repurchase.

### 2.2.3. Micro vs Macro Loyalty Levels

Loyalty research into brand loyalty and/or consumer loyalty dates back well more than 40 years. According to Jacoby and Chestnut (1978), its history extends back to Copeland's (1923) study on a phenomenon that he labelled as "brand insistence."

Generally, there are two different approaches to loyalty: A micro level one where the focus is on a brand, a space and/or an influence maker; A macro or aggregate level one where the focus refers-to a big company (Jacoby & Chestnut, 1978)

There are different analyses at a micro level, either in a hotel, in a event or even in a unique country. However, in a macro concept where the concept of loyalty is related in a country of destination perspective, there are a lack of studies.

### 2.2.4. Destination Loyalty

As mentioned, repeat purchases or recommendations to other people are most usually referred to as consumer loyalty in the marketing literature (micro level). The success of a marketing campaign is directly related to the level of loyalty that consumers reach. (Flavian, Martinez, & Polo, 2001).

Similarly, travel destinations can be considered as products, and tourists may revisit or recommend travel destinations to other potential tourists such as friends or relatives.

However, the study of the usefulness of the concept of loyalty and its applications to tourism products or services has been limited, even though loyalty has been thought of as one of the major driving forces in the competitive market (Dimanche & Havitz, 1994).

In the last decade, tourism and leisure researchers have incorporated the concept of consumer loyalty into tourism products, destinations, or leisure/recreation activities (Backman & Crompton, 1991; Baloglu & Mangaloglu, 2001; Mazanec, 2000; Pritchard & Howard, 1997; Selin, et al, 1988).

Usually, loyalty is categorized in the following ways: (1) the behavioural approach, (2) the attitudinal approach, and (3) the composite approach (Jacoby & Chestnut, 1978).

The behavioural approach is related to consumers' brand loyalty and has been characterized as frequent purchase or probability of purchase. The measurement of this approach lacks a conceptual standpoint and does not attempt to explain the factors that affect customer loyalty. Namely, the desire to revisit or recommend a country to other potential tourists, may not be fully explained in full by loyalty. (Dick & Basu, 1994).

The attitudinal approach is based on consumer brand preferences or intention to buy and consumer loyalty and it is related to the loyalty phases described by Oliver (1999). This approach relates loyalty in terms of psychological commitment.

In the field of tourism, this refers to cases in which the image that the tourist has on a country is not the best, but the tourist continues to have a desire to visit it (Backman & Crompton, 1991).

Lastly, the combination or composite approach is a confluence between the attitudinal approaches and the behaviour approach (Backman & Crompton, 1991). Usually, it is recognized that customers who purchase and have loyalty to particular brands must have a positive attitude toward those brands. This approach has limitations and even some researchers have discounted only the behavioural or attitudinal approach and have suggested integrating the two (Backman & Crompton, 1991). Thus, the reviewed literature suggests that it is necessary to build a relationship of satisfaction and motivation around the client so that there is a bond of loyalty (Isoraité, 2016).

**2.3. Loyalty determinant's – Travel Phases**

2.3.1.      Pre-Travel

      a.   Tourist

A classic example deals with the terms used to denote persons who travel. While words like ", traveller", "visitor", "passenger", and "guest" may connote indifferent, acceptable or positive meanings, the term "tourist" often connotes a rude, humorously dressed, misbehaving, unsophisticated oaf tangled in the straps of a camera. While most people who are traveling to places outside of their own home environments willingly label themselves as travellers, guests, or visitors, they seldom call themselves "tourists" reserving the term for other people (D. Hunt & Layne, 1991).

A tourist is an individual *"who travels for a period of 24 hours or more in a country other than that in which he usually resides"* (Shaw & Williams, 2009).

      b.   Personal Characteristics

The personal characteristics of the tourist, such as place of residence, age, income, level of education, place of residence, personal motivation to travel, previous visit experiences, gender and marital status are important factors to determine satisfaction (Ragavan, Subramonian & Sharif, 2014; Sarra, Di zio & Cappucci, 2015; Lu et al., 2015).

The income has influence in motivations and satisfactions of tourist. Acording to Jarvis et al. (2016), a person with a higher income tend to me more satisfaied with their trip. On the other hand, Lu et al. (2015) concluded that regarding with travel motivations and

tourist satisfaction, income has a negative influence which means that a higher income leads to a lower level of travel motivations and satisfaction.

Travel motivations have a positive correlation with shopping behaviour and satisfaction, which means that a tourist, motivated to travel, will have a higher consumption of products/experiences and a higher satisfaction too (Lu et al., 2015).

Motivation has been referred to as psychological/ biological needs and wants, including integral forces that arouse, direct, and integrate a person's behavior and activity (Dann, 1981; Woo, Et al, 2011).

In tourism research, this motivation concept can be classified into two forces, Push and pull motivations (Dann, 1981). Push motivations are related to the tourists' desire, while pull motivations are associated with the attributes of the destination choices (Cha, McCleary, & Uysal, 1995; Crompton, 1979; Dann, 1981; Woo, Et al, 2011).

Push motivations are related with the desire for rest and relaxation, the sensation of escape, in search of prestige, fitness and healt, social interacions and adventure, exicetment and family (Crompton, 1979).

Pull motivations explores the destination's attractiveness, such as cultural, nature or commerce such as shopping, parks, beaches, recreation facilities, cultural attractions, entertainment, natural scenery.

Studies have proven that push and pull motivations have been primarily utilized in studies of tourist behavior. The discoveries and issues undoubtedly play a useful role in attempting to understand a wide variety of different needs and wants that can motivate and influence tourist behavior.

For tourism companies maximize, the loyalty of tourists requires careful management of points of interest and all factors related to the image of the country thus ensuring the satisfaction of tourists.

c. Destination Image

Destination Image constitutes as an overall impression with some emotional condition (Oxenfeldt, 1974). That impression is a perceptual phenomenon that turns up in consumer's emotional interpretation with cognitive and affective components.

Destination image is an interactive system of thoughts, opinions, feelings, visualizations, and intentions toward a destination (Költringer & Dickinger, 2015; Tasci

& Gartner, 2007). An overall or composite result from interactions between these consumer choice attitudes (Lin, et al, 2007; Pike & Ryan, 2004).

Destination image influences tourists in the process of choosing a destination and revisiting the destination in the future (Hosany & Prayag, 2013; Zeugner-Roth, Diamantopoulos & Žabkar, 2015).

The concept of destination image is related with destination loyalty. It also has been operationalized as consisting of two components: a perceptual-cognitive component that captures knowledge and beliefs about a destination's attributes and an affective component that describes feelings toward a destination (Beerli & Martín, 2004).

The cognitive and affective components work to influence the overall image of a destination in the mind of past or prospective tourists (Baloglu & McCleary, 1999). Cognitive and affective image, with some of the cognitive components influencing the affective and overall image perceptions, suggest that cognitive image components are not evaluated all the same.

### d. Word-of-Mouth

Word-of-Mouth is defined as *"informal communication between private parties concerning evaluations of goods and services"* (Anderson 1998).

Word-of-mouth happens in the context of a specific situation (Allsop, Bassett, and Hoskins, 2007) and may contain both cognitive and emotive elements (Sweeney, Soutar, and Mazzarol, 2012).

The impact of word-of-mouth could be tremendous because consumers who have not experienced a product or service will to be credible and trustworthy coming from an experienced personal source (Allsop, Bassett, and Hoskins, 2007). It has been shown to be the information source most frequently used by tourists (e.g., Andereck and Caldwell 1994; Bieger and Laesser, 2004) and is therefore of critical importance to tourism destinations and businesses.

### 2.3.2. On travel

#### a. Frequency

Frequency, in this context, is the amount of travelling that a person does in a year. The number of trips carried out in a year usually varies between one and thirty trips per year, and in the case of travel, these tend to travel three times a year (Losada et al., 2016).

Available time, gender and economic status are the main characteristics that influence the frequency of travel. Studies say that, in terms of gender, the elderly women travel more frequently than men. This result may be related to the greater number of women in the third idea and the growth of female independence in the last decades (Vieira, 2016). Refer to the financial advice, and fulfilling the specs, people with more money tend to make more trips. Simultaneously, people with more available time tend to travel more (Losada et al., 2016).

b. Length of Stay

Duration or Length of stay (LOS) is the time (in days) that tourists stay travelling, i.e., staying in a place different than their homes. There are differente atributions and caharacteristics that influence de Length of Stay, such as travel cost, nationality, socio-demographic characteristics and destination attributes (Peypoch et al., 2012).

c. Behaviour

To adquire travel experiences, products or services are influenced by Tourist behaviour who represent tourist personality and characteristics. According to Gazley and Watling (2015) the fact that a person travels directly influences his behavior. People while travelling try different experiences, such as, meeting different people, going to an unusual kind of events, try new food and normally enjoying knowing and experience the host culture.

The behaviour is also influenced by the way a person expresses feelings and this is essentially shaped by experiences during the journey. Self-expression can influence behavior and is different from person to person.

People who loves to share their achieving at the travel to themselves and to others tend to be more self-expressive. Tourists who are less self-expressive will travel more and will not be concerned about acquiring products or experiences with the intent of showing others (Gazley & Watling, 2015).

A person who is more self-expressive tends to constantly be having different experiences and is less interested about the meaning of products and experiences they consume.

2.3.3.    Post-travel

Trust and consumer satisfaction are the seeds for behavioral loyalty not only because they increase attitudinal loyalty in a high-involvement, high-service product market but also because they directly or indirectly persuade the consumer to invest in specific assets. Marketers should not count on satisfaction alone to induce consumers to invest in specific assets. Loyalty programs and properly trained personnel are but two examples. In the future, database assets may prove critical. Marketers can build up a consumer database to accumulate data on past usage, purchases, complaining behaviours, and returns (Vieira, 2016).

a.   Satisfaction

The tourism experience is unique, emotionally charged, and with a high personal value (McIntosh & Siggs, 2005). Companies and Governaments should provide unforgettable, satisfactory, and extraordinary experiences to their customers (Nikolova & Hassan, 2013).

The uses of satisfaction, to evaluate past consumer experience, are the performance of products and services, and the perceptions of the physical environment such as a neighbourhood and tourist destinations in the tourism context (Ekinci, Preciado e Sirakaya-Turk, 2013; Ross & Iso-Ahola, 1991).

There is no definition of satisfaction derived from the fact that researchers are unable to reach a consensus on it. Yet this is usually associated with a feeling, affective and evaluative that evolves according to your experience of the product (Wu, 2015).

Oliver (1997) defines satisfaction as "*the consumer's fulfilment response, the degree to which the level of fulfilment is pleasant or unpleasant*". In a meta-analysis performed for satisfaction, Geyskens et al. (1999) found that this construct is the most popular one among empirical investigations of channel relationships. According to Ruekert and Churchill (1984), the construct of satisfaction is of key importance in understanding channel relationships that lead to loyalty.

Satisfaction is a cumulative construct that includes not only satisfaction with a specific product, service or travel experience but also with the various aspects of the organization, such as the physical facilities and the interaction with employees. Moreover, satisfaction is positively related to trust (Anderson & Narus, 1990). According to Michell, Reast, and

Lynch (1998), satisfaction is the foundation of trust. In this study satisfaction is releated not only with trust but also with loyalty.

b. Thrust

Trust has been defined as: (1) a set of specific beliefs dealing primarily with the integrity, benevolence, and ability of another party (Doney & Cannon, 1997); (2) a general belief that another party can be trusted (Gefen, 2000; Hosmer, 1995), sometimes also called trusting intentions (McKnight, Cummings & Chervany, 1998) or "the willingness' of a party to be vulnerable to the actions of another" (Mayer, Davis & Sshoorman 1995); (3) the effect reflected in "feelings" of confidence and security in the caring response of the other party (Yee, 2004).

Despite the different definition of trust that exist in the past literatures, according to Gefen et al (2003), the conceptualization of trust is seen as a set of specific beliefs which includes integrity, benevolence, ability and predictability to be applied in e-commerce, specifically online shopping context. The definition is aligned with the past literatures where it has been widely used in studies related to ongoing economic relationship dealing with buyer-seller and business interactions (Yee, 2004).

In consonance with the definition of trust of Gefen et al. (2003), Lin and Wang (2006) who conducted their study in the m-commerce context have adopted the same view by defining trust as a set of specific beliefs dealing primarily with the integrity (trustee honesty and promise keeping), benevolence (trustee caring and motivation to act in the truster's interest), competence (ability of trustee to do what the truster needs) and predictability (trustee's behavioral consistency) of a particular m-vendor (Yee, 2004).

# Chapter 3 - CRISP-DM

## 3.1. Business understanding

As mentioned in the literature review, the loyalty of tourists is influenced by satisfaction, and this, in other hand, is determined by different factors, whether be in pré-travel or on travel phase. Moreover, the socio-demographic characteristics of tourists may be associated to the loyalty type.

With that, it is extremely important to the tourism companies to have this knowledge and try to predict where tourists will travel. In addition, there are several countries that have a high dependence on tourism which has a high impact on the local economy.

Therefore, the data studied is from Flash Eurobarometer 414 "Preferences of Europeans towards tourism 2015". Moreover, only the tourists from EU who travelled at least once from 2014 to 2016.

Regarding the business goals, the first one is evaluate the loyalty of European tourists to the country of destination in a global prespective. The second objective is to identify profiles of loyal tourists with respect to their destination country. These two objectives allow travel agencies and Nacional organisms to better comprehend the travel preferences of the tourists and to develop better marketing campaigns promoting the destination.

In order to identify profiles of European tourists in a Big data prespective, it was necessary to gather the data collected in the different years, as well as to select the valid answers for analysis based on a treatment of variables and a filtering of answers. Finally, different classification algorithms or parametrizations should be tested in order to extrat useful knowledge.

The resources to be used are the IBM SPSS Statistics (v. 23) and the IBM SPSS Modeler (v. 18.).

### 3.2. Data understanding and preparation

As mentioned before, the data used in this study was collected from the Flash Eurobarometer 414 "Preferences of Europeans towards tourism" in the 28 countries of the European Union, conducted by TNS Political & Social at the request of the European Commission, Directorate-General for Internal Market, Industry, Entrepreneurship and SMEs (European Commission, 2015). The interviews were conducted by phone in the respective national language. The analysis already made to this questionnaire is very simple and descriptive and do not explore complex relationships among variables.

The data collected refer to the years 2014, 2015 and 2016, and it was necessary to combine the data so that it is possible to carry out the analyzes. This process forced a variable treatment since the questionnaires presented different variables with different structures, and also, different names were assigned to the variables according to the year. After the treatment of the variables and identification of their names, the data understang was started.

The data available consists of 91.328 tourists (around 30.000 for each year) and 629 variables, which need to be observed and cleaned in order to meet this study goals. For this study, only the tourists from EU who have traveled at least once to a European country and pretend to travel again to a European country in the follow year.

After analyzing the final data, it turned out that most people had responded that they wanted to travel to more than one country, just as they had traveled to more than one country last year. On the other hand, since few valid results were obtained, a transformation of the data was carried out, where it was considered only a trip made and a trip that was intended to travel. After this transformation the sample comprises 73.252 tourists.

From all the variables included in de data file retrieved, only the ones that can be related to tourist loyalty were maintained, resulting in 114 final variables to be studied. These variables were divided into three different subgroups: socio-demographic characteristics, holidays taken last year, and holidays plan this year.

Given the choice of increasing the number of entries, we did not take into account the satisfaction variables since the data did not allow us to know the satisfaction level for all trips.

In the data preparation phase, it was necessary to create different variables that allowed results more consistent in the modeling phase. Thus, both the nationality variable, holiday taken last year, and the variable holidays plan this year, were transformed simultaneously.

It was used to create a new variable D3_r1 (area of the European Union that the country of nationality belongs) to split the nationality countries into the European demographic divisions.

It was also created the variable D3_r2 (currency of the nationality country) to assign the value "Euro" or "Non-euro" to each country. The same was complete for the remaining variables (Holidays taken and holidays plan).

The country division was made using the following characterization (Vieira, 2016):

- **North European countries:** Denmark, Finland and Sweden.
- **Central European countries:** Austria, Belgium, France, Germany, Ireland, Italy, Luxembourg, Malta, Netherlands and United Kingdom.
- **Countries of the Iberian Peninsula:** Portugal and Spain.
- **Eastern European countries:** Czech Republic, Hungary, Poland, Rep. of Cyprus and Romania.
- **Balkan countries:** Bulgaria, Greece, Latvia, Republic of Croatia, Slovakia and Slovenia.
- **Baltics:** Estonia and Lithuania.

The creation of the variable D3_R2 was made using the following information (UE, 2016):

- **Euro:** Germany, Austria, Belgium, Cyprus, Slovakia, Slovenia, Spain, Estonia, Finland, France, Greece, Ireland, Italy, Latvia, Lithuania, Luxembourg, Malta, Netherlands and Portugal.
- **Non-euro:** Bulgaria, Croatia, Hungary, Poland, Czech Republic, Romania, Sweden, Denmark and United Kingdom.

Six new variables were also created to categorize whether planned and realized trips were made to the own country, the foreign country, and the neighboring country.

Table 1 to 5 are related to the tourist characterization and social demographic new variables.

*Table 1 - Nationality*

| Nationality related characteristics | | Nº | % |
|---|---|---|---|
| **Nationality** | Austria | 4909 | 6.7 |
| | Belgium | 3330 | 4.5 |
| | Bulgary | 1438 | 2.0 |
| | Croatia | 1316 | 1.8 |
| | Cyprus | 949 | 1.3 |
| | Czech Republic | 3558 | 4.9 |
| | Denmark | 3992 | 5.4 |
| | Estonia | 1040 | 1.4 |
| | Finland | 3761 | 5.1 |
| | France | 3707 | 5.1 |
| | Germany | 4844 | 6.6 |
| | Greece | 1454 | 2.0 |
| | Hungary | 1597 | 2.2 |
| | Irland | 4384 | 6.0 |
| | Italy | 2645 | 3.6 |
| | Latvia | 1058 | 1.4 |
| | Lithuania | 762 | 1.0 |
| | Luxembourg | 1991 | 2.7 |
| | Malta | 718 | 1.0 |
| | Netherlands | 4939 | 6.7 |
| | Poland | 3218 | 4.4 |
| | Portugal | 1152 | 1.6 |
| | Romenia | 1303 | 1.8 |
| | Slovakia | 2590 | 3.5 |
| | Slovenia | 1759 | 2.4 |
| | Spain | 2313 | 3.2 |
| | Sweden | 3938 | 5.4 |
| | United Kingdom | 4587 | 6.3 |
| | **Total** | **73252** | **100.0** |
| **Nationality region** | North European Countries | 11691 | 16.0 |
| | Central Europena Countries | 36054 | 49.2 |
| | Countries of the Iberian Peninsula | 3465 | 4.7 |
| | Eastern European Countries | 10625 | 14.5 |
| | Balkan Countries | 9615 | 13.1 |
| | Baltics | 1802 | 2.5 |
| | **Total** | **73252** | **100.0** |
| **National currency** | Euro | 48305 | 65.9 |
| | Non Euro | 24947 | 34.1 |
| | **Total** | **73252** | **100.0** |

*Table 1 - Nationality*

*Table 2 - Distribution of the socio-demographic characteristics*

| Social-demographic characteristics | | Nº | % |
|---|---|---:|---:|
| **Age binned** | 15 - 24 Years | 4989 | 6.8 |
| | 25 - 39 Years | 12699 | 17.3 |
| | 40 - 54 Years | 21299 | 29.1 |
| | 55 Years and older | 34265 | 46.8 |
| | **Total** | **73252** | **100.0** |
| **Gender** | Male | 33413 | 45.6 |
| | Female | 39839 | 54.4 |
| | Total | 73252 | 100.0 |
| **Age education (11 categories)** | Up to 14 years | 1841 | 2.5 |
| | 15 years | 2084 | 2.9 |
| | 16 years | 3550 | 4.9 |
| | 17 years | 3355 | 4.6 |
| | 18 years | 9374 | 12.9 |
| | 19 years | 6460 | 8.9 |
| | 20 years | 4896 | 6.7 |
| | 21 years | 4768 | 6.6 |
| | 22 years and older | 32306 | 44.4 |
| | Still studying | 3922 | 5.4 |
| | No full-time education | 184 | 0.3 |
| | **Total** | **72740** | **100.0** |
| **Age education (5 categories)** | Up to 15 years | 3925 | 5.4 |
| | 16-19 years | 22739 | 31.3 |
| | 20 Years and older | 41970 | 57.7 |
| | Still studying | 3922 | 5.4 |
| | No full-time education | 184 | 0.3 |
| | **Total** | **72740** | **100.0** |
| **Type of community** | Rural area or vilalge | 21784 | 29.9 |
| | Small or middle sized town | 27615 | 37.8 |
| | Large town | 23566 | 32.3 |
| | **Total** | **72965** | **100.0** |
| **Household size - aged 15 (binned)** | 1 | 13941 | 19.1 |
| | 2 | 38080 | 52.1 |
| | 3 | 11340 | 15.5 |
| | 4+ | 9701 | 13.3 |
| | **Total** | **73062** | **100.0** |
| **Mobile phone** | Yes | 71070 | 97.1 |
| | No | 2144 | 2.9 |
| | **Total** | **73214** | **100.0** |
| **Landline (phone)** | Yes | 57642 | 78.7 |
| | No | 15572 | 21.3 |
| | Total | 73214 | 100.0 |
| **Phone available** | Mobile Only | 15572 | 21.3 |
| | Landline Only | 2144 | 2.9 |
| | Mobile and landline | 55498 | 75.8 |
| | **Total** | **73214** | **100.0** |

*Table 2 - Distribution of the socio-demographic characteristics*

*Table 3 - Distribution of the occupation related characteristics*

| Occupation characteristics | | Nº | % |
|---|---|---|---|
| **Occupation** | Farmer, forester, fisherman (self-employed) | 475 | 0.7 |
| | Owner of Shop, craftsman (self-employed) | 1495 | 2.0 |
| | Professional (sel-employed lawyer, medical practitioner, ...) | 2767 | 3.8 |
| | Manager of a company (self-employed) | 2091 | 2.9 |
| | Other (self-employed) | 842 | 1.2 |
| | Professional (employed doctor, lawyer, accountant, architect, ...) | 5733 | 7.8 |
| | General management, director or top management | 2051 | 2.8 |
| | Middle management | 6176 | 8.5 |
| | Civil servant | 5556 | 7.6 |
| | Office clerk | 4211 | 5.8 |
| | Other employee (salesman, nurse, ...) | 6830 | 9.3 |
| | Other (employee) | 452 | 0.6 |
| | Supervision\ foreman (team manager, ...) | 667 | 0.9 |
| | Manual worker | 2724 | 3.7 |
| | Unskilled manual worker | 546 | 0.7 |
| | Other (manual worker) | 85 | 0.1 |
| | Looking after the home | 2946 | 4.0 |
| | Student (full time) | 3695 | 5.1 |
| | Retired | 21574 | 29.5 |
| | Sneeking a job | 1632 | 2.2 |
| | Other (without a professional activity) | 524 | 0.7 |
| | **Total** | **73072** | **100.0** |
| **Occupation type** | Self-employed | 7670 | 10.5 |
| | Employees | 31009 | 42.4 |
| | Manual Workers | 4022 | 5.5 |
| | Not Working | 30371 | 41.6 |
| | **Total** | **73072** | **100.0** |

*Table 3 - Distribution of the occupation related characteristics*

*Table 4 - Holidays Taken Last Year*

| Holidays taken last year | | N° | % |
|---|---|---|---|
| **Country** | Austria | 4577 | 6.2 |
| | Belgium | 1893 | 2.6 |
| | Bulgary | 1317 | 1.8 |
| | Croatia | 2331 | 3.2 |
| | Cyprus | 978 | 1.3 |
| | Czech Republic | 1985 | 2.7 |
| | Denmark | 1731 | 2.4 |
| | Estonia | 1263 | 1.7 |
| | Finland | 3259 | 4.4 |
| | France | 7306 | 10.0 |
| | Germany | 6110 | 8.3 |
| | Greece | 3161 | 4.3 |
| | Hungary | 1830 | 2.5 |
| | Irland | 3112 | 4.2 |
| | Italy | 4750 | 6.5 |
| | Latvia | 566 | 0.8 |
| | Lithuania | 521 | 0.7 |
| | Luxembourg | 351 | 0.5 |
| | Malta | 939 | 1.3 |
| | Netherlands | 2393 | 3.3 |
| | Poland | 2423 | 3.3 |
| | Portugal | 1678 | 2.3 |
| | Romenia | 1057 | 1.4 |
| | Slovakia | 1291 | 1.8 |
| | Slovenia | 3056 | 4.2 |
| | Spain | 6017 | 8.2 |
| | Sweden | 3252 | 4.4 |
| | United Kingdom | 4105 | 5.6 |
| | **Total** | **73252** | **100.0** |
| **Region** | North European Countries | 12994 | 17.7 |
| | Central Europena Countries | 30784 | 42.0 |
| | Countries of the Iberian Peninsula | 7695 | 10.5 |
| | Eastern European Countries | 8273 | 11.3 |
| | Balkan Countries | 11722 | 16.0 |
| | Baltics | 1784 | 2.4 |
| | **Total** | **73252** | **100.0** |
| **Currency** | Euro | 53223 | 72.7 |
| | Non Euro | 20029 | 27.3 |
| | **Total** | **73252** | **100.0** |

*Table 5- Holidays plan this year*

| Holidays plan this year | | N° | % |
|---|---|---:|---:|
| **Country** | Austria | 4303 | 5.9 |
| | Belgium | 1230 | 1.7 |
| | Bulgary | 1293 | 1.8 |
| | Croatia | 2857 | 3.9 |
| | Cyprus | 866 | 1.2 |
| | Czech Republic | 1694 | 2.3 |
| | Denmark | 1444 | 2.0 |
| | Estonia | 1095 | 1.5 |
| | Finland | 3325 | 4.5 |
| | France | 7102 | 9.7 |
| | Germany | 5251 | 7.2 |
| | Greece | 3508 | 4.8 |
| | Hungary | 1573 | 2.1 |
| | Irland | 3506 | 4.8 |
| | Italy | 6001 | 8.2 |
| | Latvia | 487 | 0.7 |
| | Lithuania | 510 | 0.7 |
| | Luxembourg | 287 | 0.4 |
| | Malta | 779 | 1.1 |
| | Netherlands | 1881 | 2.6 |
| | Poland | 2293 | 3.1 |
| | Portugal | 1970 | 2.7 |
| | Romenia | 935 | 1.3 |
| | Slovakia | 1153 | 1.6 |
| | Slovenia | 3393 | 4.6 |
| | Spain | 7073 | 9.7 |
| | Sweden | 3264 | 4.5 |
| | United Kingdom | 4179 | 5.7 |
| | **Total** | **73252** | **100.0** |
| **Region** | North European Countries | 14034 | 19.2 |
| | Central Europena Countries | 28518 | 38.9 |
| | Countries of the Iberian Peninsula | 9043 | 12.3 |
| | Eastern European Countries | 7361 | 10.0 |
| | Balkan Countries | 12691 | 17.3 |
| | Baltics | 1605 | 2.2 |
| | **Total** | **73252** | **100.0** |
| **Currency** | Euro | 53720 | 73.3 |
| | Non Euro | 19532 | 26.7 |
| | **Total** | **73252** | **100.0** |

*Table 5- Holidays plan this year*

### 3.3. Modelling

In order to reach the desired objectives, decision tree models were created, which are the easier way to comprehend and analyze, especially using a large number of data (Chen 2003). CHAID (chi-squadred automatic interection detector) and the regression and classification tree (CART) were tested in a specific way,

CHAID is a widely used method of segmentation of tourism and this is based on a categorization of at least two variables where it allows analyzing and segmenting with the use of independent variables.

The advantage of using this model is that it accepts any type of variable accepting both nominal, interval or continuous as independet variable (predictors). Another advantage is related to the application objectives of this model, which is a very versatile model and suits different uses.

CART is an algorithm that construct binary trees. This is designed for the use of static databases i.e. with boundary. This translates into a learning model that becomes more specific until the best result is achieved. Being a decision tree-based model, this is composed by nodes and branches.

As already mentioned, the advantage of using these two models is its ease of comprehension, in this way, both models were tested and evaluated. In this way a validation was performed through the training and test sample and this forced to balance the weight of loyal and non-loyal tourists. This process was carried out through the random exclusion of loyal tourists. This phase was essential derived. Otherwise the model would predict practically all tourists as loyal (what happened in molel G of CHAID)

*Table 6 - Predicting models to classify loyalty using CART algorithm*

| Parametrization | Model A<br>CART (R) | Model B<br>CART (R1) | Model C<br>CART(R2) | Model D<br>CART (R3)<br>boosting | Model E<br>CART (R4)<br>bagging<br>(voting) | Model F<br>CART (R5)<br>bagging<br>(voting) |
|---|---|---|---|---|---|---|
| Tree depth | 5 | 9 | 11 | 9 | 11 | 11 |
| Minimum records in parent branch | 2 | 2 | 2 | 2 | 2 | 2 |
| Minimum records in child node | 1 | 1 | 1 | 1 | 1 | 1 |
| Balanced sample | Yes | Yes | Yes | Yes | Yes | Yes |
| Training sample | 70% | 70% | 70% | 70% | 70% | 70% |
| Testing sample | 30% | 30% | 30% | 30% | 30% | 30% |

*Table 7- Predicting models to classify loyalty using CHAID algorithm*

| Parametrization | Model G CHAID (R) | Model H CHAID (R1) | Model I CHAID (R2) | Model J CHAID (R3) | Model K CHAID (R4) boosting (10) | Model L CHAID (R5) boosting (10) | Model M CHAID (R6) boosting (10) | Model N CHAID (R7) bagging (voting - 10) | Model O CHAID (R8) bagging (voting - 10) |
|---|---|---|---|---|---|---|---|---|---|
| Tree depth | 5 | 5 | 9 | 11 | 5 | 7 | 9 | 9 | 11 |
| Minimum records in parent branch | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 |
| Minimum records in child branch | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| Significance level | 0.05 | 0.05 | 0.05 | 0.05 | 0.05 | 0.05 | 0.05 | 0.05 | 0.05 |
| Balanced sample | No | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Training sample | 70% | 70% | 70% | 70% | 70% | 70% | 70% | 70% | 70% |
| Testing sample | 30% | 30% | 30% | 30% | 30% | 30% | 30% | 30% | 30% |

In order to evaluate the models, different techniques were used, namely accuracy, sensibility, specificity, AUC - ROC and Gini.

One of the measures most used in relation to the evaluation of models is the accuracy, which measures the rate of correctness, regardless of the class. Sensibility evaluate the percentage of true positives that a model produces. Specificity evaluates the percentage of true negatives that a model produces (Delen et al., 2013).

QUINLAN (1993) prosed the Gain Ratio, that is how much information which is what the information gain as an evaluation criterion. In the first step, the information gain for all attributes is calculated. It is then selected which data gives more information.

The Italian Corrado Gini in 1912 proposed a model which employs a statistical dispersion index. This index is widely used in economic and social analysis, for example, to quantify the distribution of income in each country. It is used in the CART algorithm (BREIMAN et al., 1984). Just as in calculating the information gain, simply calculate the difference between gini index before and after division. Thus, you select the attribute that generates a higher value for Gini.

A classification and avialization model for the so - called AUC - ROC models was also used. ROC is a probability curve and the AUC serve to measure the separability. In short, it conveys the information on how efficient a model is and whether it is able to distinguish the variables. A model with better AUC is usually a better in terms of distributing the variables in this case in tourist loyalty.

## Chapter 4 – Ressults and Discussion

To achieve the first goal outlined, which refers to the identification of pairs of countries visited, a crosstable between travel taken last year and travel plan this year was made. In table 8 it is possible to identify the countries where there is a stronger loyalty, as well as the destination travel trends.

The first result that stands out, refers to the tendency of tourists to return to their country of destination. In this way, regardless of the destination site, the next year the highest percentage for the planned trip goes to that destination.

The biggest example of this is the Bulgary, where tourists who traveled to Bulgary last year, say that this year most people intend to return to Bulgary (51.0%) followed by Greece (9.9%) and Germany (6.4%).

Another example is Poland, where tourists who traveled to Poland last year have a tendency to return to that country (41.4%) followed by Italy (5.1%) and Germany (4.8%).

These results, whether relating to Bulgaria or Poland can be explained derived from the fact that most people who traveled to these countries are people of that nationality.

On the other hand, concerning tourists who traveled to Luxembourg, there is a preference to travel to France (14%) and only then appears the intention to return to the country 11.4%. Finally, 10.0% of the tourists who traveled to Luxembourg showed interest in traveling to Germany.

This result is explained given the size of the country, so that there is no return trend to explore the country. France and Germany are neighboring countries and can explain the destination country alternatives.

Regarding to Portugal, the tourists showed an intention to return to the country, being this their first destination of preference (35.8%). Tourists then show the destination country as the neighboring country of Spain (14.7%) followed by Italy (8.3%) and France (8.2%).

The results regarding portugal can be explained derived from the existence of several points of interest throughout the country and the fact that the interest of the tourists by the Iberian Peninsula wants to explore it. Italy and France are considered the countries with a culture more similar to the countries of the peninsula being a trend of destination countries.

Table 8 - Crosstable - Holidays taken last year x Holiday plan this year.

| Holiday plan this year | Holidays taken last year | | | | | | | | | | | | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Austria | Belgium | Bulgaria | Croatia | Cyprus | Czech Republic | Denmark | Estonia | Finland | France | Germany | Greece | Hungary | Irland | Italy | Latvia | Lithuania | Luxembourg | Malta | Netherlands | Poland | Portugal | Romenia | Slovakia | Slovenia | Spain | Sweden | United Kingdom |
| Austria | 28.4% | 4.2% | 2.5% | 5.5% | 4.0% | 5.8% | 3.5% | 2.1% | 3.1% | 4.8% | 7.2% | 4.1% | 5.1% | 5.3% | 4.9% | 3.5% | 3.5% | 5.1% | 4.7% | 4.5% | 2.6% | 2.7% | 4.1% | 6.7% | 4.4% | 3.3% | 2.6% | 3.6% |
| Belgium | 0.8% | 17.1% | 0.5% | 0.6% | 0.4% | 0.4% | 0.6% | 0.5% | 2.0% | 2.4% | 1.8% | 0.9% | 0.7% | 0.8% | 1.1% | 0.7% | 2.3% | 5.4% | 2.4% | 3.6% | 0.5% | 1.6% | 0.5% | 0.5% | 0.9% | 1.1% | 0.9% | 1.0% |
| Bulgaria | 0.8% | 0.6% | 51.0% | 0.8% | 1.5% | 1.5% | 0.3% | 0.5% | 0.2% | 0.5% | 1.0% | 3.2% | 1.2% | 0.6% | 0.8% | 0.4% | 1.2% | 0.9% | 0.5% | 0.8% | 0.9% | 0.5% | 3.4% | 1.9% | 0.5% | 0.5% | 0.3% | 0.8% |
| Croatia | 5.6% | 1.4% | 1.5% | 40.6% | 1.1% | 6.9% | 1.2% | 1.0% | 0.7% | 1.7% | 3.1% | 2.1% | 6.9% | 0.5% | 4.1% | 1.9% | 2.3% | 1.4% | 0.6% | 1.6% | 3.4% | 1.4% | 2.2% | 7.5% | 5.9% | 1.8% | 1.5% | 1.3% |
| Cyprus | 1.0% | 0.5% | 0.8% | 0.4% | 27.3% | 0.7% | 0.6% | 0.3% | 0.9% | 0.6% | 0.4% | 3.0% | 0.4% | 1.3% | 0.3% | 0.7% | 0.6% | 0.9% | 1.2% | 1.3% | 0.2% | 0.5% | 3.2% | 0.7% | 0.6% | 0.5% | 0.8% | 1.5% |
| Czech Republic | 2.1% | 1.1% | 1.3% | 2.8% | 1.4% | 28.4% | 1.2% | 1.4% | 0.8% | 1.4% | 2.0% | 1.5% | 2.6% | 0.9% | 1.4% | 1.2% | 1.9% | 1.1% | 1.1% | 0.9% | 3.3% | 0.3% | 0.9% | 9.4% | 1.2% | 1.1% | 0.9% | 0.8% |
| Denmark | 1.1% | 1.0% | 0.4% | 0.7% | 0.9% | 1.0% | 27.4% | 3.0% | 0.8% | 1.2% | 3.0% | 1.0% | 0.4% | 0.5% | 1.3% | 2.5% | 1.5% | 0.9% | 0.7% | 1.3% | 0.9% | 1.1% | 0.3% | 0.5% | 0.8% | 1.8% | 3.3% | 1.0% |
| Estonia | 0.5% | 0.6% | 0.5% | 0.4% | 1.0% | 0.6% | 3.2% | 29.3% | 3.9% | 0.6% | 0.6% | 0.6% | 0.7% | 0.4% | 0.6% | 6.5% | 5.0% | 0.9% | 0.4% | 0.4% | 0.9% | 0.4% | 0.4% | 0.6% | 0.9% | 1.3% | 2.5% | 0.2% |
| Finland | 2.1% | 5.0% | 0.8% | 1.3% | 3.3% | 2.7% | 2.3% | 14.2% | 37.6% | 2.7% | 1.9% | 0.9% | 2.8% | 6.3% | 0.8% | 4.1% | 3.8% | 2.8% | 7.7% | 1.6% | 3.0% | 1.5% | 2.0% | 1.2% | 6.8% | 2.0% | 7.4% | 1.7% |
| France | 7.1% | 15.9% | 3.0% | 3.6% | 5.0% | 5.4% | 6.8% | 3.6% | 5.1% | 34.6% | 7.4% | 4.6% | 3.0% | 6.7% | 10.0% | 4.6% | 3.8% | 14.0% | 9.6% | 10.5% | 3.5% | 8.2% | 4.2% | 3.2% | 4.9% | 9.7% | 5.1% | 8.9% |
| Germany | 7.9% | 7.8% | 6.4% | 5.2% | 4.3% | 4.9% | 8.7% | 4.2% | 3.2% | 5.2% | 27.6% | 4.8% | 3.9% | 4.2% | 6.1% | 4.6% | 3.5% | 10.0% | 3.8% | 9.2% | 4.8% | 4.3% | 5.0% | 3.1% | 4.4% | 4.8% | 4.5% | 4.7% |
| Greece | 3.6% | 2.6% | 9.9% | 3.4% | 10.9% | 3.6% | 2.8% | 2.8% | 1.4% | 2.5% | 3.2% | 40.3% | 2.8% | 1.3% | 4.4% | 2.7% | 2.3% | 2.6% | 1.5% | 2.7% | 2.3% | 2.9% | 6.9% | 4.1% | 0.9% | 3.3% | 2.4% | 4.1% |
| Hungary | 1.7% | 0.8% | 0.6% | 2.2% | 0.8% | 2.2% | 0.9% | 0.8% | 1.6% | 0.7% | 1.4% | 1.2% | 33.8% | 1.1% | 1.1% | 0.9% | 0.4% | 2.3% | 0.6% | 0.9% | 2.2% | 0.5% | 2.8% | 4.0% | 2.4% | 0.6% | 2.6% | 0.7% |
| Irland | 4.5% | 3.3% | 1.8% | 1.3% | 5.4% | 1.9% | 2.0% | 1.9% | 7.8% | 4.0% | 2.0% | 1.3% | 2.6% | 35.1% | 1.7% | 1.9% | 1.3% | 5.4% | 7.3% | 2.8% | 2.3% | 3.6% | 2.6% | 2.1% | 7.0% | 2.7% | 3.8% | 6.3% |
| Italy | 9.3% | 7.2% | 3.1% | 8.4% | 2.5% | 7.3% | 6.0% | 3.7% | 1.7% | 8.3% | 8.7% | 7.7% | 6.3% | 2.3% | 34.1% | 4.2% | 5.4% | 5.1% | 4.6% | 6.1% | 5.1% | 8.3% | 4.8% | 5.4% | 2.2% | 8.7% | 4.6% | 6.3% |
| Latvia | 0.3% | 0.3% | 0.3% | 0.2% | 0.4% | 0.7% | 0.8% | 2.7% | 0.5% | 0.3% | 0.5% | 0.3% | 0.4% | 0.3% | 0.3% | 27.4% | 6.9% | 0.3% | 0.4% | 0.3% | 0.9% | 0.2% | 0.4% | 0.6% | 0.2% | 0.2% | 0.6% | 0.2% |
| Lithuania | 0.5% | 1.0% | 0.3% | 0.3% | 0.4% | 0.6% | 0.5% | 1.6% | 0.7% | 0.3% | 0.5% | 0.3% | 0.4% | 0.0% | 0.3% | 8.8% | 29.4% | 0.3% | 0.3% | 0.2% | 1.0% | 0.2% | 0.2% | 0.9% | 0.2% | 0.3% | 0.4% | 0.4% |
| Luxembourg | 0.5% | 0.9% | 0.2% | 0.1% | 0.1% | 0.1% | 0.1% | 0.2% | 0.3% | 0.5% | 0.5% | 0.2% | 0.2% | 0.4% | 0.3% | 0.4% | | 11.4% | 0.2% | 0.8% | 0.2% | 0.5% | 0.2% | 0.3% | 0.4% | 0.1% | 0.5% | 0.2% |
| Malta | 0.6% | 1.8% | 0.6% | 0.3% | 0.8% | 0.3% | 0.8% | 0.3% | 2.0% | 1.2% | 0.5% | 0.4% | 0.7% | 1.2% | 0.6% | 0.2% | 1.5% | 0.6% | 21.7% | 0.6% | 0.7% | 0.2% | 0.3% | 0.4% | 1.4% | 0.6% | 0.6% | 0.9% |
| Netherlands | 1.5% | 4.9% | 0.8% | 1.2% | 1.8% | 0.7% | 1.3% | 1.0% | 0.9% | 2.3% | 3.1% | 1.5% | 1.1% | 1.3% | 1.7% | 1.4% | 1.0% | 5.1% | 1.6% | 28.5% | 0.9% | 1.2% | 1.5% | 0.5% | 0.9% | 1.6% | 1.3% | 1.9% |
| Poland | 1.3% | 1.1% | 0.8% | 1.5% | 1.4% | 3.6% | 1.4% | 1.4% | 2.6% | 1.2% | 2.7% | 0.7% | 2.7% | 1.7% | 1.8% | 3.0% | 4.4% | 2.0% | 3.8% | 1.0% | 41.4% | 1.2% | 1.6% | 3.6% | 3.3% | 1.2% | 2.2% | 1.3% |
| Portugal | 1.5% | 2.7% | 1.1% | 0.9% | 0.5% | 1.1% | 1.5% | 1.1% | 0.7% | 3.2% | 1.9% | 1.8% | 1.0% | 1.7% | 2.5% | 2.3% | 1.9% | 2.8% | 1.2% | 1.8% | 1.2% | 35.8% | 0.9% | 0.9% | 0.4% | 3.9% | 1.0% | 2.6% |
| Romenia | 1.2% | 0.4% | 1.6% | 0.5% | 5.1% | 0.3% | 0.3% | 0.4% | 0.5% | 0.4% | 0.6% | 1.2% | 2.5% | 0.9% | 0.8% | 0.7% | 1.0% | | 0.4% | 1.1% | 0.6% | 0.5% | 37.7% | 1.2% | 0.3% | 0.4% | 0.1% | 0.8% |
| Slovakia | 1.6% | 0.5% | 1.3% | 2.6% | 0.5% | 6.3% | 0.1% | 0.5% | 0.3% | 0.4% | 0.9% | 0.9% | 3.4% | 0.8% | 1.0% | 1.2% | 2.5% | 0.6% | 0.5% | 0.3% | 2.2% | 0.5% | 1.2% | 28.8% | 0.7% | 0.6% | 0.3% | 1.3% |
| Slovenia | 3.2% | 2.9% | 1.4% | 7.3% | 3.5% | 3.4% | 1.9% | 3.3% | 8.1% | 3.1% | 2.1% | 1.0% | 4.8% | 7.6% | 1.1% | 1.4% | 1.0% | 5.1% | 8.6% | 2.6% | 4.2% | 0.6% | 1.6% | 1.6% | 38.6% | 1.2% | 5.6% | 1.1% |
| Spain | 4.6% | 8.7% | 4.4% | 3.6% | 3.6% | 6.4% | 9.5% | 9.0% | 3.6% | 8.8% | 8.4% | 7.3% | 4.0% | 6.2% | 9.7% | 5.3% | 5.8% | 5.1% | 4.6% | 7.9% | 4.7% | 14.7% | 4.7% | 4.5% | 2.5% | 39.5% | 6.5% | 10.6% |
| Sweden | 1.9% | 2.1% | 0.8% | 2.2% | 1.8% | 1.6% | 10.3% | 6.7% | 7.1% | 2.5% | 3.2% | 1.6% | 3.9% | 3.4% | 2.2% | 4.2% | 3.5% | 5.1% | 4.5% | 1.8% | 2.8% | 2.1% | 1.0% | 0.6% | 5.8% | 2.9% | 33.6% | 2.7% |
| United Kingdom | 4.7% | 4.0% | 2.4% | 1.9% | 10.0% | 1.7% | 4.2% | 2.6% | 2.0% | 4.8% | 3.6% | 5.5% | 1.7% | 7.0% | 4.9% | 3.2% | 2.5% | 2.8% | 5.2% | 5.1% | 3.3% | 4.4% | 5.5% | 5.1% | 1.5% | 4.1% | 4.3% | 33.2% |
| Total | 100.0% | 100.0% | 100.0% | 100.0% | 100.0% | 100.0% | 100.0% | 100.0% | 100.0% | 100.0% | 100.0% | 100.0% | 100.0% | 100.0% | 100.0% | 100.0% | 100.0% | 100.0% | 100.0% | 100.0% | 100.0% | 100.0% | 100.0% | 100.0% | 100.0% | 100.0% | 100.0% | 100.0% |

Notes: n=73250; χ2(729)=30039.05;p<0.001; V de Cramer=0.318

This dissertation aims to verify, evaluate and categorize the loyalty of European tourists to the destination country. In order to achieve this goal, and answer the objective number two, evaluating the loyalty of tourists, whether on an individual level (trips to the own country, foreign country and strangeness). To that end, he was considered a loyal tourist if he returned to his country last year.

In Table 9 it's possible to verify the distribution of country type visited and to visit. In this way, it is possible to verify that the majority of the people (always superior to 70%) did not travel to its own country. Being that this questionnaire was done to people who have made at least one trip by plane to a European country, it is normal to have a majority for trips to the stranger. However, travel to the country itself accounts for about 25% of travel (26.4%) and planned travel (25.4%).

*Table 9 - Distribution of country type visited and to visit*

| Country visited and to visit | | Nº | % |
|---|---|---|---|
| Country visited | Abroad | 53917 | 73.6 |
| | Own country | 19335 | 26.4 |
| | **Total** | **73252** | **100.0** |
| Country to visit | Abroad | 54680 | 74.6 |
| | Own country | 18572 | 25.4 |
| | **Total** | **73252** | **100.0** |

Referring to table 10, this shows the loyalty of the tourists that in general, either referring to a specific country as well as by type of loyalty (tourist loyal to the country itself, loyal to a stranger and not loyal country).

As it is possible to verify in table 10, 77.2% are loyal, that is, they intend to return to the country where they had vacations. On the other hand, it is also possible to verify that 62.7% are loyal to a foreign country, that is, they had vacations last year in a foreign country and intend again to leave their own country this year again. Concerning to people traveling to their own country, it is possible to say that 14.5% of people are loyal. Finally, it is important to note that 33.6% of people are loyal to a specific country regardless of whether or not this is their own country.

*Table 10 - Distribution of loyalty related chacacteristics*

| Loyalty related chacacteristics | | Nº | % |
|---|---|---|---|
| **Loyalty** | No | 16701 | 22.8 |
| | Yes | 56551 | 77.2 |
| | **Total** | **73252** | **100.0** |
| **Loyalty to a specific country** | No | 48672 | 66.4 |
| | Yes | 24580 | 33.6 |
| | **Total** | **73252** | **100.0** |
| **Loyalty type** | Abroad | 45948 | 62.7 |
| | Own country | 10603 | 14.5 |
| | Non-loyal | 16701 | 22.8 |
| | **Total** | **73252** | **100.0** |

Table 11 refers to a cross-table comparing the countries visited and country to visit by type, that is, whether it is your own country or a stranger country. As already mentioned, it is possible to verify that if a person has made a trip to the stranger, in the following 85.2% intends to return to a stranger country. On the other hand, if a person traveled this year to their own country, there is a tendency (54.8%) of this to choose their own country as holiday destination.

*Table 11 - Distribution of country type to visit this year by country type visited last year*

| | | | Country visited | | Total |
|---|---|---|---|---|---|
| | | | Abroad | Own country | |
| **Country to visit** | Abroad | Nº | 45948 | 8732 | **54680** |
| | | % | 85.2% | 45.2% | **74.6%** |
| | Own country | Nº | 7969 | 10603 | **18572** |
| | | % | 14.8% | 54.8% | **25.4%** |
| **Total** | | **Nº** | **53917** | **19335** | **73252** |
| | | **%** | **100.0%** | **100.0%** | **100.0%** |

Notes: $\chi2(1)=12066.64$; $p<0.001$; V de Cramer=0.406

Regarding the third objective of identifying socio-demographic characteristics that could explain loyalty, the CART and CHAID models were applied. These were tested for different parameterizations and compared the models based on the ROC and the gain chart (Figures 2 to 5).
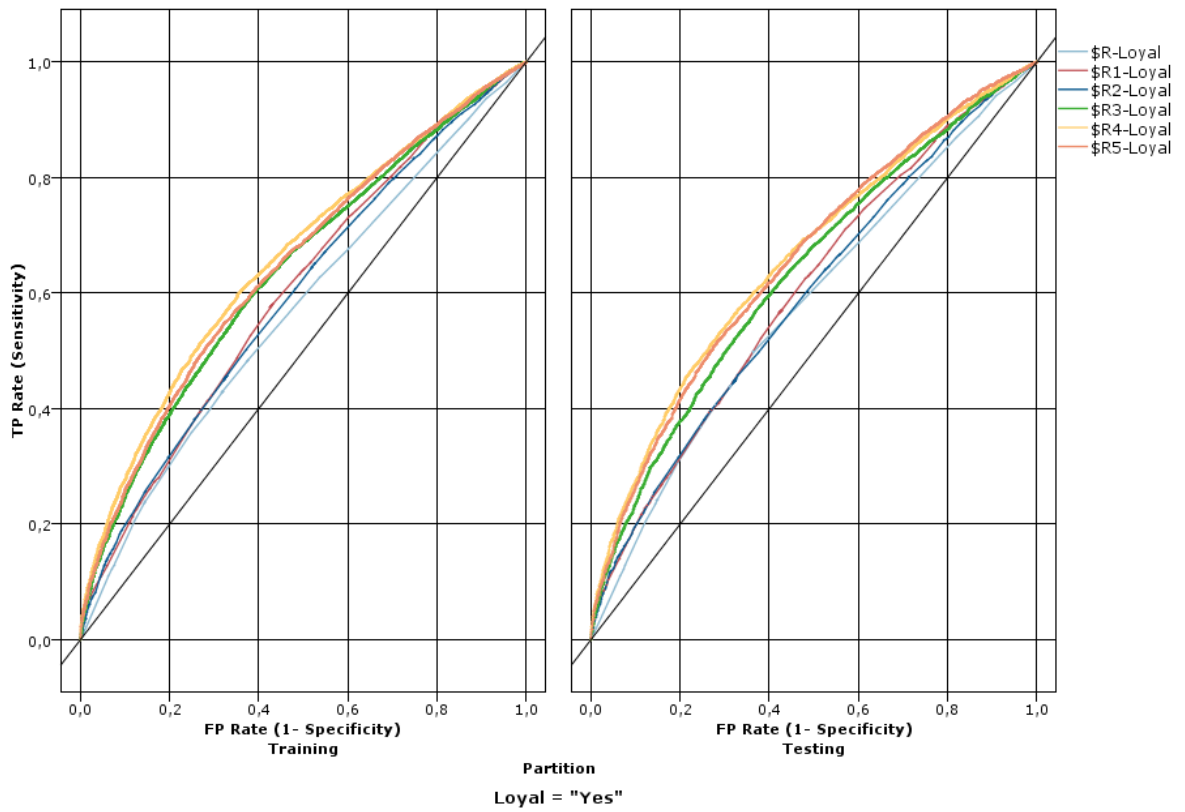
*Figure 2 - ROC for the different CART models*
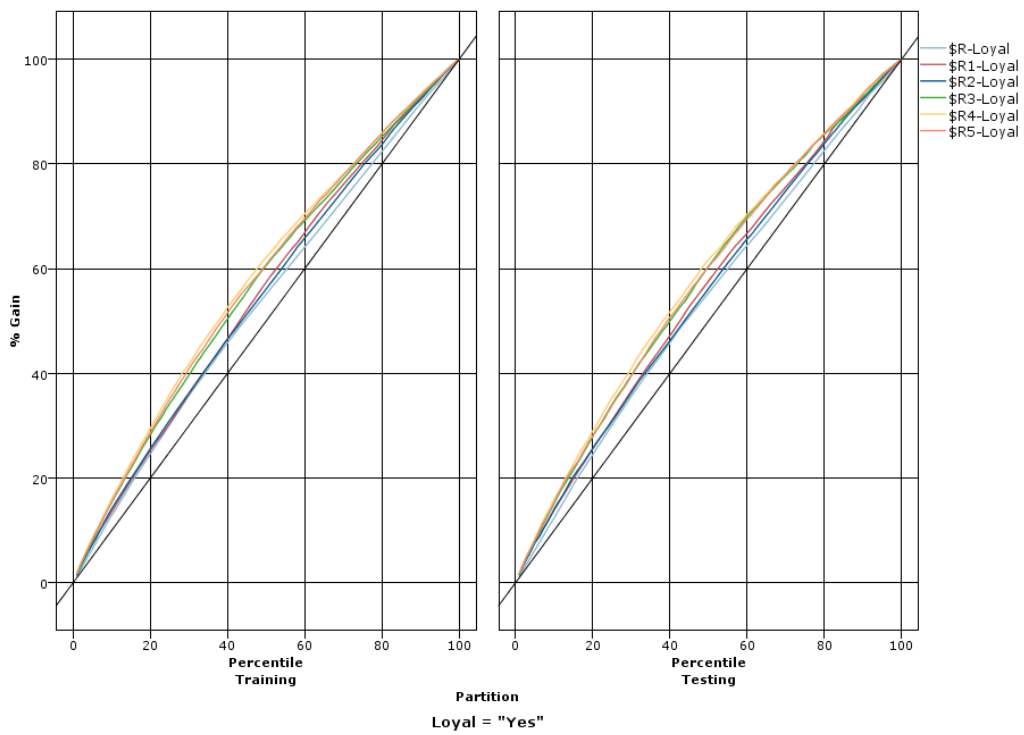


*Figure 3 - Gains chart for the different CART models*
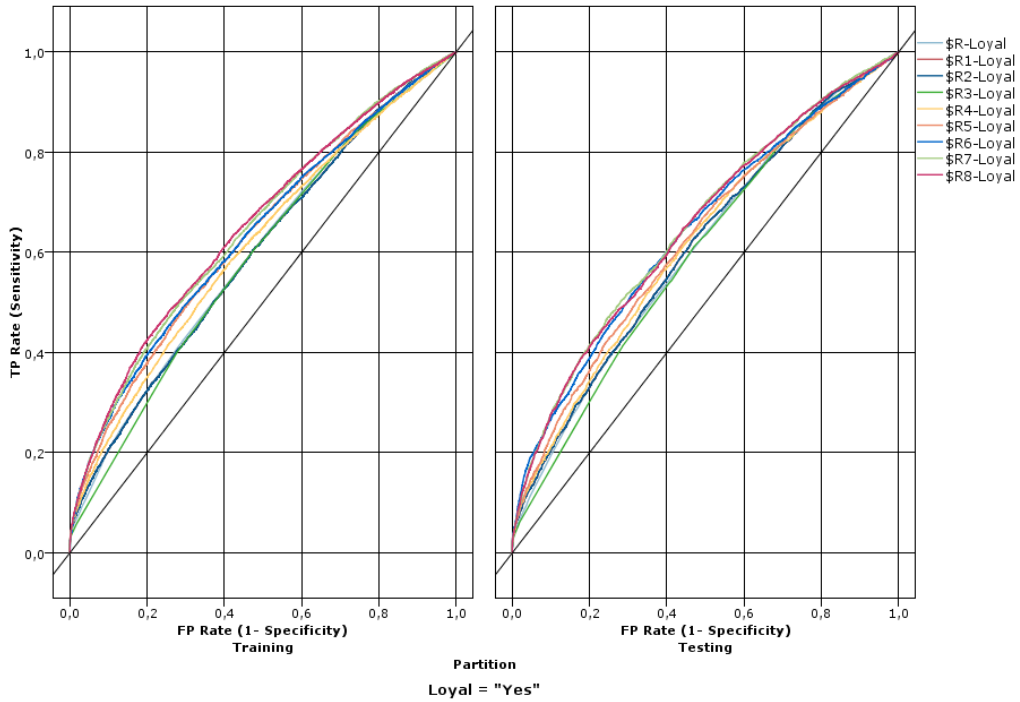
*Figure 4 - ROC for the different CHAID models*
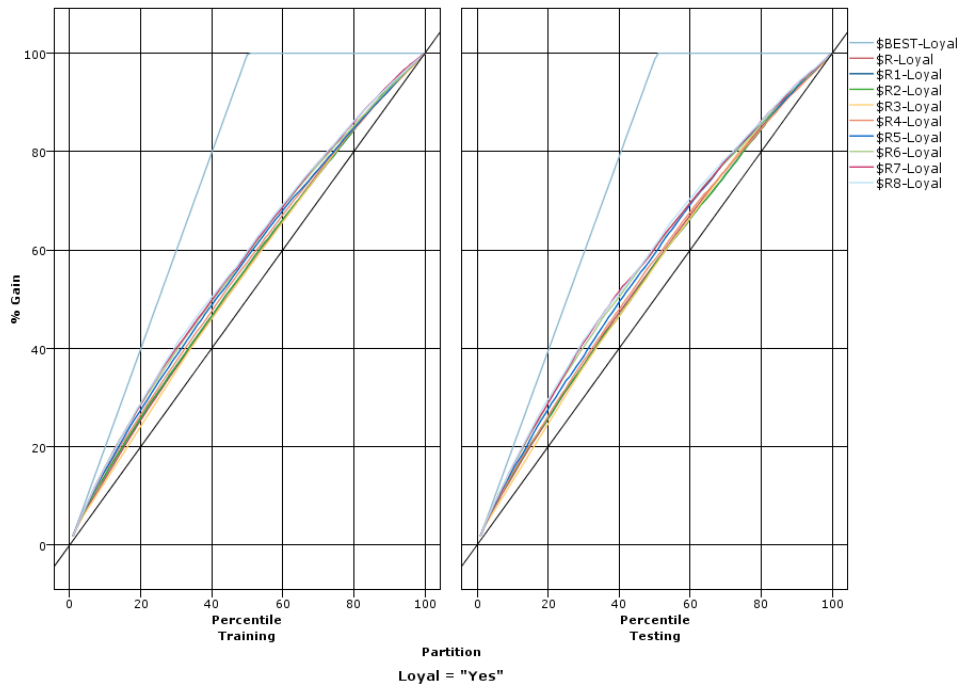


*Figure 5 - Gains chart for the different CHAID models*

The results of the best models of each algorithm are presented in Table 12. It is observed that the best model is the one obtained with the CART algorithm, being slightly better than the CHAID model in all the metrics except for the sensitivity (ability to classify loyal tourists well).

*Table 12- Results of the best predicting models to classify loyalty*

| Evaluation metrics | CART model E (R4) | | CHAID model O (R8) | |
| --- | --- | --- | --- | --- |
| | Train | Test | Train | Test |
| Accuracy | 61.62% | 61.98% | 60.24% | 60.71% |
| Sensibility | 55.27% | 56.24% | 57.75% | 58.13% |
| Specificity | 67.95% | 67.77% | 62.75% | 63.27% |
| AUC | 0.652 | 0.662 | 0.653 | 0.661 |
| Gini | 0.303 | 0.323 | 0.307 | 0.321 |

It should be noted that the predictive capacity of the model is low, overall. However, it allows to identify the profiles of loyal tourists and also of non-loyal tourists, based on the most reliable terminal nodes (the rule set is presented in the annex). Thus, it was identify profiles of tourists that have a strong possibility of being loyal, considering a confinement equal to or greater than 0.8 and a minimum of 10 tourists as support:

- **IF** Nationality region in [ North European Countries Countries of the iberian Peninsula Balkan Countries ] **AND** Nationality in [ Denmark Slovakia ] **AND** Household size - aged 15+ > 1,500 **AND** Occupation in [ Manager of a company (self-employed) Other (self-employed) ] **AND** Age education (11 categories) in [ 15 years 16 years 17 years 19 years 20 years 22 years and older ] **THEN** Loyal = Yes (Support=40; Confidence=0.875);

- **IF** Nationality region in [ North European Countries Countries of the iberian Peninsula Balkan Countries] **AND** Nationality in [ Denmark Greece Latvia Slovakia Slovenia] **AND** Household size - aged 15+ > 4,500 **AND** Occupation in [ Professional (sel-employed lawyer, medical practitioner, ...) Manager of a company (self-employed) Other (self-employed) **THEN** Loyal = Yes (Support=16; Confidence=1.0);

- **IF** Nationality region in [ North European Countries Countries of the iberian Peninsula Balkan Countries] **AND** Nationality in [ Denmark Greece Latvia Slovakia Slovenia] **AND** Occupation in [ Farmer, forester, fisherman (self-employed) Owner of Shop, crafsman (self-employed) General management, director or top management Middle management Civil servant Office clerk Other employee (salesman, nurse, ...) Other (employee) Supervision\ foreman (team manager, ...) Manual worker Unskilled manual worker Other (manual worker) Looking after the home Student (full time) Retired Sneeking a job Other (without

a professional activity)] **AND** Age education (11 categories) in [ Up to 14 years] **THEN** Loyal = Yes (Support=13; Confidence=1.0);

- **IF** Nationality region in [ North European Countries Countries of the iberian Peninsula Balkan Countries] **AND** Nationality in [ Denmark Greece Latvia Slovakia Slovenia] **AND** Occupation in [ General management, director or top management Other employee (salesman, nurse, ...) Other (employee) Unskilled manual worker] **AND** Age education (11 categories) in [ 15 years 16 years 17 years 18 years 19 years 20 years 21 years 22 years and older Still studying No full-time education] **AND** Age <= 63.5 **THEN** Loyal = Yes (Support=14; Confidence=1.0);

- **IF** Nationality region in [ Central Europena Countries Eastern European Countries Baltics] **AND** Nationality in [ Luxembourg] **AND** Phone available in [ Mobile and landline] **THEN** Loyal = Yes (Support=285; Confidence=0.926);

- **IF** Nationality region in [ Central Europena Countries Eastern European Countries Baltics] **AND** Nationality in [ Belgium Cyprus Malta Netherlands] **AND** Occupation in [ Farmer, forester, fisherman (self-employed) Manager of a company (self-employed) General management, director or top management Civil servant Other (employee) Manual worker Looking after the home Student (full time) Retired ] **AND** Age education (11 categories) in [ 20 years ] **AND** Age <= 70,500 **THEN** Loyal = Yes (Support=78; Confidence=0.846);

The same model was used in relation to non-loyal tourist profiles, that is, confidence equal or superior to 0.8 and minimum support of 10:

- **IF** Nationality region in [ North European Countries Countries of the iberian Peninsula Balkan Countries ] **AND** Nationality in [ Greece Latvia Slovenia ] **AND** Occupation in [ Professional (sel-employed lawyer, medical practitioner, ...) Manager of a company (self-employed) Other (self-employed)] **AND** Household size - aged 5+ <= 4,500 **AND** Age education (11 categories) in [16 years 17 years 19 years 20 years 22 years and older ] **AND** Type of community in (Rural area or village) **THEN** Loyal = NO (Support=10; Confidence=0.8);

- **IF** Nationality region in [ North European Countries Countries of the iberian Peninsula Balkan Countries] **AND** Nationality in [ Greece Latvia ] **AND** Occupation in [ Professional (sel-employed lawyer, medical practitioner, ...) Manager of a company (self-employed) Other (self-employed)] **AND** Household

size - aged 5+ <= 4,500 **AND** Age education (11 categories) in [19 years 20 years Still studying ] **THEN** Loyal = NO (Support=15; Confidence=0.933);

- **IF** Nationality region in [ North European Countries Countries of the iberian Peninsula Balkan Countries] **AND** Nationality in [ Denmark Greece Latvia Slovakia Slovenia] **AND** Occupation in [ Farmer, forester, fisherman (self-employed) Owner of Shop, craftsman (self-employed) General management, director or top management Middle management Civil servant Office clerk Other employee (salesman, nurse, ...) Other (employee) Supervision\ foreman (team manager, ...) Manual worker Unskilled manual worker Other (manual worker) Looking after the home Student (full time) Retired Sneeking a job Other (without a professional activity)] **AND** Age education (11 categories) in [ 16 years 17 years 21 years] **AND** Age > 68.5 **THEN** Loyal = No (Support=40; Confidence=0.8);

- **IF** Nationality region in [ North European Countries Countries of the iberian Peninsula Balkan Countries] **AND** Nationality in [ Bulgaria] **AND** Occupation in [Owner of Shop, craftsman (self-employed) General management, director or top management Office clerk Other (employee) Student (full time)] **AND** Gender in [Female] **AND** Household size - aged 15+ <= 4,500 **THEN** Loyal = No (Support=58; Confidence=0.81);

- **IF** Nationality region in [ North European Countries Countries of the iberian Peninsula Balkan Countries] **AND** Nationality in [ Bulgary Croatia Finland Portugal Spain Sweden ] **AND** Occupation in [ Manager of a company (self-employed) Other (self-employed) Other (without a professional activity) ] **AND** Gender in [Female] **AND** Household size - aged 15+ <= 4.5 **AND** Age education (11 categories) in [ 18 years 22 years and older ] **AND** Type of community in [ Small or middle sized town Large town ] **AND** Age > 42.5 **THEN** Loyal = No (Support=58; Confidence=0.81);

- **IF** Nationality region in [ North European Countries Countries of the iberian Peninsula Balkan Countries] **AND** Nationality in [ Bulgary Croatia Finland Portugal Spain Sweden ]**AND** Occupation [ Owner of Shop, crafsman (self-employed) Manager of a company (self-employed) Other (self-employed) General management, director or top management Office clerk Other (employee) Student (full time) Other (without a professional activity) ] **AND** Household size - aged 15+ > 4,500 **THEN** Loyal = No (Support=50; Confidence=0.82);

- **IF** Nationality region in [Central Europena Countries Eastern European Countries Baltics] **AND** Nationality in [ Austria France] **AND** Occupation [ Farmer, forester, fisherman (self-employed) Professional (sel-employed lawyer, medical practitioner, ...) Manager of a company (self-employed) Middle management Other employee (salesman, nurse, ...) Other (employee) Unskilled manual worker Other (without a professional activity) ] **AND** Age education (11 categories) in [ Up to 14 years 16 years 20 years 21 years Still studying ] **AND** Age binned in [ 15 - 24 Years 40 - 54 Years ]**THEN** Loyal = No (Support=72; Confidence=0.806);

- **IF** Nationality region in [Central Europena Countries Eastern European Countries Baltics] **AND** Nationality in [ Austria France] **AND** Occupation [ Professional (sel-employed lawyer, medical practitioner, ...) Manager of a company (self-employed) Middle management Unskilled manual worker Other (without a professional activity) ] **AND** Age education (11 categories) in [ Up to 14 years 16 years 20 years 21 years] **AND** Age binned in [ 25 - 39 Years 55 Years and older] **THEN** Loyal = No (Support=22; Confidence=0.818);

- **IF** Nationality region in [Central Europena Countries Eastern European Countries Baltics] **AND** Nationality in [ Germany Hungary Italy Poland] **AND** Occupation [ Owner of Shop, craftsman (self-employed) Manager of a company (self-employed) Other (self-employed) General management, director or top management Looking after the home Sneeking a job ] **AND** Age > 37,500 **AND** Age binned in [ 15 - 24 Years 25 - 39 Years ] **AND** Type of community in [ Small or middle sized town Large town ] **THEN** Loyal = No (Support=10; Confidence=1.0);

It is still relevant to highlight that nationality is the most important predictor in the classification, following by the nationalty region and thirdly the occupation (Figure 6).
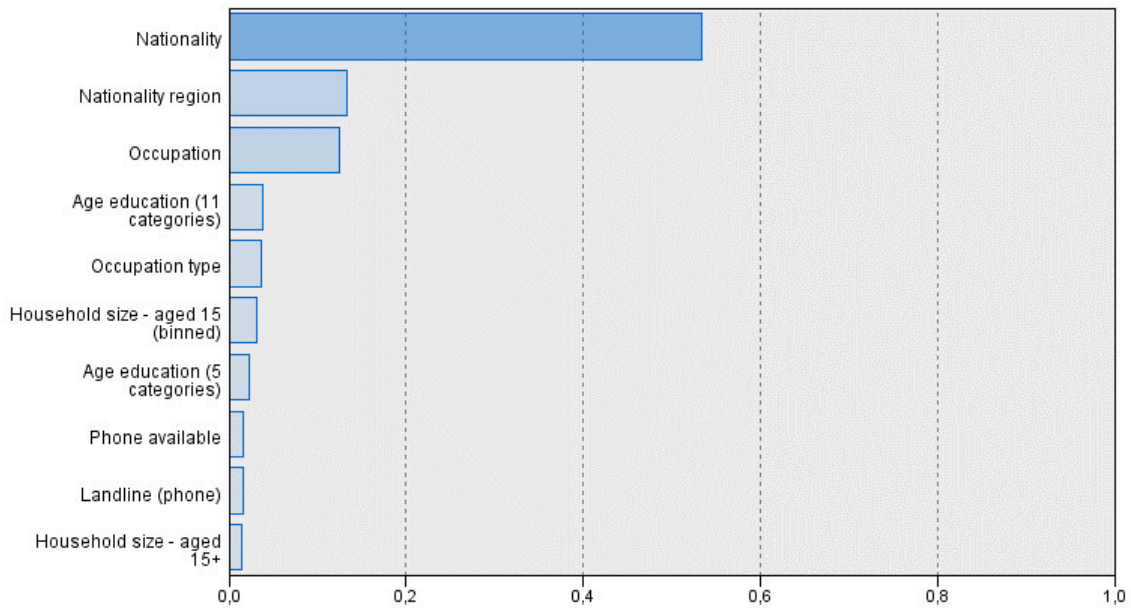
*Figure 6 - Predictor importance of loyalty*

In order to understand how the three predictors influence the loyalty of tourists, figures 7 to 9 are shown. In Figure 7 it is possible to verify that the countries with the highest propensity to have loyal tourists are Luxembourg, Malta, Netherlands, Belgium, Cyprus and Czech Republic (all with more than 60% of loyal tourists). On the other hand, the least likely to be loyal are Austria, Bulgaria, Croatia, France, Sweden, Finland, Estonia, Portugal, Italy, Romania, Slovenia and Spain.
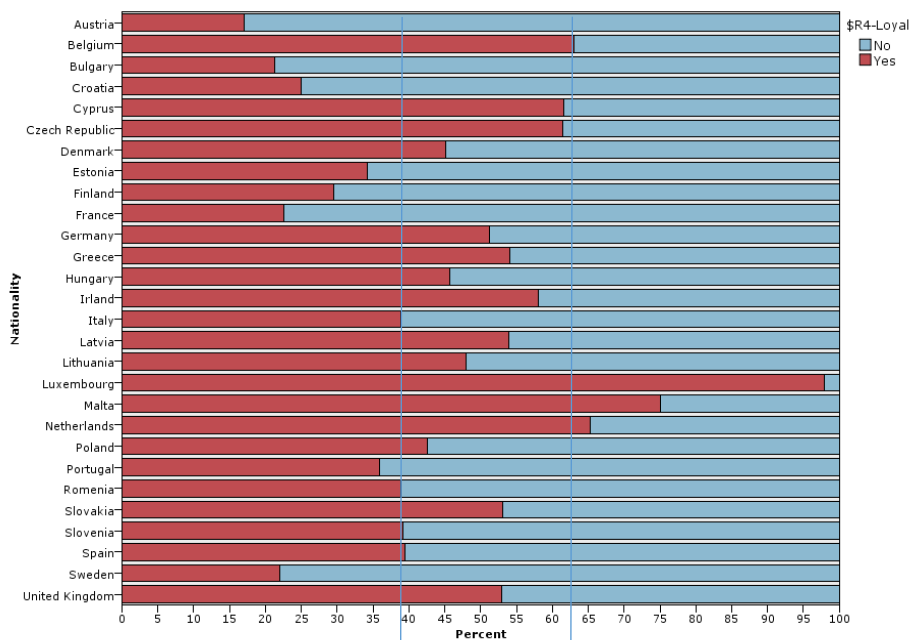


*Figure 7 - Distribution of the predicted loyalty by nationality*

Regarding nationality region the North European countries and the Iberian Coutries are the coutries with more than 60% of non-loyas tourists. On the other hand, the countries

36

of central Europe and eastern European countries have around 50% of loyal tourists. Nevertheless, in terms of the group, the differences are not so significant in relation to the Nationality (by country).
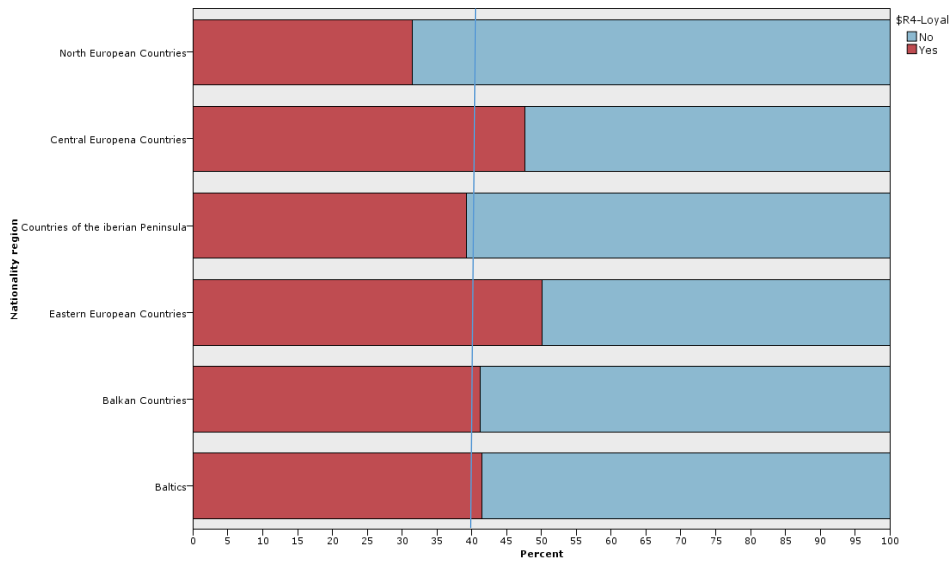


*Figure 8 - Distribution of the predicted loyalty by nationality region*

Reggarding to the occupation, it is possible to verify that the managers of a company as well as the looking after the home, are those that present a greater loyalty (60%). On the other hand, an intermediary manager, a manual worker and supervisors are the ones with the lowest loyalty (only 30%).
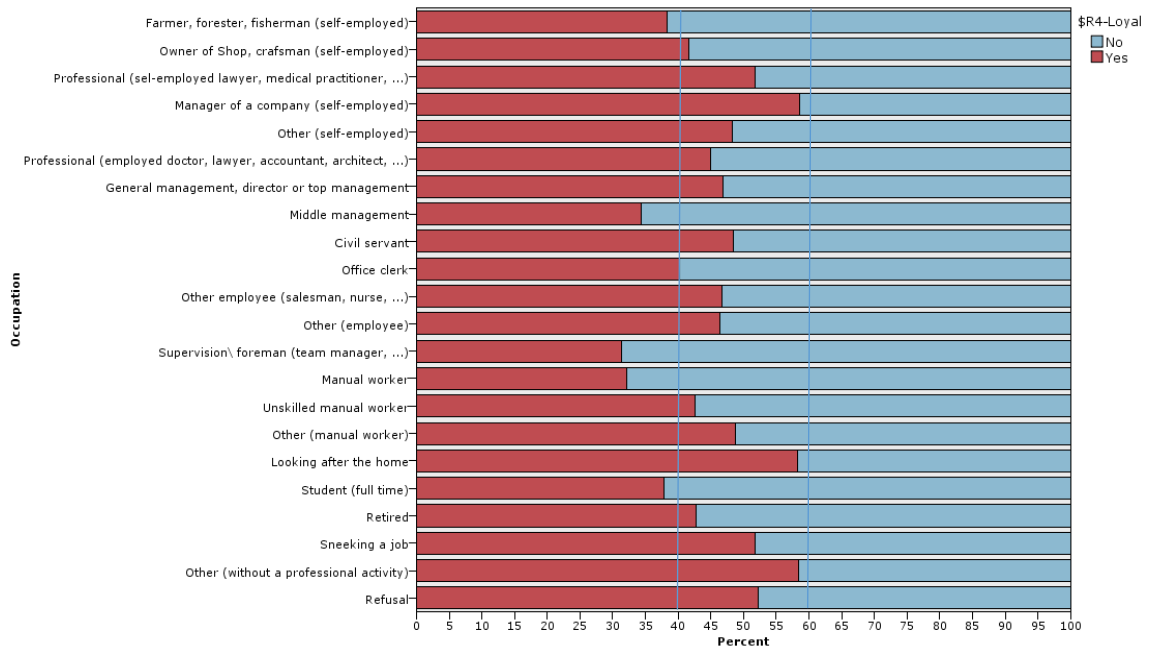


*Figure 9 - Distribution of the predicted loyalty by occupation*

# Chapter 5 – Conclusion and Limitations

## 5.1. Conclusions

Tourism is an important economic activity in most countries around the world; In Europe tourism is the third largest socio-economic activity in the EU (according European Commission) and has an important impact on the economic. This research aims in evaluate the loyalty of European tourists in a perspective of destination country.

To achieve the objectives proposed, and therefore to evaluate the loyalty of the tourists, data analysis and processing techniques were used through the IBM SPSS tools. At the structure level, this project follows a structure based on CRISP-DM being this a successful methodology with proven success cases in business problems.

The results allowed to conclude that there are pairs of countries with higher propensity to be visited in two following years comparing to others that do not tend to be visited in two consecutive years by the same tourist. Furthermore, results showed that there are tourists loyal to their own-country or to a foreign country. In addition, there are non-loyal tourists in Europe. Thus, using classification predictive models we concluded that there are few profiles of loyal and non-loyal tourists, being the nationality, nationality region and the occupation the most important predictors of loyalty.  is an important economic activity in most countries around the world, In Europe Tourism is the third largest socio-economic activity in the EU (according European Commission) and has an important impact on the economic. This research aims in evaluate the loyalty of European tourists in a prespective of destinations country.

## 5.2. Contributions

The knowledge acquired with this study has contribution to the literature of loyalty in tourism and for the professionals of tourism related companies.

First, to the literature, this study contributes on presenting interesting results in macro-level perspective, i.e., in a country loyalty perspective. Moreover, this research comprises a big sample of multinational tourists, for a period of three years, giving support to the results about country loyalty and its determinants; and adopting the CRISP-DM methodology successfully showed that this data mining methodology could be applied in tourism research. Finally, this study alerts researchers that there are many data sets available, like the ones from Eurobarometer, in scientific and non-scientific repositories

and organizations, which should be analysed in order to extract new useful knowledge for the different tourism stakeholders by applying statistics and data mining techniques.

Second, this study allows tourism professionals access to new knowledge to collect information in order to optimize, for instance, their own advertising to reach the desired people. Indeed, knowing the socio-demographic profiles of loyal and non-loyal tourist companies can use the information to implement strategies to retain tourists and to increase the number of tourists by converting non-loyal into loyal tourists.

Concluding, this study contributes to help countries avoiding stagnation and decline of their tourism industry.

## 5.3. Limitations of the Research

This project had some limitations that conditioned the results of the project, namely:

- The questionnaires collected over the 3 years were different and structured differently. This work of data processing would lead to that at the junction of the data either was forced to remove the variables, or it was necessary to make a presumption for the joining of every year.
- Another limitation is the number of valid questionnaires, since it was necessary to prove the loyalty of the tourists and only the questionnaires that had filled the field of nationality, trips and planned trip were considered valid, from the high number of initial responses this was quickly diminished, forcing new data to be created in order to obtain any results.

## 5.4. Further Research

This study allows a range of future work regarding the evaluation of tourist loyalty.

Since satisfaction variables were not considered, it would be interesting to study a cross between satisfaction and loyalty, in order to obtain valid association rules with greater relevance.

On the other hand, with the use of satisfaction variables and socio-demographic variables, another approach would be to create a predictive model where it would be possible to predict whether a tourist is loyal or not and to which country.

Finally, it would be relevant to obtain a greater range of responses, perhaps using a larger number of years of analysis, to try to bridge the limitation of valid cases obtained in the project.

# References

Allsop, Dee, Bryce Bassett, e James Hoskins. "Word-of-Mouth Research: Principles and Applications." *Journal of Advertising Research*, 2007.

Andereck, Kathleen, e Linda Caldwell. "The influence of tourists' characteristics on ratings of information sources for an attraction." 1994.

Anderson, Eugene W. "Customer Satisfaction and Word of Mouth." *Journal of Service Research*, 1998.

Anderson, J. C., e J. A. Narus. "A Model of Distributor Firm and Manufacturer Firm Working Partnerships." *Journal of Marketing*, 1990.

Backman, Sheila, e John Crompton. "The usefulness of selected variables for predicting activity loyalty." 1991.

Baloglu, S., e M. Mangaloglu. "Tourism Destination Images of Turkey, Egypt, Greece, and Italy as Perceived by US-Based Tour Operators and Travel Agents." 2001.

Baloglu, Seyhmus, e Ken McCleary. "A model of destination image formation." *Annals of Tourism Research*, 1999.

Beerli, A., e J.D. Martin. "Factors Influencing Destination Image." *Annals of Tourism resarch*, 2004.

Bieger, Thomas, e Christian Laesser. "Information Sources for Travel Decisions: Toward a Source Process Model." *Journal of Travel Research*, 2004.

Butler, R.W. "The concept of a tourist area cycle of evolution: implications for management of resources." *The canadian Geographer*, Março 1980: 5-12.

Campón, Ana, Helena Alves, e José Hernández. *Loyalty Measurement in Tourism: A Theoretical Reflection.* 2013.

Cejas, Flora. "An Application of the CHAID Algorithm to Study the Environmental Impact of Visitors to the Teide National Park in Tenerife, Spain." 2017.

Cha, Sukbin, Muzaffer Uysal, e Ken Mccleary. "Travel Motivations of Japanese Overseas Travelers: A Factor-Cluster Segmentation Approach." 1995.

Chiou, Jyh-shen, e Cornelia Droge. "Service Quality, Trust, Specific Asset investment, and expertise: Direct and indirect effects in a satisfaction-lotalty framework." 2006.

Chpaman, Pete, et al. "CRISPDM 1.0 - Step-by-step data mining guide. USA: SPSS Inc." 2000.

Christaller, W. "Some considerations of tourism location in Europe: the peripheral regions - under-developed countries - recreation areas." *Regional Science Association Papers*, 1963: 103.

Dann, Graham. "Tourist motivation an appraisal." 1981.

Dick, Alan, e Kunal Basu. "Customer loyalty: Toward an integrated conceptual framework." 1994.

Dimanche , Frederic, e M.E: Havitz. "Consumer Behavior and Tourism: Review and Extension of Four Study Areas." 1994.

Doney, P.M., e J.P. Cannon. "An Examination of the Nature of Trust in Buyer-Seller Relationships." *Journal of Marketing*, 1997.

Ekinci, Yuksel, Sandra Preciado, e Ercan Sirakaya-Turk. "Symbolic consumption of tourism destination brands." *Journal of Business Reseach*, 2013.

Esen, Şaban, e Hande Uyar. "COMPETITIVENESS OF TOURISM AND THE EVALUATION OF TURKEY ACCORDING INTERNATIONAL TOURISM COMPETITIVE CRITERIA." 2010.

Feinauer, Ilse, e Amanda lourens. "The loyalty of the literary reviser: Author, source text, target text or reader?" 2017.

Gazley, A., e L. Watling. "Me, My Tourist-Self, and I: The Symbolic Consumption of Travel." *Jounal of Travel & Tourism Marketing*, 2015.

Gedenk, Karen, e Scott Neslin. "The role of retail promotion in determining future brand loyalty: its effect on purchase event feedback." 1999.

Gefen, D. "E-commerce: the role of familiarity and trust." 2000.

Hongmei, zhang, xiaoxaio Fu, Liping Cai, e Lin Lu. "Destination Image and tourist loyalty : A meta- annalysis." 2013: 10.

Hosany, Sameer, e Girish Prayah. "Patterns of tourists' emotional responses, satisfaction, and intention to recommend." *Journal of Business Research*, 2013.

Hosmer, L.T. "Trust: The Connecting Link between Organizational Theory and Philosophical Ethics." *Academy of management Review*, 1995.

Isoraite, Margarida. "CUSTOMER LOYALTY THEORETICAL ASPECTS." 2016.

Isoraité, margarita. "Costumoer Loyalty Theoretical Aspects." s.d.

Jacoby, Jacob, e Robert Chestnut. "Brand Loyalty, Measurement and Management." 1978: 120.

Kandampully, Jay. "Customer loyalty in the hotel industry: the role of customer satisfaction and image." 2000.

Kim, Yeong Gug, Bo Suh, e Anita Eves. "The relationships between food-related personality traits, satisfaction, and loyalty among visitors attending food events and festivals." 2010.

Költringer, Clemens, e Astrid Dickinger. "nalyzing destination branding and image from online sources: A web content mining approach." *Journal of Business Research*, 2015.

Kozak, Metin, e Mike Rimmington. "Tourist Satisfaction with Mallorca, Spain, as an Off-Season Holiday Destination." 2014.

Kuhl, J., e J. Beckmann. "Action control: From cognition to behavior. Heidelberg, Berlin, New York: Springer." 2012.

Lelper, Nell. *The Framework of tourism.* Australia, 1979.

Lin, Chien, e Beryl Kuo. "escalation of loyalty and the decreasing impact of perceived value and satisfaction over time." 2013.

Lin, Chung-Hsien, Duarte Morais, Deborah Kerstetter, e Jing-Shoung Hou. "Examining the Role of Cognitive and Affective Image in Predicting Choice Across Natural, Developed, and Theme-Park Destinations." *Journal of Travel Research*, 2007.

Losada, N., E. Alén, T. Domínguez, e J. L. Nicolau. "Travel frequency of seniors tourists." 2016.

Lu, C, Y. Hsu, Y Lu, e W. Lin. "Measuring tourist satisfaction by motivation, travel behavior and shoopping behaviour: the case of lake scenic area in taiwan." *International Journal of organizational Innovation*, 215.

Madlberger, Maria, e Sabine Matook. *Theorizing E-Commerce Business Models: On the Impact of Partially and Fully Supported Transaction Phases on Customer Satisfaction and Loyalty.* Vol. 21. Australia: Australasian Journal of Information Systems, 2017.

Marinez, Eva, Carlos Flávian, e Yolanda Polo. "Loyalty to grocery stores in the Spanish market of the 1990s." 2001.

Martinez, Luis, e Aristides Ferreira. *Análise de dados com SPSS.* Portugal, 2008.

Mayer, R.C., J. H. Davis, e F. D. Sshoorman . "An Integration Model of Organizational Trust." *Academy of Management Review*, 1995.

Mazanec, J. "Market Segmentation." 2000.

McIntosh, Alison, e Anna Siggs. "An Exploration of the Experiential Nature of Boutique Accommodation." *Journal of Travel Research*, 2005.

Mcknight, D. H., L. L. Cummings, e N. L. Chervany. "Initial Trust Formation in New Organizational Relationshops." *Academy of Management Review*, 1998.

Mohammed, Shaimaa, e Fairooz Hamdi. "Factors Affecting Tourism Loyalty: A Study on Touristic Destinations in Duhok City." 2017.

Munoz, Julo. "Towards a new definition for "visiting friends and relatives"." 2016.

Nikolova, Milena, e Salah Hassan. "Nation branding effects on retrospective global evaluation of past travel experiences." *Journal of Business Research*, 2013.

Noronha, R. " Review of the Sociological Literature on Tourism." *New York: World Bank*, 1976.

Oliver, Richard L. "Whence Consumer Loyalty?" *Journal of Marketing*, 1999: 33-44.

Partner, Frederick, Robert Bain, e Christopher Bain. "The loyalty effect - the relationship between loyalty and profits." 2000.

Peypoch, N., R. Randriamboarison, F. Rasoamananjara, e B. Solonandrasana. "The length of stay of tourists in Madagascar." 2012.

Phu, Bihn. "Country image, country attachment, country loyalty, and life satisfaction of foreign residents in Vietnam." 2015.

Pike, S., e C. Ryan. "Destination Positioning Analysis Through a Comparison of Cognitive, Affective, and Conative Perceptions." *Journal of Travel Research*, 2004.

Pritchard, Mark, Dennis Howard, e Mark Havitz. *Loyalty Measurement: A Critical Examination.* United Kingdom, 1992.

Pritchard, Mark, e Dennis Howard. "The Loyal Traveler: Examining a Typology of Service Patronage." 1997.

Ragavan, Neethiahnanthan, Hema Subramonian, e Saeed Sharif. "Tourists' Perceptions of Destination Travel Attributes: An Application to International Tourists to Kuala Lumpur." 2014.

Resino, Juan, e Edyta Andrzejak. "The Relational Benefits and the Quality of Services in Tourism." 2017.

Riley, Michael, Outi Niininen, Edith Szivas, e Tonny Willis. "The Case for Process Approaches in Loyalty Research in Tourism." 2001.

Ross, Elizabeth, e Seppo iso-ahola. "Sightseeing Tourists' Motivation and Satisfaction." *Annals of Tourism Research*, 1991.

Sarra, A, S. Di Zio, e M Cappucci. "A quantitative valuation of tourist experience in Lisbon." 2015.

Selin, Steven, Dennis Howard, Edward Udd, e Ted Cable. "An Analysis of Consumer Loyalty to Municipial Recreation Programs." 1988.

Shaw, G, e M Williams. "Knowledge Transfer and Management in Tourism Organisations: an emergin reserach agenda." 2009.

Sweney, Jillian, Geoffrey Soutar, e Tim Mazzarol. "Word of mouth: measuring the power of individual messages." s.d.

Talib, Salim, Mohd Zahari, J.M. Abdul-Rahim, e M. Z. Suhaimi. "International tourists revisit intention: Has it prevalence in United Arab Emirates?" 2015.

Tasci, Asli, e William Gartnet. "Destination Image and Its Functional Relationships." *journal of Travel Research*, 2007.

Vieira, Lisa. "European Tourist Perspective on Destination Satisfaction: A Business Analytics approach." 2016.

Woo, Eunju, Medet Yolal, Fatmagul Cetinel, e Muzaffer Uysal. "A Comparative Study of Motivation across Different Festival Products." 2011.

World Tourism Organization. *Tourism Satellite Account : Recommended Methodological Framework.* Luxembourg, 2001.

Wu, Chih. "Destination loyalty modeling of the global tourism." 2015.

Yee, Beh Yin. "A Study on the determinants of customer loyalty using internet bankin website in malaysia." 2004.

Yee, Beh Yin. "A Study on the Determinants of Customer Loyalty Using Internet Banking Website in Malaysia." 2004.

Yfantidou, Georgia, e Michela Matarazzo. "The future of sustainable tourism in develping countries." 2016.

Zeugner-Roth, Katharina, Adamantios Diamantopoulos, e Vesna Zabkar. "Consumer Ethnocentrism, National Identity, and Consumer Cosmopolitanism as Drivers of Consumer Behavior: A Social Identity Theory Perspective." *Journal of International Marketing*, 2015.

# Annex 1

- Nationality region in [ North European Countries Countries of the iberian Peninsula Balkan Countries ] [ Mode: No ] (5,638)
  - Nationality in [ Denmark Greece Latvia Slovakia Slovenia ] [ Mode: Yes ] (2,336)
    - Occupation in [ Professional (sel-employed lawyer, medical practitioner, ...) Manager of a company (self-employed) Other (self-employed) Professional (employed doctor, lawyer, accountant, architect, ...) ] [ Mode: Yes ] (398)
      - Household size - aged 15+ <= 4,500 [ Mode: Yes ] (382)
        - Nationality in [ Denmark Slovakia ] [ Mode: Yes ] (218)
          - Household size - aged 15+ <= 1,500 [ Mode: Yes ] ⇒ Yes (54; 0,556)
          - Household size - aged 15+ > 1,500 [ Mode: Yes ] (164)
            - Occupation in [ Manager of a company (self-employed) Other (self-employed) ] [ Mode: Yes ] (43)
              - Age education (11 categories) in [ 18 years ] [ Mode: No ] ⇒ No (3; 0,667)
              - Age education (11 categories) in [ 15 years 16 years 17 years 19 years 20 years 22 years and older ] [ Mode: Yes ] ⇒ Yes (40; 0,875)
            - Occupation in [ Professional (sel-employed lawyer, medical practitioner, ...) Professional (employed doctor, lawyer, accountant, architect, ...) ] [ Mode: Yes ] ⇒ Yes (121; 0,661)
        - Nationality in [ Greece Latvia Slovenia ] [ Mode: Yes ] (164)
          - Age education (11 categories) in [ 16 years 17 years 18 years 21 years 22 years and older ] [ Mode: Yes ] (146)
            - Type of community in [ Rural area or vilalge ] [ Mode: No ] (16)
              - Age <= 43 [ Mode: Yes ] ⇒ Yes (6; 0,667)
              - Age > 43 [ Mode: No ] ⇒ No (10; 0,8)
            - Type of community in [ Small or middle sized town Large town ] [ Mode: Yes ] ⇒ Yes (130; 0,631)
          - Age education (11 categories) in [ 19 years 20 years Still studying ] [ Mode: No ] (18)
            - Nationality in [ Greece Latvia ] [ Mode: No ] ⇒ No (15; 0,933)
            - Nationality in [ Slovenia ] [ Mode: Yes ] ⇒ Yes (3; 1,0)
      - Household size - aged 15+ > 4,500 [ Mode: Yes ] ⇒ Yes (16; 1,0)
    - Occupation in [ Farmer, forester, fisherman (self-employed) Owner of Shop, crafsman (self-employed) General management, director or top management Middle management Civil servant Office clerk Other employee (salesman, nurse, ...) Other (employee) Supervision\ foreman (team manager, ...) Manual worker Unskilled manual worker Other (manual worker) Looking after the home Stude...
      - Age education (11 categories) in [ Up to 14 years ] [ Mode: Yes ] ⇒ Yes (13; 1,0)
      - Age education (11 categories) in [ 15 years 16 years 17 years 18 years 19 years 20 years 21 years 22 years and older Still studying No full-time education ] [ Mode: No ] (1,925)
        - Age <= 68,500 [ Mode: Yes ] (1,583)
          - Occupation in [ Owner of Shop, crafsman (self-employed) Civil servant Supervision\ foreman (team manager, ...) Manual worker Other (manual worker) Looking after the home Sneeking a job ] [ Mode: No ] (423)
            - Age binned in [ 25 - 39 Years 40 - 54 Years ] [ Mode: Yes ] (303)
              - Nationality in [ Denmark Greece Slovakia ] [ Mode: Yes ] ⇒ Yes (223; 0,538)
              - Nationality in [ Latvia Slovenia ] [ Mode: No ] (80)
                - Age <= 30,500 [ Mode: Yes ] ⇒ Yes (11; 0,727)
                - Age > 30,500 [ Mode: No ] ⇒ No (69; 0,667)
            - Age binned in [ 15 - 24 Years 55 Years and older ] [ Mode: No ] ⇒ No (120; 0,683)
          - Occupation in [ Farmer, forester, fisherman (self-employed) General management, director or top management Middle management Office clerk Other employee (salesman, nurse, ...) Other (employee) Unskilled manual worker Student (full time) Retired Other (without a professional activity) ] [ Mode: Yes ] (1,160)
            - Age <= 63,500 [ Mode: Yes ] ⇒ Yes (954; 0,523)
            - Age > 63,500 [ Mode: Yes ] (206)
              - Occupation in [ General management, director or top management Other employee (salesman, nurse, ...) Other (employee) Unskilled manual worker ] [ Mode: Yes ] ⇒ Yes (14; 1,0)
              - Occupation in [ Farmer, forester, fisherman (self-employed) Middle management Office clerk Retired Other (without a professional activity) ] [ Mode: Yes ] (192)
                - Age education (11 categories) in [ 15 years 16 years 17 years 18 years 20 years 21 years ] [ Mode: No ] ⇒ No (62; 0,516)
                - Age education (11 categories) in [ 19 years 22 years and older No full-time education ] [ Mode: Yes ] ⇒ Yes (130; 0,654)
        - Age > 68,500 [ Mode: No ] (342)
          - Age education (11 categories) in [ 16 years 17 years 21 years ] [ Mode: No ] ⇒ No (40; 0,8)
          - Age education (11 categories) in [ 15 years 18 years 19 years 20 years 22 years and older No full-time education ] [ Mode: No ] (302)
            - Household size - aged 15+ <= 2,500 [ Mode: No ] ⇒ No (277; 0,567)
            - Household size - aged 15+ > 2,500 [ Mode: Yes ] ⇒ Yes (25; 0,64)
  - Nationality in [ Bulgary Croatia Finland Portugal Spain Sweden ] [ Mode: No ] (3,302)
    - Occupation in [ Farmer, forester, fisherman (self-employed) Professional (sel-employed lawyer, medical practitioner, ...) Professional (employed doctor, lawyer, accountant, architect, ...) Middle management Civil servant Other employee (salesman, nurse, ...) Supervision\ foreman (team manager, ...) Manual worker Unskilled manual worker Other (manual worker) Looking after the home Retire...
      - Age education (11 categories) in [ 15 years 16 years 17 years 18 years 19 years 20 years 21 years 22 years and older ] [ Mode: No ] (2,477)
        - Household size - aged 15+ <= 1,500 [ Mode: No ] (493)
          - Phone available in [ Mobile Only Landline Only ] [ Mode: No ] (205)
            - Occupation in [ Professional (sel-employed lawyer, medical practitioner, ...) Civil servant Unskilled manual worker Retired ] [ Mode: Yes ] (121)
              - Type of community in [ Rural area or vilalge Small or middle sized town ] [ Mode: No ] ⇒ No (80; 0,562)
              - Type of community in [ Large town ] [ Mode: Yes ] ⇒ Yes (41; 0,683)
            - Occupation in [ Farmer, forester, fisherman (self-employed) Professional (employed doctor, lawyer, accountant, architect, ...) Middle management Other employee (salesman, nurse, ...) Manual worker Sneeking a job ] [ Mode: No ] ⇒ No (84; 0,702)
          - Phone available in [ Mobile and landline ] [ Mode: Yes ] (288)
            - Occupation in [ Professional (sel-employed lawyer, medical practitioner, ...) Middle management Manual worker Looking after the home ] [ Mode: Yes ] ⇒ Yes (43; 0,744)
            - Occupation in [ Professional (employed doctor, lawyer, accountant, architect, ...) Civil servant Other employee (salesman, nurse, ...) Supervision\ foreman (team manager, ...) Other (manual worker) Retired Sneeking a job ] [ Mode: No ] (245)
              - Age education (11 categories) in [ 15 years 21 years ] [ Mode: Yes ] ⇒ Yes (29; 0,724)
              - Age education (11 categories) in [ 16 years 17 years 18 years 19 years 20 years 22 years and older ] [ Mode: No ] ⇒ No (216; 0,542)
        - Household size - aged 15+ > 1,500 [ Mode: No ] (1,984)
          - Nationality in [ Croatia Portugal Spain ] [ Mode: No ] (659)
            - Age education (11 categories) in [ 17 years 21 years ] [ Mode: No ] ⇒ No (75; 0,653)
            - Age education (11 categories) in [ 15 years 16 years 18 years 19 years 20 years 22 years and older ] [ Mode: No ] (584)
              - Occupation in [ Professional (sel-employed lawyer, medical practitioner, ...) Professional (employed doctor, lawyer, accountant, architect, ...) Middle management Civil servant Other employee (salesman, nurse, ...) Retired ] [ Mode: No ] (457)
                - Age education (11 categories) in [ 15 years 18 years 19 years 20 years ] [ Mode: Yes ] ⇒ Yes (149; 0,523)
                - Age education (11 categories) in [ 16 years 22 years and older ] [ Mode: No ] ⇒ No (308; 0,588)
              - Occupation in [ Farmer, forester, fisherman (self-employed) Supervision\ foreman (team manager, ...) Manual worker Unskilled manual worker Looking after the home Sneeking a job ] [ Mode: Yes ] ⇒ Yes (127; 0,583)
          - Nationality in [ Bulgary Finland Sweden ] [ Mode: No ] (1,325)
            - Age binned in [ 25 - 39 Years ] [ Mode: No ] (198)
              - Occupation in [ Professional (employed doctor, lawyer, accountant, architect, ...) Civil servant Supervision\ foreman (team manager, ...) Manual worker Looking after the home ] [ Mode: No ] (124)
                - Age education (11 categories) in [ 21 years ] [ Mode: Yes ] ⇒ Yes (4; 0,75)
                - Age education (11 categories) in [ 18 years 19 years 20 years 22 years and older ] [ Mode: No ] ⇒ No (120; 0,742)
              - Occupation in [ Professional (sel-employed lawyer, medical practitioner, ...) Middle management Other employee (salesman, nurse, ...) Unskilled manual worker Sneeking a job ] [ Mode: No ] ⇒ No (74; 0,514)
            - Age binned in [ 15 - 24 Years 40 - 54 Years 55 Years and older ] [ Mode: No ] ⇒ No (1,127; 0,571)
      - Age education (11 categories) in [ Up to 14 years Still studying No full-time education ] [ Mode: Yes ] (123)
        - Occupation in [ Farmer, forester, fisherman (self-employed) Middle management Civil servant ] [ Mode: No ] ⇒ No (5; 1,0)
        - Occupation in [ Professional (sel-employed lawyer, medical practitioner, ...) Professional (employed doctor, lawyer, accountant, architect, ...) Other employee (salesman, nurse, ...) Supervision\ foreman (team manager, ...) Manual worker Unskilled manual worker Looking after the home Retired Sneeking a job ] [ Mode: Yes ] (118)
          - Age <= 80,500 [ Mode: Yes ] ⇒ Yes (115; 0,609)
          - Age > 80,500 [ Mode: No ] ⇒ No (3; 1,0)
    - Occupation in [ Owner of Shop, crafsman (self-employed) Manager of a company (self-employed) Other (self-employed) General management, director or top management Office clerk Other (employee) Student (full time) Other (without a professional activity) ] [ Mode: No ] (702)
      - Household size - aged 15+ <= 4,500 [ Mode: No ] (652)
        - Gender in [ Male ] [ Mode: No ] ⇒ No (301; 0,575)
        - Gender in [ Female ] [ Mode: No ] (351)
          - Occupation in [ Owner of Shop, crafsman (self-employed) General management, director or top management Office clerk Other (employee) Student (full time) ] [ Mode: No ] (295)
            - Nationality in [ Bulgary ] [ Mode: No ] ⇒ No (58; 0,81)
            - Nationality in [ Croatia Finland Portugal Spain Sweden ] [ Mode: No ] (237)
              - Age education (11 categories) in [ 18 years 20 years ] [ Mode: No ] (45)
                - Occupation in [ Owner of Shop, crafsman (self-employed) General management, director or top management Other (employee) ] [ Mode: No ] ⇒ No (9; 0,889)
                - Occupation in [ Office clerk Student (full time) ] [ Mode: Yes ] ⇒ Yes (36; 0,583)
              - Age education (11 categories) in [ 15 years 17 years 19 years 21 years 22 years and older Still studying ] [ Mode: No ] ⇒ No (192; 0,698)
          - Occupation in [ Manager of a company (self-employed) Other (self-employed) Other (without a professional activity) ] [ Mode: No ] (56)
            - Age education (11 categories) in [ 18 years 22 years and older ] [ Mode: No ] (40)
              - Type of community in [ Rural area or vilalge ] [ Mode: Yes ] ⇒ Yes (6; 0,833)
              - Type of community in [ Small or middle sized town Large town ] [ Mode: No ] (34)
                - Age <= 42,500 [ Mode: Yes ] ⇒ Yes (13; 0,538)
                - Age > 42,500 [ Mode: No ] ⇒ No (21; 0,81)
            - Age education (11 categories) in [ Up to 14 years 16 years 19 years 20 years 21 years Still studying ] [ Mode: Yes ] (16)

- Nationality in [ Bulgary Spain ] [ Mode: No ] ⇨ No (8; 0,5)
- Nationality in [ Croatia Finland Sweden ] [ Mode: Yes ] ⇨ Yes (8; 1,0)
- Household size - aged 15+ > 4,500 [ Mode: No ] ⇨ No (50; 0,82)
- Nationality region in [ Central Europena Countries Eastern European Countries Baltics ] [ Mode: Yes ] (10.656)
  - Nationality in [ Austria France ] [ Mode: No ] (2.216)
    - Age education (11 categories) in [ Up to 14 years 16 years 20 years 21 years Still studying ] [ Mode: No ] (523)
      - Occupation in [ Owner of Shop, crafsman (self-employed) Professional (employed doctor, lawyer, accountant, architect, ...) General management, director or top management Civil servant Office clerk Supervision\ foreman (team manager, ...) Manual worker Looking after the home Student (full time) Retired Sneeking a job ] [ Mode: No ] (411)
        - Age education (11 categories) in [ 16 years 20 years ] [ Mode: Yes ] (166)
          - Mobile phone in [ Yes ] [ Mode: Yes ] (160)
            - Occupation in [ Professional (employed doctor, lawyer, accountant, architect, ...) Civil servant Looking after the home ] [ Mode: No ] ⇨ No (26; 0,654)
            - Occupation in [ Owner of Shop, crafsman (self-employed) General management, director or top management Office clerk Supervision\ foreman (team manager, ...) Manual worker Retired Sneeking a job ] [ Mode: Yes ] ⇨ Yes (134; 0,56)
          - Mobile phone in [ No ] [ Mode: Yes ] ⇨ Yes (6; 1,0)
        - Age education (11 categories) in [ Up to 14 years 21 years Still studying ] [ Mode: No ] (245)
          - Occupation in [ Owner of Shop, crafsman (self-employed) Supervision\ foreman (team manager, ...) Looking after the home ] [ Mode: Yes ] ⇨ Yes (6; 0,833)
          - Occupation in [ Professional (employed doctor, lawyer, accountant, architect, ...) General management, director or top management Civil servant Office clerk Manual worker Student (full time) Retired Sneeking a job ] [ Mode: No ] (239)
            - Household size - aged 15+ <= 1,500 [ Mode: Yes ] ⇨ Yes (44; 0,545)
            - Household size - aged 15+ > 1,500 [ Mode: No ] ⇨ No (195; 0,621)
      - Occupation in [ Farmer, forester, fisherman (self-employed lawyer, medical practitioner, ...) Manager of a company (self-employed) Middle management Other employee (salesman, nurse, ...) Other (employee) Unskilled manual worker Other (without a professional activity) ] [ Mode: No ] (112)
        - Age binned in [ 15 - 24 Years 40 - 54 Years ] [ Mode: No ] ⇨ No (72; 0,806)
        - Age binned in [ 25 - 39 Years 55 and older ] [ Mode: No ] (40)
          - Occupation in [ Other employee (salesman, nurse, ...) ] [ Mode: Yes ] ⇨ Yes (15; 0,733)
          - Occupation in [ Professional (sel-employed lawyer, medical practitioner, ...) Manager of a company (self-employed) Middle management Unskilled manual worker Other (without a professional activity) ] [ Mode: No ] (25)
            - Age education (11 categories) in [ Up to 14 years 16 years 20 years 21 years ] [ Mode: No ] ⇨ No (22; 0,818)
            - Age education (11 categories) in [ Still studying ] [ Mode: Yes ] ⇨ Yes (3; 1,0)
    - Age education (11 categories) in [ 15 years 17 years 18 years 19 years 22 years and older No full-time education ] [ Mode: No ] (1.693)
      - Occupation in [ Farmer, forester, fisherman (self-employed) Professional (sel-employed lawyer, medical practitioner, ...) Civil servant Other employee (salesman, nurse, ...) Supervision\ foreman (team manager, ...) Manual worker Sneeking a job ] [ Mode: No ] (437)
        - Type of community in [ Rural area or vilalge ] [ Mode: No ] ⇨ No (205; 0,644)
        - Type of community in [ Small or middle sized town Large town ] [ Mode: No ] (232)
          - Age education (5 categories) in [ 16-19 years ] [ Mode: Yes ] ⇨ Yes (53; 0,604)
          - Age education (5 categories) in [ Up to 15 years 20 Years and older ] [ Mode: No ] ⇨ No (179; 0,57)
      - Occupation in [ Owner of Shop, crafsman (self-employed) Manager of a company (self-employed) Other (self-employed) Professional (employed doctor, lawyer, accountant, architect, ...) General management, director or top management Middle management Office clerk Other (employee) Unskilled manual worker Looking after the home Student (full time) Retired Other (without a professio
  - Nationality in [ Belgium Cyprus Czech Republic Estonia Germany Hungary Irland Italy Lithuania Luxembourg Malta Netherlands Poland Romenia United Kingdom ] [ Mode: Yes ] (8.440)
    - Nationality in [ Luxembourg ] [ Mode: Yes ] (296)
      - Phone available in [ Mobile Only Landline Only ] [ Mode: Yes ] (11)
        - Age education (11 categories) in [ 18 years 22 years and older ] [ Mode: Yes ] ⇨ Yes (4; 1,0)
        - Age education (11 categories) in [ 19 years Still studying ] [ Mode: No ] ⇨ No (7; 0,714)
      - Phone available in [ Mobile and landline ] [ Mode: Yes ] ⇨ Yes (285; 0,926)
    - Nationality in [ Belgium Cyprus Czech Republic Estonia Germany Hungary Irland Italy Lithuania Malta Netherlands Poland Romenia United Kingdom ] [ Mode: Yes ] (8.144)
      - Nationality in [ Belgium Cyprus Malta Netherlands ] [ Mode: Yes ] (1.995)
        - Occupation in [ Farmer, forester, fisherman (self-employed) Manager of a company (self-employed) General management, director or top management Civil servant Other (employee) Manual worker Looking after the home Student (full time) Retired ] [ Mode: Yes ] (1.131)
          - Age education (11 categories) in [ 20 years ] [ Mode: Yes ] (90)
            - Age <= 70,500 [ Mode: Yes ] ⇨ Yes (78; 0,846)
            - Age > 70,500 [ Mode: No ] (12)
              - Type of community in [ Rural area or vilalge Small or middle sized town ] [ Mode: No ] ⇨ No (9; 1,0)
              - Type of community in [ Large town ] [ Mode: Yes ] ⇨ Yes (3; 1,0)
          - Age education (11 categories) in [ Up to 14 years 15 years 16 years 17 years 18 years 19 years 21 years 22 years and older Still studying ] [ Mode: Yes ] ⇨ Yes (1.041; 0,632)
        - Occupation in [ Owner of Shop, crafsman (self-employed) Professional (sel-employed lawyer, medical practitioner, ...) Other (self-employed) Professional (employed doctor, lawyer, accountant, architect, ...) Middle management Office clerk Other employee (salesman, nurse, ...) Supervision\ foreman (team manager, ...) Unskilled manual worker Sneeking a job Other (without a professio
          - Nationality in [ Malta ] [ Mode: Yes ] ⇨ Yes (50; 0,74)
          - Nationality in [ Belgium Cyprus Netherlands ] [ Mode: Yes ] (814)
            - Household size - aged 15+ <= 1,500 [ Mode: Yes ] (134)
              - Age education (11 categories) in [ Up to 14 years 19 years 20 years 21 years 22 years and older ] [ Mode: Yes ] ⇨ Yes (98; 0,735)
              - Age education (11 categories) in [ 16 years 17 years 18 years Still studying ] [ Mode: No ] ⇨ No (36; 0,639)
            - Household size - aged 15+ > 1,500 [ Mode: Yes ] (680)
              - Age education (11 categories) in [ Up to 14 years 16 years 18 years 19 years 20 years 21 years 22 years and older ] [ Mode: No ] (639)
                - Occupation in [ Professional (sel-employed lawyer, medical practitioner, ...) Professional (employed doctor, lawyer, accountant, architect, ...) Middle management Office clerk Other employee (salesman, nurse, ...) ] [ Mode: Yes ] ⇨ Yes (529; 0,52)
                - Occupation in [ Owner of Shop, crafsman (self-employed) Other (self-employed) Supervision\ foreman (team manager, ...) Unskilled manual worker Sneeking a job Other (without a professional activity) ] [ Mode: No ] ⇨ No (110; 0,609)
              - Age education (11 categories) in [ 15 years 17 years Still studying ] [ Mode: Yes ] ⇨ Yes (41; 0,732)
      - Nationality in [ Czech Republic Estonia Germany Hungary Irland Italy Lithuania Poland Romenia United Kingdom ] [ Mode: Yes ] (6.149)
        - Occupation in [ Owner of Shop, crafsman (self-employed) Manager of a company (self-employed) Other (self-employed) General management, director or top management Looking after the home Sneeking a job ] [ Mode: Yes ] (965)
          - Age binned in [ 15 - 24 Years 25 - 39 Years ] [ Mode: Yes ] (261)
            - Nationality in [ Germany Hungary Italy Poland ] [ Mode: No ] (99)
              - Age <= 37,500 [ Mode: No ] ⇨ No (86; 0,57)
              - Age > 37,500 [ Mode: No ] (13)
                - Type of community in [ Rural area or vilalge ] [ Mode: Yes ] ⇨ Yes (3; 0,667)
                - Type of community in [ Small or middle sized town Large town ] [ Mode: No ] ⇨ No (10; 1,0)
            - Nationality in [ Czech Republic Estonia Irland Lithuania Romenia United Kingdom ] [ Mode: Yes ] ⇨ Yes (162; 0,568)
          - Age binned in [ 40 - 54 Years 55 Years and older ] [ Mode: Yes ] (704)
            - Age education (11 categories) in [ 18 years 21 years Still studying ] [ Mode: Yes ] (179)
              - Landline (phone) in [ Yes ] [ Mode: No ] (142)
                - Nationality in [ Czech Republic Hungary Irland Italy Poland ] [ Mode: No ] ⇨ No (82; 0,622)
                - Nationality in [ Estonia Germany Lithuania Romenia United Kingdom ] [ Mode: Yes ] ⇨ Yes (60; 0,583)
              - Landline (phone) in [ No ] [ Mode: Yes ] ⇨ Yes (37; 0,784)
            - Age education (11 categories) in [ Up to 14 years 15 years 16 years 17 years 19 years 20 years 22 years and older No full-time education ] [ Mode: Yes ] ⇨ Yes (525; 0,644)
        - Occupation in [ Farmer, forester, fisherman (self-employed) Professional (sel-employed lawyer, medical practitioner, ...) Professional (employed doctor, lawyer, accountant, architect, ...) Middle management Civil servant Office clerk Other employee (salesman, nurse, ...) Other (employee) Supervision\ foreman (team manager, ...) Manual worker Unskilled manual worker Other (manual w
          - Nationality in [ Estonia Germany Hungary Poland Romenia ] [ Mode: No ] (2.327)
            - Age education (11 categories) in [ 15 years 16 years 19 years 20 years ] [ Mode: No ] ⇨ No (486; 0,591)
            - Age education (11 categories) in [ Up to 14 years 17 years 18 years 21 years 22 years and older Still studying No full-time education ] [ Mode: No ] (1.841)
              - Occupation in [ Farmer, forester, fisherman (self-employed) Professional (employed doctor, lawyer, accountant, architect, ...) Civil servant Office clerk Supervision\ foreman (team manager, ...) Unskilled manual worker Student (full time) Retired ] [ Mode: Yes ] (1.267)
                - Household size - aged 15+ <= 2,500 [ Mode: No ] ⇨ No (917; 0,513)
                - Household size - aged 15+ > 2,500 [ Mode: Yes ] ⇨ Yes (350; 0,58)
              - Occupation in [ Professional (sel-employed lawyer, medical practitioner, ...) Middle management Other employee (salesman, nurse, ...) Other (employee) Manual worker Other (manual worker) Other (without a professional activity) ] [ Mode: No ] ⇨ No (574; 0,561)
          - Nationality in [ Czech Republic Irland Italy Lithuania United Kingdom ] [ Mode: Yes ] (2.857)
            - Household size - aged 15+ <= 3,500 [ Mode: Yes ] (2.446)
              - Occupation in [ Professional (employed doctor, lawyer, accountant, architect, ...) Other employee (salesman, nurse, ...) Other (employee) Supervision\ foreman (team manager, ...) Manual worker Student (full time) ] [ Mode: Yes ] ⇨ Yes (753; 0,502)
              - Occupation in [ Farmer, forester, fisherman (self-employed) Professional (sel-employed lawyer, medical practitioner, ...) Middle management Civil servant Office clerk Unskilled manual worker Other (manual worker) Retired Other (without a professional activity) ] [ Mode: Yes ] (1.693)
                - Age <= 86,500 [ Mode: Yes ] ⇨ Yes (1.678; 0,558)
                - Age > 86,500 [ Mode: No ] ⇨ No (15; 0,733)
            - Household size - aged 15+ > 3,500 [ Mode: No ] (411)
              - Age <= 72,500 [ Mode: No ] ⇨ No (404; 0,557)
              - Age > 72,500 [ Mode: Yes ] ⇨ Yes (7; 1,0)