MASTERARBEIT / MASTER'S THESIS

Titel der Masterarbeit / Title of the Master‘s Thesis

"Influence of sociodemographic factors, fundraising methods and past behaviour in the lifetime value of donors. A case study of SOS Children's Villages"

verfasst von / submitted by

Diana Carolina Platarrueda Acuña

angestrebter akademischer Grad / in partial fulfilment of the requirements for the degree of

Master of Science (MSc)

Wien, 2020 / Vienna 2020

Studienkennzahl lt. Studienblatt / degree programme code as it appears on the student record sheet:	UA 066 914
Studienrichtung lt. Studienblatt / degree programme as it appears on the student record sheet:	Masterstudium Internationale Betriebswirtschaft
Betreut von / Supervisor:	ao. Univ.-Prof. i.R. Mag. Dr. Johann Brazda

Acknowledgements

This thesis is dedicated to my parents Sara Acuña and Darío Platarrueda and all my family who taught me good values, prepared me for big things in life and supported me in every step along the way, even in the distance.

I would like to thank my colleagues in SOS Children's Villages, the biggest family in the world, for making this possible, letting me work together with you. Special thanks to the Strategy and Analysis team for your patience and support during these study years.

Thank you to the University of Vienna, my supervisor and his team for facilitating this work, my teachers, and the people I met during my studies here, some of whom are dear friends now.

Agradecimientos

Esta tesis está dedicada a mis padres Sara Acuña y Darío Platarrueda y a toda mi familia, quienes me enseñaron buenos valores, me prepararon para cosas grandes en la vida y me apoyaron en cada paso en este camino, incluso en la distancia.

Agradezco a mis colegas de Aldeas Infantiles SOS, la familia más grande del mundo, por hacer esto posible permitiendome trabajar con ustedes. Agradecimiento especial a mi equipo de Estrategia y Análisis por su paciencia y apoyo durante estos años de estudio.

También agradezco a la Universidad de Viena, a mi supervisor y su equipo por facilitar este trabajo, a mis profesores y a las personas que conocí durante mis estudios, algunos de los cuáles se conviertieron en apreciados amigos.

Abstract

As non-profit organisations face increasing competition in the sector, the need to target and build a long-term relationship with the best possible donors becomes essential for the effective use of limited resources. Research shows a wide range of factors that influence giving behaviour and donor lifetime value. This study addresses the questions of which factors influence the value of donors as well as the common traits of high-value donors. A statistical analysis using CHAID was conducted on a sample of 292,478 existing donors of SOS Children's Villages, an international NGO in the field of childcare, to identify which factors influence high-value donors. Results indicate that behavioural factors are highly influential to donor value, whereas fundraising methods and socio-demographics are of relative importance for specific donor categories. The main recommendation is to conduct further research using different approaches in the field of giving behaviour and its effects on donor value.

Abstract

Gemeinnützige Organisationen sind zunehmendem Wettbewerb ausgesetzt. Um eine effektive Nutzung begrenzter Ressourcen zu ermöglichen, ist Planung und Aufbau langfristiger Beziehungen zu den bestmöglichen Spendern essenziell. Die vorangehende Forschung auf diesem Feld beleuchtet ein breites Spektrum von Faktoren, die das Spendenverhalten und den Spendergesamtwert beeinflussen. Die vorliegende Studie beschäftigt sich mit den Faktoren, die den Spenderwert beeinflussen sowie die gemeinsamen Merkmale des Spenders von hohem Wert. Eine statistische Analyse wurde durch die Verwendung von dem CHAID-Algorithmus an einer Stichprobe von 292.478 bestehenden Spendern für SOS-Kinderdörfer, einer internationalen Nichtregierungsorganisation im Bereich der Kinderbetreuung, durchgeführt. Ziel war die Bestimmung der Faktoren, die den Wert von Spender von hohem Wert beeinflussen. Die Ergebnisse deuten darauf hin, dass verhaltensbezogene Faktoren den Spenderwert stark beeinflussen, während Fundraising-Methoden sowie Soziodemographie für bestimmte Spenderkategorien von relativer Bedeutung sind. Weitere Forschungsarbeiten mit unterschiedlichen Ansätzen auf dem Gebiet des Geberverhaltens und seiner Auswirkungen auf den Spenderwert sind Teil der Empfehlungen.

Table of Content

List of Tables **5**

List of Figures **7**

List of Abbreviations **8**

1 Introduction **9**

1.1 Status and trends of Individual Giving fundraising 10

1.1.1 People donating money, as a percentage of the total population 10

1.1.2 Statistics in giving from individual donors 10

1.1.3 Trends in Individual Giving Fundraising in NGOs 11

1.2 About SOS Children’s Villages 12

1.2.1 Individual Giving fundraising in SOS CV 13

1.2.2 Motivation and scope 16

1.3 Research questions 17

2 Literature review and theoretical framework **18**

2.1 Donor lifetime val ue: how much are donors worth 19

2.1.1 Donor value measures 20

2.1.2 Summary of relevant literature addressing DLTV 23

2.2 Giving behaviour of individuals 25

2.3 Donor segmentation 30

2.4 Theoretical framework 32

2.4.1 Factors influencing DLTV 32

2.4.2 Hypotheses 32

3 Methodology **34**

3.1 Selection of the statistical method: CHAID 34

3.2 Data and variables 35

3.3 Donor lifetime value 37

3.4 Socio-demographic variables 38

3.5 Behaviour variables 39

3.5.1 Donor types, categories and subgroups 39

3.5.2 Behaviour variables 42

3.6 Fundraising methods variables 42

3.7 Application of the CHAID algorithm 44

4 Results **47**
4.1 CHAID model for Sponsors 48
4.1.1 Results for Sponsors 48
4.1.2 Evaluation of the model for Sponsors 52
4.2 CHAID model for Committed Givers 53
4.2.1 Results for Committed Givers 53
4.2.2 Evaluation of the model for Committed Givers 56
4.3 CHAID model for Mid and Major donors 58
4.3.1 Results for Mid and Major donors 59
4.3.2 Evaluation of the model for Mid and Major donors 61
4.4 CHAID model for Single Givers 63
4.4.1 Results for Single Givers 63
4.4.2 Evaluation of the model for Single Givers 68
5 Discussion **69**
5.1 Methodological considerations 69
5.2 Results and recommendations 70
6 Conclusion **73**
References **75**
Appendix **82**

List of Tables

Table 1. *SOS CVI Revenue* 13

Table 2. *Donor lifetime value measures.* 21

Table 3**.** *Summary of studies regarding donor lifetime value analysis* 23

Table 4. *Annual Average Donor Lifetime Value (AADLTV) Classes* 37

Table 5. *Socio-demographic variables and categories* 38

Table 6. *Donor Category and Donor Type description* 39

Table 7. *Donor Category and Donor Type summary* 40

Table 8. *Complete donor records by Donor Category* 40

Table 9. *Subgroups by Donor Category and Donor Type* 41

Table 10. *Categories for Donations in the first year and Lifespan* 42

Table 11. *Fundraising methods: channels and products* 43

Table 12. *Independent variables used for the CHAID model, according to Donor Category* 44

Table 13. *Parameters of the chaid_control function in R* 45

Table 14. *Distribution of "High" and "Low-mid" AADLTV classes per donor category* 47

Table 15. *Variables and categories for Sponsors* 48

Table 16. *High-value nodes from the model for Sponsors* 51

Table 17. *Confusion matrix and accuracy of models for Sponsors* 52

Table 18. *Variables and categories for Committed Givers* 53

Table 19. *High-value nodes from the model for Committed Givers* 54

Table 20. *Confusion matrix and accuracy of models for Committed Givers* 56

Table 21. *Variables and categories for Mid and Major Donors* 58

Table 22 *High-value nodes from the model for Mid and Major donors* 59

Table 23. *Confusion matrix and accuracy of models for Mid and Major donors* 61

Table 24. *Variables and categories for Single Givers* 63

Table 25. *CHAID control arguments modified for Single Givers model* 64

Table 26. *High-value nodes from the model for Single Givers* 66

Table 27. *Confusion matrix and accuracy of models for Single Givers* 68

List of Figures

Figure 1. A basic model of giving behaviour. Adapted from: Noor, et al., 2015; Chang, 2007; Lee & Chang, 2008 27

Figure 2. Model of Individual Charity Giving Behaviour. Adapted from (Sargeant, 1999, S. 218)........ 27

Figure 3. Theoretical framework 32

Figure 4. Analysis model framework 36

Figure 5. CHAID Model for Sponsors........ 50

Figure 6. CHAID Model for Committed Givers........ 57

Figure 7. CHAID Model for Mid- and Major Donors........ 62

Figure 8. CHAID Model for Single Givers 65

List of Abbreviations

AADLTV: Average Annual Donor Lifetime Value

CAF: Charities Aid Foundation

CHAID: CHi-squared Automated Interaction

CRM: Customer Relationship Management

DLTV: Donor Lifetime Value

HNWI: High Net Worth Individual

NGOs: Non-governmental Organisations

SOS CV: SOS Children’s Villages

1 Introduction

#The later income stream, Individual Giving fundraising, refers to the monetary contributions that private individuals provide to non-profit organisations and has grown in relevance as a source of support for non-profit organisations (Srnka, Grohs, & Eckler, 2003, S. 70). With the growing importance of donations by individuals and looking for to improve efficiency of fundraising in this income stream, international NGOs have turned towards a relational approach in their fundraising, focused on the setting of regular donations and building long-term profitable relationships with their donors, supported by the structuration of a professional Individual Giving fundraising strategy that considers marketing concepts (Magson, 1999; Schlegelmilch, Diamantopoulos, & Love, 1997).

The case for SOS Children's Villages (SOS CV) has been similar. SOS CV is an NGO with international presence "working to protect and care for children who have lost parental care, or who stand at risk of losing it" (SOS Children's Villages International, 2017). SOS CV has established fundraising strategies for many of the countries in which it operates; however, it finds challenges and opportunities to gain and retain individual supporters and understand their value. This Master Thesis will analyse the factors and trends that indicate how much SOS CV supporters are worth to the organisation, employing a statistical study of the existing database of donors in a selected country in which SOS CV operates. The results aim to provide strategic input to the Individual Giving fundraising strategy of the organisation and to contribute to the area of research of marketing for individual donors in NGOs.

Chapter 1 introduces the context and motivation of this Master Thesis and defines the research questions. Chapter 2 covers the literature review that explores the trends in individual giving fundraising that lead to the proposal of a theoretical framework. Chapter 3 describes the methodology for the empirical study. Chapter 4 shows the results and Chapter 5 discusses the findings. Chapter 6 will cover the conclusions of the study.

1.1 Status and trends of Individual Giving fundraising

1.1.1 People donating money, as a percentage of the total population

The 2018 publication of the Charities Aid Foundation (CAF) World Giving Index, a study about trends in generosity across the globe, indicates that private individuals are increasingly engaged in charitable behaviour in the last years. According to CAF, 29.1% of individuals donated money in 2017 globally, representing a downward trend for the second consecutive year (Charities Aid Foundation, 2018).

CAF (2018) analyses the percentage of people donating money according to gender, age, region and economic status of the country. Results in 2017 show that, at a global level, this percentage is slightly higher for men than for women (0.2 percentage points). In terms of age, although until 2016 it was consistently proven that likelihood to donate money increases with age at global level, in 2017 the "report shows that those aged 50+ are now no more likely to donate than those aged 30-49 years, and both of these age groups are now significantly less likely to report donating money than they were previously. The proportion of younger people (aged 15-29 years) donating money across the globe remains stable at around a quarter" (Charities Aid Foundation, 2018). For developed countries, the percentage in 2017 is higher and increased versus 2016 (from 40 to 42%), compared to developing countries (from 25 to 24%). In terms of regions, Oceania has the highest percentage (70%) of people who donated money in 2017, followed by Europe (37%).

1.1.2 Statistics in giving from individual donors

In terms of the volume of donations, there is no global report that estimates the total amount of giving by individuals. Statistics from the International Fundraising Leadership Forum, a group of large INGOs, shows that the total volume of donations from individuals to 15 INGOs in this group in 2018 amount to € 7.2 billion globally, representing 35% of the total income. Individual giving increased since 2014 but decreased by 4.3% in 2018 compared to 2017 (International Fundraising Leadership Forum, 2019). Thi decline is a concern for NGOs, especially after public news such as the Haiti OXFAM sexual misconduct scandal (BBC News, 2018) has proven to affect the public image and credibility of the NGO sector, resulting in loss of supporters for various NGOs across countries.

The Global Trends in Giving Report (Nonprofit Tech for Good, 2018) researches trends of how donors give to charity in an annual survey. The following are some of the main results and statistic found about individual donor giving in 2018:

- 45% of donors are part of in a regular giving program
- 54% of donors prefer to pay using a credit or debit card online
- 41% of donors donated to crowdfunding campaigns from other individuals, which causes that 16% of them donate less to organisations
- 31% give to organisations outside of their residency country
- 18% gave using Facebook fundraising tools, and 88% of them are likely to do it in the future

The report also presents the results categorised by gender, generation, ideology, religion and donor size, as factors that influence the trends, and find differences for specific donor groups. New peer to peer donation tools such as Crowdfunding and Facebook, which offer donors alternative ways to give than donating through NGOs, are growing in importance and may potentially affect results for NGOs in the future.

1.1.3 Trends in Individual Giving Fundraising in NGOs

With the development of the third sector, NGOs face increasing competition with each other and with other organizations in the sector, for people's support and contributions (de Vries, Reis, & Moscato, 2015; Durango-Cohen, Torres, & Durango-Cohen, 2013), as well as scrutiny from the public and the press, who put pressure to ensure that the operations and allocation of funds by NGOs are transparent and effective.

This environment leads NGOs to look for ways to analyse their association with the public and their supporters, often based on the sophistication of their marketing activities, to achieve their long-term support from them. Trends like the exploration of what leads donor behaviour (Schlegelmilch, Diamantopoulos, & Love, 1997; Sargeant, 1999; Sargeant, Ford, & West, 2006; Chang, 2007; Lee & Chang, 2008; Srnka, Grohs, & Eckler, 2003; Snipes & Oswald, 2010; Noor, et al., 2015), donor clustering and segmentation (Shelley & Polonsky, 2002; Srnka, Grohs, & Eckler, 2003; Durango-Cohen, Torres, & Durango-Cohen, 2013; Rupp, Kern, & Helmig, 2014; de Vries, Reis, & Moscato, 2015) and the application of a relational approach to donors through Lifetime Value (Magson, 1999; Sargeant, 2001; Aldrich, 2000; Masters, 2000; Bennett, 2006; Abdolvand, Albadvi, & Koosha, 2014) rose in academic relevance and in application in the third sector, mainly in NGOs doing Individual Giving fundraising, with the aim of finding and targeting the more likely and high quality supporters for their causes and organizations.

In the face of high competition for support from individuals, the focus is on the relationship with supporters. NGOs have over the years, built a database of their supporters, usually using Customer Relationship Management systems to record the donors and their contributions. Sargeant & McKenzie (1999) claim that "many non-profits are sitting on a veritable goldmine of information – a rich source of data about their donors; their individual characteristics and behaviours". Nowadays, in the era of technology and data, donor data is a powerful source to learn from their current supporters, maintain them, and give input to acquire new ones. International non-profits are finding ways to catch-up in the application of data to concepts of Individual Giving fundraising, such as donor behaviour, segmentation and lifetime value (i.e. the total contribution of donors during their tie as supporters for the organisations). The use of digital technologies and donor databases facilitates this process and optimises the targeting of "better" donors and their engagement, ultimately improving the outcomes of the fundraising activities.

1.2 About SOS Children's Villages

SOS Children's Villages (SOS CV) is a global federation that works to protect and care for the group of children who have lost parental care, or who stand at risk of losing it, with the vision that "every child belongs to a family and grows up with love, respect and security" (SOS Children's Villages International, 2017). To achieve the vision, SOS CV is set to work with stakeholder like communities, partners and states, seeking to ensure the fulfilment of the rights of all children. Present in over 134 countries and territories, SOS CV provides quality alternative care and services to prevent family breakdown, as well as other services to safeguard children, advocate for their rights, education and protection in case of emergencies (SOS Children's Villages International, 2017).

In 2018 SOS CV raised € 1.26 billion globally from various revenue streams, as shown in Table 1. private individuals is the largest source, representing 50% of total income, followed by Governmental and Institutional funding (36%), Corporations and Foundations (7%), Emergency appeals and other income (7%) (SOS Children's Villages International, 2019).

Table 1. *SOS CVI Revenue*

Revenue	**2015**	**2016**	**2017**	**2018**
Individuals	591,345	619,758	647,951	628,485
Corporations / foundations	90,589	87,442	93,221	83,583
Public funds	358,892	393,812	427,358	456,782
Emergency appeals	17,224	5,652	5,081	2,334
Other	88,516	100,144	100,390	89,844
Total Revenue	**1,146,566**	**1,206,808**	**1,274,000**	**1,261,028**

Source: SOS Children's Villages International (2019; 2018; 2017)

After years of an increasing trend, global income from SOS CV decreased in 2018, with a slight decline of 1%, driven by a reduction across various income streams, mainly Individual Giving.

1.2.1 Individual Giving fundraising in SOS CV

In 2018, € 629 million were raised by SOS CV from private individuals globally, in three categories (SOS Children's Villages International, 2019, S. 63):

- Single donations: individuals provide a one-off gift to the organisation. This category reached € 302 million in 2018 (24% of total income)
- Sponsorships/Committed Giving: individuals commit to donating funds regularly with a defined frequency, linked to a specific support type. These donations raised € 299 million in 2018 (24% of total income).
- Major Donors: Large donations that are made by High Net Worth Individuals. SOS CV raised € 26 million in this stream in 2018

After several years of continued growth in this income segment, the last years has seen a slowdown in growth, reaching a decline in 2018, driven by a decrease in European markets. Europe is the largest region in terms of income volume and contributes the most significant amounts of international funds to different regions around the globe (SOS Children's Villages

International, 2019); therefore, its decline creates concerns about the sustainability of the organisation's programmes in the long term.

1.2.1.1 Fundraising channels in SOS CV

SOS CV uses different channels to engage and acquire new donors across different countries, according to the local possibilities and realities. From the various methods for raising funds used by the organisation, the following are most representative:

- Face to face: an agent that represents SOS CV approaches and requests donations directly from potential donors in the streets, public or private sites (e.g. parks, metro stations, shopping malls, fairs, conferences)
- Telemarketing: an SOS CVI representative makes telephone calls to potential donors to request for donations
- Digital: Donors fill-in an online formulary in a website with their details and sign-up for donations online
- Direct mail: Prospect donors receive a letter to their address with an ask and instructions for how to donate
- DRTV: SOS CV uses televisión broadcast to present viewers with a request to contribute via a transfer, text message, or another payment channel

1.2.1.2 Fundraising products in SOS CV

The products relate to the value propositions and commitment that donors select to support the organisation. SOS CV works with the following types of products to receive donations[1]:

- Sponsorships: the donation is linked and restricted to a specific purpose. The donors are called *sponsors*. Within this category, there are subtypes:
    - National/International Child sponsorship: the donor sponsors and receives information about a specific child (in the country or another country). This product usually generates the highest emotional connection with the sponsor and therefore, a long relationship with the organisation.

[1] Regular donations can be paid monthly, quarterly, bi-annually or yearly.

- National/International Village sponsorship: the donation is allocated to cover costs in a specific villages location worldwide
- Other Sponsorship: the organisation allocates the donation to a specific purpose, different from child or village sponsorships, which vary from country to country

- Committed Giving: the committed givers provide regular donations linked to the work of the organisation, however, not restricted to a specific use
- Single donations: one-off contributions to the organisation, without a commitment to regular donations. Single giving is a common way in which prospects make their first donations and can switch to regular donors later on
- Mid-major donations: single or regular donations with a significantly higher amount than average. Typically, donors with mid-major donations receive differentiated and customised treatment by the organisation.

1.2.1.3 Challenges and opportunities for Individual Giving in SOS CV

The results of a global internal survey made with fundraiser staff globally and internal communications with a global Fundraising expert provided the insights to understand the current challenges that SOS CV faces in the area of Individual Giving Fundraising, that delimit the problem and scope of the present study.

The results of the internal strategy survey (SOS Children's Villages International, 2019) indicate that in countries with established fundraising strategies, the segment of individuals, in particular, regular donations, is one of the priorities. The main barriers identified are weaknesses in the strategy, increasing competition and lack of resources for fundraising activities. When asked about what is the target audience for individual donors is, around 40% (a significant percentage) did not know. For those developed countries where this was known, Matures, Baby Boomers and Generation X stood as the more critical generations to target, with differences between countries. Respondent's comments in this regard reveal that most donors tend to be female and that targeting younger audiences below 25 to 35 years old are avoided as does not bring the desired engagement and results, while they regard mature audiences as more committed and the highest contributors.

The interview with internal global Fundraising strategy expert (W. van Rijn, personal communication, 2019) highlighted some key issues and points of interest for the organisation.

Firstly, the saturation of fundraising activities, increasing competition and media attention to scandals in the third sector represent a risk to SOS CV, mainly in developed European countries as income in this region stagnates or decreases in some of the key countries for the organisation. On the other hand, regions with developing countries (Asia, Latin America and Eastern Europe) face a fast development of the fundraising sector, accompanied by increasing competitiveness of incoming organisation starting operations in these, now attractive, markets. Secondly, there are gaps in the knowledge about existing and potential donors, and what influences their support and engagement to the organisation. Although this is identified as an essential issue to tackle, there is no systematic analysis to identify the characteristics and motivation of donors or to perform their segmentation that supports the design of strategies to retain and attract supporters; the level of knowledge and analyses vary from country to country. Factors such as gender, age, income, and motivations are assumed to influence donations, however, without a precise indication as to how. One of the opportunities identified is the use of donor lifetime value analysis, as input for managers and fundraisers about which should be the organisational fundraising focus that can bring sustainability of income to achieve the strategic goals with efficiency in the use of limited available resources. Lastly, fundraising experts acknowledge that there is vast available data in the existing donor databases across the organisation that is ready to be further explored and could provide valuable insights and inputs to the definition of the fundraising strategies.

1.2.2 Motivation and scope

The results of the survey and the expert interview reveal a gap in the systematic knowledge of donors and the extensive use of existing information available in the organisation's databases. Accordingly, the scope of the present study is to carry over a pilot analysis, using existing historical donor data available in a selected country[2], that covers the factors influencing the lifetime value of donors during their relationship with SOS CV, as well as insights regarding which factors lead to high-value donors.

[2] Given confidentiality agreement the country of the study will not be mentioned

This research is complementary to other related studies being carried out by SOS CV at an international level in parallel, that aim at a holistic overview of aspects related to individuals. For example, one research is studying the psychographic motivations and profiles of donors who support SOS CV; another project aims at benchmarking Individual Giving performance indicators for different channels and donor groups.

This study aims to broaden the academic literature in the field of the lifetime value of individuals to charity. A big part of the relevant authors produced the literature related to donor lifetime value analysis for non-profits and analytical case studies between the 1990s and 2000s (Magson, 1999; Sargeant & McKenzie, 1999; Aldrich, 2000; Sargeant, 2001; Bennett, 2006) and there have been limited academic papers with practical cases in the field in recent years. Furthermore, while the studies related to giving behaviour widely consider donor characteristics (Sargeant, 1999; Chang, 2007; Lee & Chang, 2008; Srnka, Grohs, & Eckler, 2003; Schlegelmilch, Diamantopoulos, & Love, 1997), the case studies related to donor lifetime value were limited to a differentiation by fundraising channel, without a thorough consideration of how the characteristics of donors or the combination of these with specific fundraising methods and other factors play a role in the lifetime support of donors. The present study will contribute to recent research in the lifetime value, with a more comprehensive approach, by carrying-out a current practical study that involves a broader range of factors influencing donor value and the effect of their interactions in the identification of factors that lead to high-value donors.

1.3 Research questions

The opportunities identified led to the formulation of the following research questions, in agreement with the representatives of SOS CV.

Question 1. Which factors influence the value of donors of SOS CV?

Question 2. What are the shared factors that lead to donors with a higher value for SOS CV?

2 Literature review and theoretical framework

With the development of a fast-paced third sector, non-profit organisations face "increasing competition from one another and other NFP[3] organisations for people's time, money and efforts" (de Vries, Reis, & Moscato, 2015, S. 2). Additionally, non-profits experience a higher level of communication and demands from individuals and public opinion. As response, non-profit organizations start to follow the steps of for-profit companies, adopting targeting strategies in order to find their most likely donor base (de Vries, Reis, & Moscato, 2015, S. 2) and methods to increase the reach, sophistication and efficiency of marketing activities, with the aim of effectively targeting the individuals that are more likely to engage in giving (Durango-Cohen, Torres, & Durango-Cohen, 2013, S. 172).

There are specific thematic trends in which the third sector, supported by academia, have focused in order to address the issues associated with fundraising from individuals. The first topic that has become increasingly relevant one is the study of giving behaviour: why and how people give to charity and what are the determinants of charitable giving, to support the understanding of the potential to receive donations from individuals (Chang, 2007; Sargeant, Ford, & West, 2006; Snipes & Oswald, 2010; Lee & Chang, 2008; Noor, et al., 2015; Yörük, 2009). The second trend identified in literature ad practice, closely linked to the determinants of giving, is the practice of donor segmentation, to analyse the groups of potential and existing donors (Srnka, Grohs, & Eckler, 2003; de Vries, Reis, & Moscato, 2015; Durango-Cohen, Torres, & Durango-Cohen, 2013). The third trend relates to non-profits looking for a relational approach to individuals in their fundraising, focused on the setting of regular donations and building profitable long-term relationships with their donors (Schlegelmilch, Diamantopoulos, & Love, 1997; Magson, 1999). This trend includes the incorporation of the concept of donor value, that analyses how much donors, through their contributions, will be worth for an organization during their lifetime as donors (Magson, 1999; Bennett, 2006; Aldrich, 2000; Masters, 2000; Sargeant, 2001; Sargeant & McKenzie, 1999).

[3] Non for profit

The advancements in the study of these thematic areas are complementary with the development of fundraising methods to approach and communicate with donors, as well as the creation of products and value propositions. The combination of these elements results in the development of Individual Giving strategies based on a relational approach that allows the effective allocation of limited resources, by defining the right targeting and segmentation of donors, using the available fundraising methods, channels and products (Schlegelmilch, Diamantopoulos, & Love, 1997). The following sections will present a more in-depth review of the concept of lifetime value in association with the other trends in Individual Giving Fundraising, as the basis for the theoretical framework to carry out the study.

2.1 Donor lifetime value: how much are donors worth

The field of marketing, and subsequently fundraising, is transitioning from a transactional approach to supporters to a relational approach (Sargeant, 2001). In this setting, fundraisers recognise that "if treated with respect, donors will want to give again, and fundraisers are therefore content to live with somewhat lower rates of return in the early stages of a relationship" (Sargeant, 2001, S. 26)

In a relational approach to donors, non-profit organisations aim to build profitable relationships with their supporters (Magson, 1999, S. 11), considering the costs of obtaining and retaining them, compared to the expected gains from their contributions. Commonly, organisations break even on the investment made to acquire new donors after several months or years and the costs of recruiting a new donor are higher than the cost of retaining the existing ones (Bennett, 2006; Masters, 2000). Fundraisers recognise that it is not fundamental to break even in the first communication and that the return over investments come in the longer term (Sargeant, 2001, S. 26).

Organisations face the need to predict the duration of the relationship with donors and the potential contributions that they will make in their lifetime as supporters, i.e. their value (Magson, 1999; Bennett, 2006; Masters, 2000; Aldrich, 2000). Ultimately, "the successful quantification in monetary terms of the value of a donor to a voluntary organisation can be a valuable aid to the subsequent development of fundraising strategy" (Sargeant, 2001, S. 25). That value will include all the contributions made during their relationship, including gifts in the testament or last will of donors, an indication the strength of the relationship built with the donor during the period of support (Bennett, 2006).

Donor value measures arose to address these issues. According to Magson, “the creation of value measures is born out of an organisational necessity to establish existing and ongoing financial objectives and benchmarks” (Magson, 1999, S. 12). Value indicators are meant to aid charities to take financial decisions about investments made to acquire and retain supporters (Aldrich, 2000; Masters, 2000; Magson, 1999). Paired with donor segmentation, the analysis of lifetime value can also support the selection of the right prospective donors (Masters, 2000; Sargeant & McKenzie, 1999). Value measures provide an understanding of which donors can become longer-term supporters and cover questions such as: how long donors will continue to support the organization, what is their attrition rate, what is their forecasted “worth” (value of their monetary contributions over time), and what are the effective strategies to recruit them and maximise the benefits for the organization (Magson, 1999, S. 12).

The most common value measure is Donor Lifetime Value (DLTV). DLTV arises as an adaptation of the commercial sector’s Customer Lifetime Value, which has been widely researched and applied in the areas of relationship marketing in the for-profit sector (Abdolvand, Albadvi, & Koosha, 2014). There are different ways to define, measure and interpret DLTV, that depend on the variables considered and the elements available for the analysis in each particular case. In a general definition, DLTV refers to the monetary contributions that a donor generates during his/her lifetime as a supporter (Sargeant, 2001). Depending on the availability of data, organisations can define the elements to use in the calculation of DLTV.

2.1.1 Donor value measures

Different definitions of DLTV exist in literature and application. There are three key considerations that organizations have to choose from when analysing DLTV: deciding on gross or net DLTV, using historic (present) or projected future measures (Sargeant, 2001; Magson, 1999) and applying DLTV at an individual donor level or for specific segments or the donor base (Sargeant, 2001; Magson, 1999; Sargeant & McKenzie, 1999; Masters, 2000; Bennett, 2006). A fourth consideration is how to address and standardise time in the equations (Aldrich, 2000; Magson, 1999).

Table 2 summarises a series of donor lifetime value measures and the definitions by Magson (1999), that include elements for consideration in DLTV analysis.

Table 2. *Donor lifetime value measures.*

Measure	Definition	Notes
Gross donation value	The total gross income generated from a donor during the relationship	Includes all contributions made in different forms to the organisation
Annual average donation value	Average gross income generated from a donor per annum	
Net lifetime value	The net income generated from a donor during the relationship, i.e. Gross donation value minus costs incurred to acquire and retain the donor	Includes the cost of acquisition (the most representative one) and ongoing costs of communications and loyalty
Discounted lifetime value	The net LTV discounted to allow for the net present value of money	

Source: Donors: how much do they give in a lifetime? *(Magson, 1999)*

2.1.1.1 Gross or net DLTV

As defined in Table 2, net DLTV takes into consideration revenue and costs of acquisition and ongoing communications, and it is, therefore, more comprehensive than gross DLTV, as it approximates better to the real value of donors, discounting the cost in which organisations.

Gross DLTV considers only the revenue coming from donors, with the limitation of not considering the costs incurred to produce that revenue. Gross DLTV becomes a suitable alternative to analyse donor value, in the absence of structured data on costs in the CRM or other systems.

2.1.1.2 Historic or predictive DLTV

One important distinction that organisations have to make is between using historical or projected future value, depending on the question that the authors want to address and the availability and structure of data.

Historic DLTV is based on an analysis of the database to date, that responds to how much donors were worth in the past (Sargeant, 2001). These measures aim to get an estimation, as accurate as possible, for particular donors at different levels. The formula to calculate past DLTV, according to Magson (1999) is:

Historic net DLTV = Past Donations-Past cost of acquiring and ongoing communication

This approach makes some assumptions that limit its applicability or interpretation, such as that all donors have a single recruitment point, a single end-date (limits itself to the current time) and does not consider a discount rate of the money over an extended period (Magson, 1999).

Future projected DLTV responds to how much is likely that a donor is worth in the future and that is according to various authors the core of the analysis of DLTV, in the perspective of a marketer (Sargeant, 2001; Magson, 1999). In practice, practitioners base the creation of future-oriented measures on past performance so far, as this is the data that exists (Magson, 1999). A commonly used formula for future predictive DLTV is as follows (Sargeant, 2001; Sargeant & McKenzie, 1999; Aldrich, 2000):

$Predictive\ net\ LTV = \sum_{i=1}^{n} Ci(1+d)^{-i}$ where,

$c = net\ contribution\ (revenue\ minus\ cost) from\ each\ year's\ fundraising\ activities$

$d = discount\ rate$

$i = the\ expected\ duration\ of\ the\ relationship\ (in\ years)$

This formula "indicates that it is necessary to calculate the likely future contribution by a donor to each year's fundraising activities, discount these future contributions, and then add them all together" (Sargeant, 2001, S. 29).

2.1.1.3 Individual or segmented DLTV

Another important consideration is whether to measure DLTV for individual donors or segments of the database (Sargeant, 2001). DLTV is commonly calculated at the level of individuals but examined in specific segments of the database, which allows for an analysis of specific donor groups and fundraising channels or appeals. The segmentation principles used differ according to the study and the availability of data from organisations in the donor bases and the financial systems.

2.1.1.4 Time considerations and annual average DLTV

One of the issues of the measurement and comparison of DLTV between different groups of donors is the consideration of the time factor. For example, if donors have been longer in the database, then the likelihood that their DLTV is higher than that of a donor added only recently. A similar case occurs, with the difference in time of introduction between the fundraising channels, with some having several years and some newer ones such as Digital, which might result in an overrepresentation of value for the older channels. To address this issue, practitioners and authors have tried to standardise the value measures to a specific timeframe (Aldrich, 2000). In practice, DLTV is measured for 1, 2 or N years, only with those donors or channels that have had the chance to behave during each specific period.

Aldrich (2000) proposed the introduction of time into the equation with the calculation of the Annual Average Donor Lifetime Value (AADLTV), a measure that considers the total Lifetime value normalised to the period that the donor or channel exists in the database. In his paper, Aldrich (2000) found that by applying this equation, the results contradict the apparent knowledge of lifetime value, to reflect the reality of the database. For example, using the regular DLTV measure, the results found that Cold Mail was the best recruitment channel, influenced by being one of the oldest channels existing in the database, and therefore having longer records and donors. However, using the AADLTV, Cold Mail was shown to be the worst channel for donor recruitment, proving the effect of not considering time in the measures.

2.1.2 Summary of relevant literature addressing DLTV

To support the basis for the theoretical framework, Table 3 presents a summary of studies in the area of donor lifetime value, its main features and the methodology used.

Table 3**.** *Summary of studies regarding donor lifetime value analysis*

Name and author(s)	Study overview	Methodology
How much are new donors worth? Making donor recruitment investment decisions	An investigation of future expected lifetime value of a non-profit's (Sight Savers International)	Future expected lifetime value analysis of the database of donors in the charity, using a time-based approach, based on annual average lifetime value

Name and author(s)	Study overview	Methodology
based on lifetime value analysis (Aldrich, 2000)	donors by recruitment source.	Segmentation made by recruitment source, comparison by descriptive statistics
Predicting the Lifetime Durations of Donors to Charities (Bennett, 2006)	Empirical study of the factors that encouraged donors to a specific charity in the UK to continue their relationship with the charity.	Bennet studied the length of stay of donors, in association with two psychometric traits and four “exchange” variables (value and frequency of donations, number of charities supported and means of donation) and the strength of enjoyment about being thanked for a gift. The author used a survey and regression analysis to measure the relationship between variables.
Donors: how much do they give in a lifetime? (Magson, 1999)	Paper that explores a variety of measures related to donor value that derived from available information, and the practicalities and issues.	Magson takes a theoretical approach that proposes measures around donor value, their calculations and issues. Through cases, the author uses descriptive statistic to compare the measures according to channel and year of recruitment.
Deciding to recruit only donors with high lifetime values (Masters, 2000)	Case study that demonstrates how the decisions of targeting donors are made, based on a small group of non-profit organisations	Masters develops a case study with responses from a small group of charities that ask and receive cash donations to understand donor targeting, using descriptive statistics for the analysis of the groups of “best” a “poor-value” donors.
The lifetime value of donors: Gaining insight through CHAID (Sargeant & McKenzie, 1999)	Review of the use of the analytical tool CHAID can be used to inform and aid the development of a fundraising strategy.	The authors provide an explanation and examples of the application of CHAID analysis tool used in lifetime value analysis of a database profiled according to age,

Name and author(s)	Study overview	Methodology
		gender, location, recruitment media and campaign, and value of the first donation.
Using Donor Lifetime Value to Inform Fundraising Strategy (Sargeant, 2001)	Review of the contribution of donor lifetime value for fundraising	Theoretical review and definition of a conceptual framework for the calculation of donor lifetime value, accompanied by an example which uses descriptive statistics to compare lifetime value from different recruitment channels.

The literature review shows mostly a theoretical approach, and only a few application cases, which focus on analysis of donor value for different communication channels, however, limited in scope of how factors related to the donors play a role and interact in donor lifetime value and how non-profit organizations can approach this issue with an analysis of their donor bases.

Some influencing factors of donor value arise from this review, such as personal characteristics (age, gender, income level), previous behaviour (e.g. amount of the first donation), and fundraising channels. The following sections expand on these factors, in connection with the trends of individual giving fundraising.

2.2 Giving behaviour of individuals

The issue of why and how individuals decide to help others has been addressed widely by different areas of study, including the "economic, clinical psychology, social psychology, anthropology and sociology literature" (Sargeant, 1999, S. 216), and more recently, literature incorporates contributions from the study of marketing (Sargeant, 1999, S. 216-217). Researches in various disciplines have worked in defining charitable behaviour and proposing and testing its possible determinants, ranging from demographic and socioeconomic, to psychological and social (Lee & Chang, 2008, p. 13). Charitable giving behaviour of individuals is determined by altruist actions taken by them and defined and measured in different ways.

The CAF World Giving Index, a study about trends in giving across the globe, for example, measures three aspects of giving: helping a stranger, volunteering time and giving money to charitable organisations (Charities Aid Foundation, 2018). In a more specific definition, Lee & Chang (2008, S. 1173) propose that “giving to charities comes in two major forms: time and money”, referring in specific to volunteering and monetary donations as the two ways of giving. Although definitions available in the literature cover the various aspects of giving, in the context of this study, the scope of giving behaviour is delimited to giving monetary contributions to charitable organisations.

Aside from the broad range of studies that analyse specific factors that influence giving behaviour, some studies have focused on the synthesis of all factors into a comprehensive model of why and how individuals decide to give. There are three approaches.

One of the latest approaches to giving behaviour is the application of behavioural economics in fundraising. Behavioural economics is a discipline based on the combination of concepts from psychology and economics to understand how individuals behave and make economic decisions (Berg, 2014). One of the main ideas behind the field is that “ample evidence in behavioural research suggests that people systematically deviate from the extreme rational assumption of such economic models” (Hochman & Ariely, 2015). Some authors have contributed to the application of behavioural economics of to fundraisings, such as Dan Ariely (Ariely, Bracha, & Meier, 2009; Hochman & Ariely, 2015) and Francesco Ambrogetti (Ambrogetti, 2016). The more recent trend in the study of giving decisions is a field that has the potential to be influential in the literature in the future.

2.2.1.1 Modelling giving behaviour

One of the conceptual frameworks used by researchers for empirical studies includes two elements: intrinsic and extrinsic factors that affect donor behaviour (Noor, et al., 2015; Chang, 2007; Lee & Chang, 2008), as seen in Figure 1. These two factors are the common denominator; however, there is a variety of determinants chosen by each author within those categories.

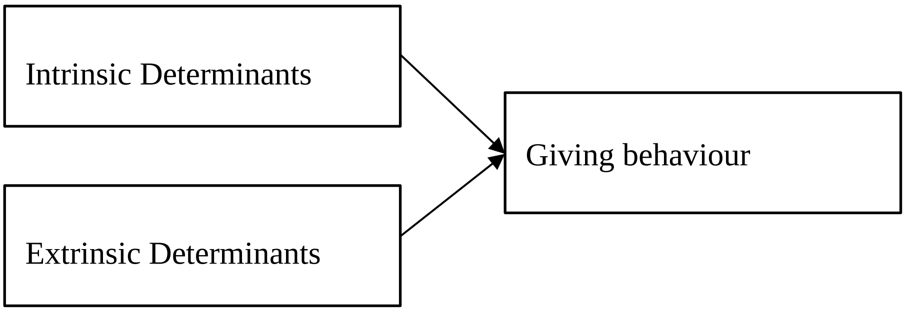

Figure 1. A basic model of giving behaviour. Adapted from: Noor, et al., 2015; Chang, 2007; Lee & Chang, 2008

Sargeant (1999) proposed a comprehensive model that considers more elements and interactions in the modelling of the decision process of giving, as seen in

Figure *2*. The model shows a process of giving that starts with the inputs (how is the donor receiving that triggers a decision, e.g. a donation appeal via a particular channel), followed by the perceptual reaction to the input, influenced by extrinsic and intrinsic determinants, that is then related to processing determinants in which the donor connects with past experiences and judgement. The process results in outputs, that is, how the donor realises the act of giving, e.g. decision to donate, type and size of the donation, loyalty (Sargeant, 1999, S. 218). Based on the model by Sargeant, the following sections explore further the elements relevant to giving behaviour, namely the inputs, intrinsic and extrinsic determinants and outputs.

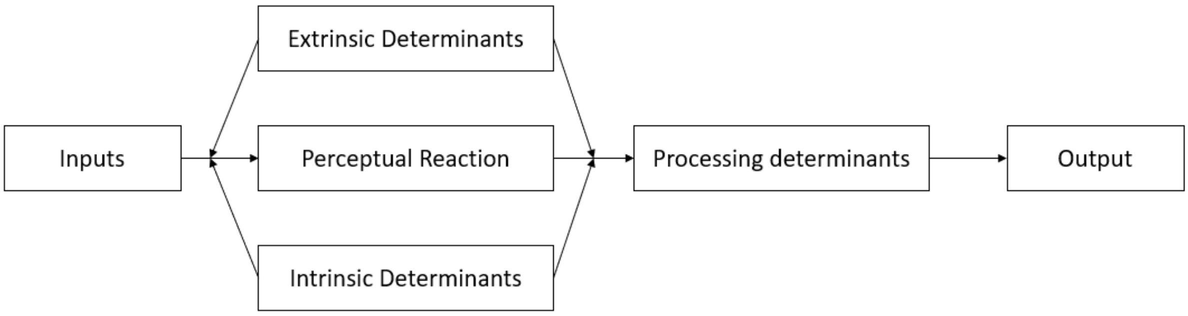

Figure 2. Model of Individual Charity Giving Behaviour. Adapted from (Sargeant, 1999, S. 218)

2.2.1.1.1 Inputs influencing the decision process of giving

In line with the incorporation of marketing theories into the study of giving behaviour, elements related to the way organisations approach donors and convey their messages affect their giving decisions are also part of the study of determinants of giving. Research has found that these factors such as the type of communication with donors, the messages and the use of different fundraising techniques and solicitation approaches, also play a role on the perception that donors have of the organization and the products (Sargeant, Ford, & West, 2006; Srnka, Grohs, & Eckler, 2003). These definitions are of practical importance to non-profits, given that these are the inputs that organisations can influence or control, and therefore the study of the product, appeals and campaigns, as well as brand positioning are essential elements in the communication and fundraising strategies of non-profits. Looking for effective ways to get a positive response and giving from individual donors has led to the adoption of diverse fundraising solicitation techniques and communications.

As the first factor for analysis, Yörük (2009) found that the act of requesting a donation (directly asking a donor to give), increases the propensity of donations by individuals. The second factor that plays a role is the channel or technique with which donors are asked to give. Charities have used a variety of channels, such as Face to Face, Door to Door, Telephone fundraising, Direct Mailing, Media advertisement and in later years, Digital channels. (Srnka, Grohs, & Eckler, 2003, S. 71; Yörük, 2012, S. 472; Sargeant, 1999, S. 217). The selection of these methods is often linked with their effectiveness to get a response from donors, the volume of donors they can attract and the incurred costs. In a comparative analysis of fundraising methods, regarding the effectiveness of these channels to persuade individuals to donate, Yörük (Yörük, 2012, S. 468) found that channels that employ personal solicitations are more effective than cold ones. A third factor to be considered is the messaging and communication presented to donors to convince them to donate, as these affect their perceptions and their reaction to a solicitation. Communicational factors may include the visual elements, the narrative and messages, the writing style, the campaigns launched, the appeals, or the product possibilities (Spears, 2002; Sargeant, 1999, S. 218). The analysis of products and its impact on fundraising performance is of especial importance in this study. For example, Child Sponsorships, one of the products used by SOS CVI, has been widespread across child-related non-profit organizations, noticeable for the high number of children assisted, increased annual income raised by NGOs and the long term support from sponsors achieved via this method (Watson, 2015; Watson & Clarke, 2014).

2.2.1.1.2 Factors influencing giving behaviour of individual donors

The research concerning why and how people donate to charity and the characteristics of charitable is vast (Lee & Chang, 2008, S. 14). In a broad classification, a variety of factors that determine giving behaviour have two categories: intrinsic and extrinsic (Lee & Chang, 2008; Sargeant, Ford, & West, 2006).

Extrinsic determinants of giving "represent the demographic and socio-economic profiles of the charity donors", referring to the characteristics of donors that are inherent to them at the moment of giving. The literature studying the relation between extrinsic factors and giving outcomes suggest that factors as age, gender, income, marital status and social class influence the giving behaviour of individuals (Chang, 2007; Noor, et al., 2015; Najev Čačija, 2013; Schlegelmilch, Diamantopoulos, & Love, 1997; Snipes & Oswald, 2010; Sargeant, 1999; Sargeant, Ford, & West, 2006); however, there are different results in relation to the nature and magnitude of the relationship between variables for different studies. The study of these factors is relevant in practice for non-profits, as it forms the basis for donor segmentation theories and giving behaviour models that aim at the identification of the donors that have a higher potential to donate.

Intrinsic determinants of giving denote the underlying motives that influence the election to donate to a charitable organisation. These can be psychographic and attitudinal (Lee & Chang, 2008, S. 13), related to perceptions, motives and emotions (Sargeant, Ford, & West, 2006, S. 156; Najev Čačija, 2013, S. 62-65). Amongst the factors that influence donor behaviour there are those related to the donor's perceptions of themselves (Schlegelmilch, Diamantopoulos, & Love, 1997), for example trust and commitment (Sargeant, Ford, & West, 2006; de Vries, Reis, & Moscato, 2015), empathy and sympathy (Chang, 2007; Sargeant, Ford, & West, 2006), or generosity and religiosity (Noor, et al., 2015); there are also factors related to the perceptions of donors from organizations, in terms of their efficiency (Harvey & McCrohan, 1998) or the familiarity and experiences that the donor has had with the charity (Najev Čačija, 2013, S. 62-65; Snipes & Oswald, 2010). The analysis of intrinsic determinants is of practical interest for charities, as it examines the role of the organisational factors and the way donors perceive organisations and their offers (Sargeant, Ford, & West, 2006, S. 156). Elements such as strong brand and visibility are relevant to induce a giving reaction, as well as to establish long-term relations with donors, especially in a context of a high number of charities present or entering the scene (Srnka, Grohs, & Eckler, 2003, S. 71).

2.2.1.1.3 Outputs of giving behaviour

The last dimension of the giving behaviour model is the output, that is, how donors support organisations with a monetary donation. The variables that determine the decision output, which non-profit organisations use a series of indicators to measure the giving behaviour of supporters, include:

- Likelihood and decision to donate (Yörük, 2009; Chang, 2007; Lee & Chang, 2008; Sargeant, 1999)
- Selection of the cause to support (Srnka, Grohs, & Eckler, 2003; Najev Čačija, 2013)
- Donation amounts and methods (Srnka, Grohs, & Eckler, 2003; Magson, 1999; Najev Čačija, 2013; Sargeant, 1999), referring to the amount of the monetary gifts made by the donors in a series of transactions with the organisation, and the nature of such transactions
- Donor retention or attrition (Magson, 1999; Sargeant, 1999), including the duration of the engagement of supporters with the organisation or commitment to regular donations
- Frequency of donations and commitment to regular support (Srnka, Grohs, & Eckler, 2003; Sargeant, 1999)

2.3 Donor segmentation

One of the crucial components that non-profit organisations use for a successful relationship marketing approach is donor segmentation and targeting, under the premise that "no organisation can be everything to everybody" (Rupp, Kern, & Helmig, 2014). Donor segmentation is vital in a context of competing for the same potential supporters with other organizations, and bring efficiency in the allocation of organizational efforts and resources (Rupp, Kern, & Helmig, 2014, S. 76; Schlegelmilch, Diamantopoulos, & Love, 1997; Srnka, Grohs, & Eckler, 2003, S. 72).

According to Srnka et al. (2003), organisations address two questions for a systematic segmentation: “(1) How to aggregate donors into similar groups for fundraising purposes; and (2) how to approach each chosen segment, if at all?” (Srnka, Grohs, & Eckler, 2003, S. 71).

Different criteria are useful to address the first question, to analyse and segments potential and existing supporters to non-profit organisations, which can be the results of multiple influencers (Srnka, Grohs, & Eckler, 2003, S. 72). A review of segmentation approaches by Rupp et al. found that related empirical studies use four main criteria for segmentation of donors: sociodemographic, psychographic, behavioural, value-based (Rupp, Kern, & Helmig, 2014, S. 78-79).

Sociodemographic and psychographic criteria are commonly present for defining donor groups in various literature (Najev Čačija, 2013; Schlegelmilch, Diamantopoulos, & Love, 1997; Srnka, Grohs, & Eckler, 2003; Sargeant & McKenzie, 1999), linked to their relationship with determinants of donor behaviour. While these tend to be static, behavioural and value-based segmentation constitutes dynamic segmentation models that “describe how segment size and membership evolve, with segments defined based on RFM statistics” (Durango-Cohen, Torres, & Durango-Cohen, 2013, S. 173). RFM is based on grouping donors according to their contributions in terms of their recency, frequency and monetary value, based on the premise that past behaviour is a predictor of future behaviour (Durango-Cohen, Torres, & Durango-Cohen, 2013, S. 173) and therefore can be associated with the end value of donors. This classification, however, is limited as it does not consider further variables and inputs that determine that behaviours, such as the demographic characteristics, intrinsic variables or the effect of selected fundraising methods. Another way of segmenting current donors and predicting the likelihood of future contributions is by incorporating the elements studied in Sargeant’s model (1999) or other models of giving behaviour (Shelley & Polonsky, 2002).

Concerning the second question of how to address each donor segment, the selection of the right channels, products and communications for defined target audiences plays a crucial role. However, these are decisions that non-profits face, according to their fundraising objectives, defined communication channels, availability of funds, among other factors.

2.4 Theoretical framework

2.4.1 Factors influencing DLTV

From the review of available literature about DLTV, combined with a review of the individual fundraising trends, the theoretical framework defines three categories of factors that influence DLTV (Figure 3) to respond to the research question.

- Socio-demographic features characteristics: sociodemographic features, such as age, gender, income level, religious background, occupation, educational background
- Fundraising methods: how the organisation carried out relationship-building activities, such as the initial acquisition channel, the appeal used to attract the donor and the ongoing communications
- Previous giving behaviour from donors, considering historical data is used to predict DLTV, that is, for example, the initial donation amount, frequency of donations or commitment to a regular giving pledge

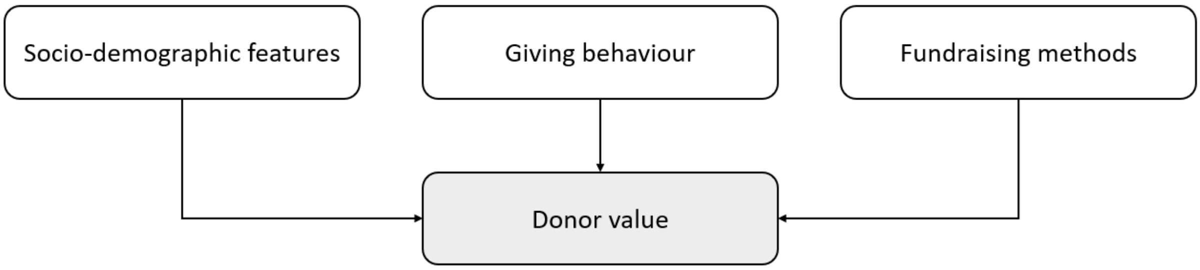

Figure 3. Theoretical framework

2.4.2 Hypotheses

Based on the motivation and scope of the study, the theoretical framework leads to two main hypotheses related to the research questions.

Hypothesis 1. Three types of factors (socio-demographic, fundraising methods and previous behaviour) influence the value of SOS CV's donors

- H1a. Socio-demographic features influence the value of SOS CV's donors

- H1b. Fundraising methods used to attract and retain donors influence the value of SOS CV's donors
- H1c. Past behaviour of SOS CV's donors influence their value

Hypothesis 2. Groups of donors with high value share attributes in terms of socio-demographic factors, fundraising methods and past behaviour

3 Methodology

3.1 Selection of the statistical method: CHAID[4]

A revision of the methods and applications of the documents in the literature review show that the statistical techniques used for the analysis of donor behaviour, segmentation and lifetime were mainly based on regression analysis (Srnka, Grohs, & Eckler, 2003; Sargeant, 1999; Snipes & Oswald, 2010; Durango-Cohen, Torres, & Durango-Cohen, 2013; Rupp, Kern, & Helmig, 2014).

The trends of collection and availability of large amounts of data in businesses lead the emergence of the field of Data Mining, from which the use of Decision Tree analysis was expanded in academia and applied in the for-profit sector (Milanovic & Stamenkovic, 2016, S. 564-565). In the case of non-profit sector analysis, Sargeant (1999) proposed to study donor lifetime value and segmentation using decision tree analysis techniques. The decision tree theoretical framework is “particularly appropriate for the purpose of exploratory knowledge discovery” (Milanovic & Stamenkovic, 2016, S. 564) and has the advantage of being “a simple, but powerful form of multiple variable analysis” (de Ville, 2006, S. 1).

The present study used the Chi-Square Automatic Interaction Detection (CHAID) algorithm as the analysis methodology. The CHAID algorithm is a methodological framework proposed by statistician Kass (1980). CHAID is a tool to determine the relationship between variables by building a prediction model to discover how independent variables (predictors) are best combined and segmented to explain the results in the dependent variable (Díaz-Pérez & Bethencourt-Cejas, 2016, S. 276; Kass, 1980).

CHAID (Kass, 1980) is a multi-variate dependency method, designed for a categorical criterion dependent variable and uses Pearson’s Chi-Square statistic and p-value, and involves two steps:

- selection of relevant variables, using the lowest p-value to hierarchically arrange the predictors in terms of their association with the dependent variable

[4] CHAID stands for Chi Square Automatic Interaction Detection

- merging of categories of each predictor to create a defined number of nodes for the tree, with statistically significant difference among them

The CHAID method “involves testing of hypotheses about the (in)dependence of two variables in each step of the algorithm’s implementation. The logic of testing and formulating the conclusions is identical to the traditional procedure for statistical hypothesis testing, whereby, a software algorithm support enables rapid computation of multiple tests and easy (user-friendly) implementation of heuristic approach in finding the best partition of the observed data set.” (Milanovic & Stamenkovic, 2016).

The strengths of employing CHAID for analysis that influenced the selection of this method are: the easiness of interpretation of results (Milanovic & Stamenkovic, 2016); the non-parametric approach that eliminates the need to prove a normal distribution of the variables (Kass, 1980); and the possibility to consider nominal variables, as well as interval variables, that are categorized, in the analysis (Díaz-Pérez & Bethencourt-Cejas, 2016).

For the specific study, the CHAID technique provides the possibility to go beyond identifying if each independent variables influence the dependent variables or not, to provide insights as to the interaction of variables and the identification of donor groups with specific characteristics, according to their value.

3.2 Data and variables

The pilot study was carried out using data from one European country in which SOS operates fundraising activities. Due to a confidentiality agreement, there is no mention of the country in the study. The analysis model followed the logic of the theoretical framework, considering the availability of data for the definition and calculation of variables in the existing dataset.

Figure 4 shows the model, and the following sections will explain the variables, with the respective calculation method and categorisation, and the steps for the application of the CHAID algorithm.

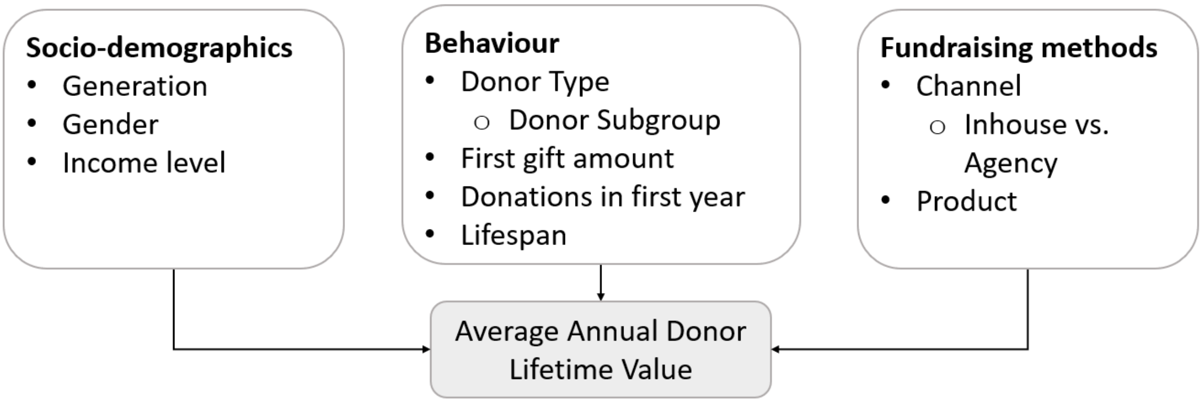

Figure 4. Analysis model framework

SOS CV provided the dataset that corresponds to existing records of the database of the private individual donors of SOS CV in the selected country, stored in the Customer Relationship Management (CRM) system. The dataset consists of a list of private individual donors in the CRM added between 2014 and 2018 and respective characteristics. Identifiable data such as name, address or ID number was not collected. The data was structured and provided in one table that contains the following information per donor:

- Estimated age
- Gender
- Donor income level
- Donor type
- Donor subgroup
- Recruitment channel
- Product type (yes/no)
- Start date (first payment)
- Last/latest payment date
- Amount of the first donation
- Sum of all donations for each donor until the reference date of (30.09.2019)

The dataset contained a total of 293,217 donor records were received. Records in which the field *Sum of all contributions* was empty or equal to zero (739 records) were excluded, resulting in a total of 292,478 records considered for analysis. From those, 33% have incomplete data in at least one of the socio-demographic variables. The decision to use the incomplete observation lies in the fact that there is a higher amount of variables known in these observations that will add to the robustness, allowing more observations to build the model.

From the data provided a series of variables and using calculations were derived, as shown in the analysis model (Figure 4) and described in the following subsections. The continuous variables were categorised according to the context, organisational definitions or the use of quartile distribution, given the need of CHAID of having categorical input and output variables,

3.3 Donor lifetime value

For the measurement of donor value, the approach was to use historic gross donor lifetime value, given that data about the costs for each donor recruitment is not available in the dataset. DLTV was calculated based on the sum of total amounts of all the gifts made by each donor (for new donors starting since January 2014) since their first payment, until the reference date of 30 September 2019. A measure of Annual Average Lifetime Value was used as the output variable, as proposed by Aldrich (2000), to standardise the outcomes considering the time factor. The equation to calculate AADLTV is as follows:

$$Historic\ Gross\ AADLTV = \frac{\sum_{i=1}^{n} Ci}{Life\ opportunity} \quad ,where$$

C= gift amount from the first to the last gift or until 30.09.2019
Life opportunity = number of months since the first donation of each donor until 30.09.2019

As one objective is to identify shared attributes in donors with a high value, donors were divided into two classes according to their AADLTV: Mid-low and High (see Table 4). Percentiles were used to identify and separate the 20% of donors with higher AADLTV (High), to aid the identification of high-value groups and the factors influencing them.

Table 4. *Annual Average Donor Lifetime Value (AADLTV) Classes*

Class	**Criteria**
Low-mid	Lower or equal to the percentile 0.8 of the data
High	Higher than the percentile 0.8 of the data

3.4 Socio-demographic variables

Age, gender and income level are the factors selected for this section, given the availability of data, forming the following variables:

- Gender: directly extracted from the database, with two possibilities (male, female).
- Generation: for age, a set of generations and respective age bands were defined using the Age field.
- Income level is a variable in the dataset that classifies donors in 5 categories, defined by external factors related to the location of the donor's residence.

Table 5 shows the variables and categories in the dataset. As mentioned before, three variables, there are unknown records, marked as such in the categories.

Table 5. *Socio-demographic variables and categories*

Variable	Code	Description
Generation		
Elderly	ELDE	Born pre-1928
Silent	SILE	1928-1945
Boomers	BOOM	1946-1964
Gen X	GENX	1965-1980
Millennials	MILL	1981-1996
Gen Z	GENZ	after 1996
Unknown	UNK	N/A
Gender		
Male	M	Is male
Female	F	Is female
Unknown	UNK	Gender not known
Income level		
Very low	1-	Income level is very low
Below average	2-	Income level is below average
Average	3-	Income level is average
Above average	4-	Income level is above average
Very high	5-	Income level is very high
Unknown	UNK	Income level not known

3.5 Behaviour variables

Giving behaviour has five variables, defined according to the literature review and possible to calculate using the existing dataset. There are two variables related to the classification of donors (Donor type and Subgroup) and three related to the initial donations of the donor (First gift amount, Donations in the first year and Lifespan).

3.5.1 Donor types, categories and subgroups

The first consideration is how to treat the data for different types of donors, as conceptually this is a critical variable that represents the treatment from the organisation to the donors, and subsequently their engagement and behaviour. Donor type denotes the type of relationship or commitment the donor has with the organisation, defining if there are regular or single donors and the volume of contributions. Therefore, it is relevant to customise the analysis for different donor types. The donor base was classified accordingly in four Donor Categories that grouped the Donor Types, as shown in Table 6 and Table 7. A separate CHAID model was built for each Donor Category defined, given the evident differences in AADLTV levels.

Table 6. *Donor Category and Donor Type description*

Donor category **Donor type**	**Description**
Single giver	A donor who made single donations and is not committed to regular donations for the organisation
Sporadic donor	Single giver with single donations
Reactivated donor	Single giver that gave in the past and was reactivated to donate again during the period of study
Committed giver (CG)	A donor with a commitment to a regular unrestricted donation
Sponsor	A donor with a commitment to a regular restricted donation that corresponds to funding a sponsorship (e.g. one child, one village)
Mid-major donor	A donor that gives significantly higher donations than other donor types that have a high impact on the organisation
Major donor	Donor engages through the Major donor program, which targets High Net Worth Individuals (HNWI), offering customised treatment
Major sponsor/CG	Sponsor or committed giver with commitments to a significantly higher regular donation than the bulk of regular donors
Mid value donor	Donor without a regular donation pledge who donated higher amounts to the organisation, not being an HNWI

Table 7. *Donor Category and Donor Type summary*

Donor category Donor type	AADLTV			Count	%
	Min	Average	Max		
Committed giver (CG)	2	142	6,000	15,615	5%
Mid-major donor	153	1,186	379,355	3,921	1%
Major donor	696	8,023	379,355	208	0%
Major sponsor/CG	222	792	7,200	1,964	1%
Mid value donor	153	815	4,667	1,749	1%
Single Giver	0.0	33	86,383	262,111	90%
Reactivated donor	1.02	22	1,469	1,181	0%
Sporadic donor	0.002	33	86,383	260,930	89%
Sponsor	0.5	383	1,478	10,831	4%
Grand Total	**0.002**	**67**	**379,355**	**292,478**	

The amount of incomplete donor records varies with the Donor Category and is especially relevant for Single and Mid-major donors, as shown in Table 8, and is relevant for the interpretation of results.

Table 8. *Complete donor records by Donor Category*

Donor category	Complete records	Total records	% complete records
Committed giver	14,464	15,615	93%
Mid-major donor	2,995	3,921	76%
Single giver	169,630	262,111	65%
Sponsor	9,030	10,831	83%
Total	**196,119**	**292,478**	**67%**

Donor Subgroup is a sub-classification of each Donor Type (see Table 9), that explores in more details for each group, according to characteristics from the donors' previous giving behaviour with the organisation.

Table 9. *Subgroups by Donor Category and Donor Type*

Type/ Subgroup	Code	Description	Count	Relative %
Single Giver			**262,111**	**90%**
Sporadic donor			**260,930**	**99.5%**
Donor group 1	SG_1	High donation frequency and amount	7,189	3%
Donor group 2	SG_2	Low donation frequency and high amount	12,708	5%
Donor group 3	SG_3	High donation frequency and low amount	3,402	1%
Donor group 4	SG_4	Low donation frequency and amount	25,083	10%
Donor group 5	SG_5	New donor	88,257	34%
Donor group 6	SG_6	One-time donor	39,739	15%
Donor group 7	SG_7	Inactive donor	21,957	8%
Donor group 8	SG_8	Donor with bad postal adress	61,469	24%
New unvalued	SG_UNV	Donor is too recent to categorize	1,126	0%
Reactivated donor			**1,181**	**0.5%**
Reactivated donor	REAC_DON	Donor lapsed who gave after a long period	1,181	100%
Committed Giver			**15,615**	**5%**
Committed giver	CG	Donor with committment to a regular unrestricted donation	13,138	84%
Support member	CG SUPP	Support members get a certificate for supporting our association	2,471	16%
Full member	CG FULL	Member of the SOS Children's Village association	6	0%
Sponsor			**10,831**	**4%**
Gift sponsor	GIFT_SPO	Someone who pays for somebody else's sponsorship	436	4%
International child sponsor	IC_SPO	Has a committment as international child sponsor	6,492	60%
International village sponsor	IV_SPO	Has a committment as sponsor to an international village	3,085	28%
National village sponsor	NV_SPO	Has a committment as sponsor to a national village	818	8%
Mid-major			**3,921**	**1%**
Major sponsor/committed giver			**1,964**	**50%**
Mid to high value Spo/CG	MJREG_HIGH	Annual donation amount >= 2,000€ and <4,000€ in one of the last 3 years or in the current year	47	2%
Mid value Spo/CG	MJREG_MID	Annual donation amount >= 600€ and <2,000€ in one of the last 3 years or in the current year	1,917	98%
Mid value donor			**1,749**	**45%**
Mid to high value donor	MID_HIGH	Sponsor or CG with annual donation amount >= 2,000€ in one of the last 3 years or the current year	246	14%
Mid value donor	MID_MID	Sponsor or CG with annual donation amount >= 600€ in one of the last 3 years or the current year	1,503	86%
Major donor			**208**	**5%**
Active major donor	AC_MAJ	Annual donation amount >= 4,000€ in one of the last 3 years or in the current year	135	65%
Former active major donor	FACT_MAJ	Annual donation amount >= 4,000€ in one year since 1996 but not in one of the last 3 years or current year	42	20%
Former top major donor	FTOP_MAJ	Annual donation amount >= 20,000€ in one year since 1996 but not in one of the last 3 years or current year	9	4%
Premium major donor	PREM_MAJ	Annual donation amount >= 75,000€ in one of the last 3 years or current year	4	2%
Top major donor	TOP _MAJ	Annual donation amount >= 20,000€ in one of the last 3 years or current year	18	9%
Total			**292,478**	

3.5.2 Behaviour variables

Three variables for giving behaviour were defined:

- First gift amount: Volume of the first-ever gift of the donor to the organisation. The field is in the dataset in the field of Amount of the first donation
- Donations in the first year: number of unique donations made during the first year after becoming a donor
- Lifespan: classifies the number of months that a donor remains as a donor, since the first gift to the last one made, or to the latest possible month (30.09.2019)

Table 10 describes de categories for Donations in the first year and Lifespan. First gift amount was classified differently for each Donor Type, as there are considerable differences between them. Each section for the donor category specifies the ranges for this variable.

Table 10. *Categories for Donations in the first year and Lifespan*

Variable	Code	Description
Donations in the first year		
Only one	One	The donor had only one donation in the first year
More than one	>= 2	The donor had more than one donation in the first year
Lifespan		
< 12 months	<12m	Last payment less than 12 months after the first gift
> = 12 months	12-24m	Last payment between 12 and 24 months from the first gift
New donor	>24m	The donor has a life opportunity lower than 12 months

3.6 Fundraising methods variables

Fundraising methods are defined using three variables: channel, product and in-house/agency. An analysis of the relation of variables showed that in the country of study, specific channels are used either with an in-house team or with an agency, not with a combination of both in most cases. Face to Face only works via agency and Digital is exclusively in-house, Direct Mail is by an agency in 99% of the cases and for other channels 99% of the cases this field is unknown. Thus, the classification of In-house vs Agency was excluded from analysis, as it was redundant and directly linked to the channel.

Table 11 shows the variables for fundraising methods.

- Channel: is the method that the organisation used to acquire the donor
- In-house vs Agency: classified according to whether the team recruiting the donors is hired internally (In-house) or through an agency (Agency).
- Product: is the value proposition that the donor is committed to when giving donations if any. Each product is listed for each donor with a "yes" or "no" dichotomy. The four possible products are Single Giving product, Committed Giving product, Village sponsorship and Child sponsorship.

An analysis of the relation of variables showed that in the country of study, specific channels are used either with an in-house team or with an agency, not with a combination of both in most cases. Face to Face only works via agency and Digital is exclusively in-house, Direct Mail is by an agency in 99% of the cases and for other channels 99% of the cases this field is unknown. Thus, the classification of In-house vs Agency was excluded from analysis, as it was redundant and directly linked to the channel.

Table 11. *Fundraising methods: channels and products*

Variable	**Code**	**Description**
Product		
Single Giving Product	Prod_SG	No commitment to regular donations
Committed Giving	Prod_CG	Commitment to an unrestricted regular donation
International Child Sponsorship	Prod_Child_Spo	Commitment to a regular donation with restricted use to sponsor a specific child
International Village Sponsorship	Prod_Vill_Spo	Commitment to a regular donation with restricted use to sponsor a specific village
Channel		
Other	OTHER	Includes Spontaneous gifts and groups channels with less frequency (Inserts, Events and Press)
Digital	DIG	The organisation sents a direct physical mail to potential donors, with a request and a payment slip
Direct Mail	DM	Donation is received via Digital channels, e.g. social media, landing page
Face to Face	F2F	Potential donors are approached in the streets or private sites by an agent to request for donations directly

Variable	Code	Description
In-house/Agency		
Agency	AGE	SOS CV contracts an agency for acquiring donors
In-house	INH	An internal team is the one recruiting donors
Unknown	UNK	Not known or non-applicable

3.7 Application of the CHAID algorithm

The statistical analysis was performed using R language (R Core Team, 2019), supported in RStudio (RStudio Team, 2018), an open-source solution for statistical analysis, using as core of the analysis the CHAID R package (The FoRt Student Project Team, 2015). Five steps were taken to build and check the CHAID model in R for each of the four Donor Categories, using the Donor Class as the dependent variable.

1. Selection of variables for the model

The first step is to determine which of the independent variables are relevant for each Donor Category, according to the particularities of the groups. Table 12 summarises the independent variables considered for each model.

Table 12. *Independent variables used for the CHAID model, according to Donor Category*

Variable	Sponsor	Committed Giver	Single Donor	Mid-major donor
Prod_SG	x	x		x
Prod_CG				x
Prod_Child_Spo				x
Prod_Vill_Spo				x
Gender	x	x	x	x
Generation	x	x	x	x
Income_level	x	x	x	x
Subgroup	x	x	x	x
Channel	x	x	x	x
Lifespan	x	x	x	x
First_Donation	x	x	x	x
Donations_first_year	x	x	x	x

The models for Sponsors and Committed Givers did not include the variables for the regular giving product, as the donor Types and Subgroups available reflect the products used in more detail. The model for Single Givers excluded the product as, by definition, this category only has single gifts. The model for Mid-major donors considered all variables.

2. Refinement of the initial CHAID model

The datasets for each donor group were loaded to RStudio, and a first model was built using the *chaid* function from the CHAID R Package (The FoRt Student Project Team, 2015), using the default parameters. Each model was assessed to define if it was necessary to exercise control in the parameters using the *chaid_control* function, described in Table 13.

Table 13. *Parameters of the chaid_control function in RStudio*

Argument	**Default**	**Description**
alpha2	0.05	Level of significance used for merging of predictor categories (step 2).
alpha3	-1	If set to a positive value < 1, level of significance used for splitting of former merged categories of the predictor (step 3). Otherwise, step 3 is omitted (the default).
alpha4	0.05	Level of significance used for splitting of a node in the most significant predictor (step 5).
minsplit	20	The number of observations in splitted response at which no further split is desired.
minbucket	7	The minimum number of observations in terminal nodes.
minprob	0.01	Minimum frequency of observations in terminal nodes.
stump	FALSE	- if TRUE, limits height to 1 (decision stump/1 rule learning)
maxheight	-1	- max height of the tree. no limit if -1

Source: R Documentation (RStudio Team, 2018)

3. Model evaluation and cross-validation

A series of criteria can be used to assess the quality of the models. Accuracy is the most common criteria, and it denotes “the proportion of correctly classified data using the designed model, contrary to the concept of error which indicates wrongly classified observations” (Milanović & Stamenković, 2016).

One of the tools to assess accuracy for classification problems is a two-dimensional matrix called Confusion Matrix. Such matrix “calculates a cross-tabulation of observed and predicted classes with associated statistics" (RStudio Team, 2018), that is, the number of correctly and incorrectly classified observations in each High and Low-mid category. Its elements "represent testing results of the predictive model" (Milanovic & Stamenkovic, 2016) and allowed the measurement of accuracy and predictability of the models. The three models were analysed for accuracy of classifications using the Confusion Matrix, with the aid of the functions *predict* and *confusionMatrix***,** functionalities available in the caret package (Kuhn, et al., 2019), analysing the following parameters:

- Accuracy: percentage of total correctly predicted observations
- Positive predictive value: percentage of "High" correctly predicted observations
- Negative predictive value: percentage of "Low-mid" correctly predicted observations

A second method used for evaluation of the models' performance in terms of its predictive capability was k-fold cross-validation, which assesses how adequately data is split into subsets/subsamples for model training, validation and testing (Milanovic & Stamenkovic, 2016). In the k-fol cross-validation, as explained by Milanovic & Stamenkovic (2016)

> *"the original set of observed data is first randomly divided into k disjunctive partitions of approximately same size, and then, the evaluation process is conducted through k iterations, as follows: in each iteration, a single subset/partition is selected for testing while the union of other subsets (k-1) is used for model training. Training and testing are carried out the same number of times".*

For each model, 10-fold cross-validation was performed using the packages *rsample* (Kuhn, Chow, & Wickham, 2019) and *caret* (Kuhn, et al., 2019). Each dataset was split into two sets, a training set of 70% of the data and a test set of 30% of the data to perform the cross/validation. The model was trained using the *train* function with the argument method "chaid". Subsequently, a confusion matrix was calculated for the resulting training set, based on the desired metric "Accuracy" and results were compared to the confusion matrix of the original model, to evaluate that there were no high deviations in the accuracy of both models and therefore no overfitting.

4 Results

The procedure described in Section 343 applied to the datasets for each donor category. Table 14 summarises the percentage of donors allocated to the AADLTV class “High” for each donor category. This section describes the considerations and results of the application of the CHAID model in each category.

Table 14. *Distribution of “High” and “Low-mid” AADLTV classes per donor category*

Variable	High		Low-mid	
	Count	%	Count	%
Sponsor	2,168	20%	8,663	80%
Committed Giver	3,376	22%	12,239	78%
Single Giver	53,589	20%	208,522	80%
Mid-major donor	791	20%	3,130	80%

For each donor category (displayed in an individual section), a summary of the observations and allocation to the variables and categories is displayed, followed by the description and plot of the resulting model, as well as a discussion of the results.

The first assessment was which variables were considered as statistically significant by the model, which indicated the influence of the factors on the value of donors. Variables that were considered by the model indicate that his factor is statistically significant to the AADLTV Class as the outcome variable. In contrast, variables not considered by the model indicated that the variable is not significant to AADLTV, and therefore the factor is not influential to the result.

The second examination was around identifying the groups of donors with “High” AADLTV and which attributes were common to these. For this purpose, the identification of terminal nodes resulting in the majority of donors in the “High” class (high-value nodes), followed by a detailed analysis of which variables are determinant to their categorisation and grouping, was an indication of shared features of high-value donors. The analysis included the breakdown of variables in each factor (Socio-demographic, Behaviour, Fundraising methods) that appeared in these nodes and how these interacted to define groups of donors with high-value. Additionally, a calculation of the average AADLTV for the nodes was displayed, as a hint that confirms that donors in high-value nodes have, on average, higher AADLTV than the total population of donors.

4.1 CHAID model for Sponsors

The category of Sponsor had 10,831 observations. Table 15 displays the distribution of the data into variables and categories. Two categories were defined for the first donation amount for Sponsors: less or equal to € 40 and over € 40. The division related to the pricing of Sponsorships (around € 30 to 40 per month) and lead to most of the observations being in the range of less or equal to € 40. A donation higher than € 40 can be the result of donors supporting multiple sponsorships or with lower donation frequency, e.g. pay once every year.

Table 15. *Variables and categories for Sponsors*

Code	Count	%
Generation		
ELDE	4	0%
SILE	363	3%
BOOM	2,303	21%
GENX	3,602	33%
MILL	4,157	38%
GENZ	324	3%
UNK	78	1%
Channel		
DIG	6,493	60%
DM	160	1%
F2F	3,104	29%
OTHER	345	3%
SPONT	729	7%
Income level		
1-	1,395	13%
2-	1,903	18%
3-	2,933	27%
4-	2,175	20%
5-	2,021	19%
UNK	404	4%

Code	Count	%
Gender		
M	4,615	43%
F	4,776	44%
UNK	1,440	13%
Donations in the first year		
One	276	3%
>= 2	10,555	97%
Lifespan		
>= 12 m	10,049	93%
< 12 m	102	1%
NEW	680	6%
First Donation		
<= 40	8,654	80%
> 40	2,177	20%
Product		
Prod_SG	No: 10145	94%
	Yes: 686	6%

4.1.1 Results for Sponsors

According to the selection of the variables, the model formula used to run the CHAID function was the following:

$$AALTV_{Class} \sim Prod\_SG + \ Gender \ + \ Generation \ + \ Income \ level + \ Subgroup \\ + \ Channel \ + \ Lifespan \ + First \ Gift \ Amount \ + \ Donations \ first \ year$$

The result was a model with 23 inner nodes and 28 terminal nodes, from which eight are terminal high-value nodes. The model used eight of the nine input variables and excluded Gender as it did not find statistically significant.

Figure 5 shows the plot for the resulting model (see the Appendix for full details of all nodes). Bars filled mostly in black indicate low-mid-value nodes, whereas bars filled mostly in white indicate high-value nodes.

The first split that the model finds is in the First donation, as the variable with higher importance for the AADLTV class. For donors with First donations lower than € 40, the following split is according to Prod_SG (single donations), and subsequently, other variables like Channel, Subgroup, Lifespan, Donations_first_year and Generation were used to split the tree further. None of the terminal nodes in this split was high-value. For donors with First donation higher than € 40, the model used Channel as the second split and subsequently Prod_SG, Lifespan, Donations in the first year, Subgroup and Income level to split the group further. There are eight high-value nodes in this split, displayed in Table 16.

These results indicated that variables related to Behaviour are highly relevant for the classification of high-value nodes, with First Donation determining the first split and the other three variables (Lifespan, Donations_first_year and Subgroup) being relevant in the categorisation of various high-value nodes. Fundraising methods are also relevant for the analysis, with both Channel and Prod_SG being part of the definition of various high-value nodes. Socio-demographic factors were not found to be relevant common traits that determine high-value nodes in the case of Sponsors. Only Income level appears in one of these nodes, however, covering all possible known categories, therefore is not a differentiating factor. Generation is not relevant for the grouping sponsors in the high-value nodes, and Gender is not significant for the model.

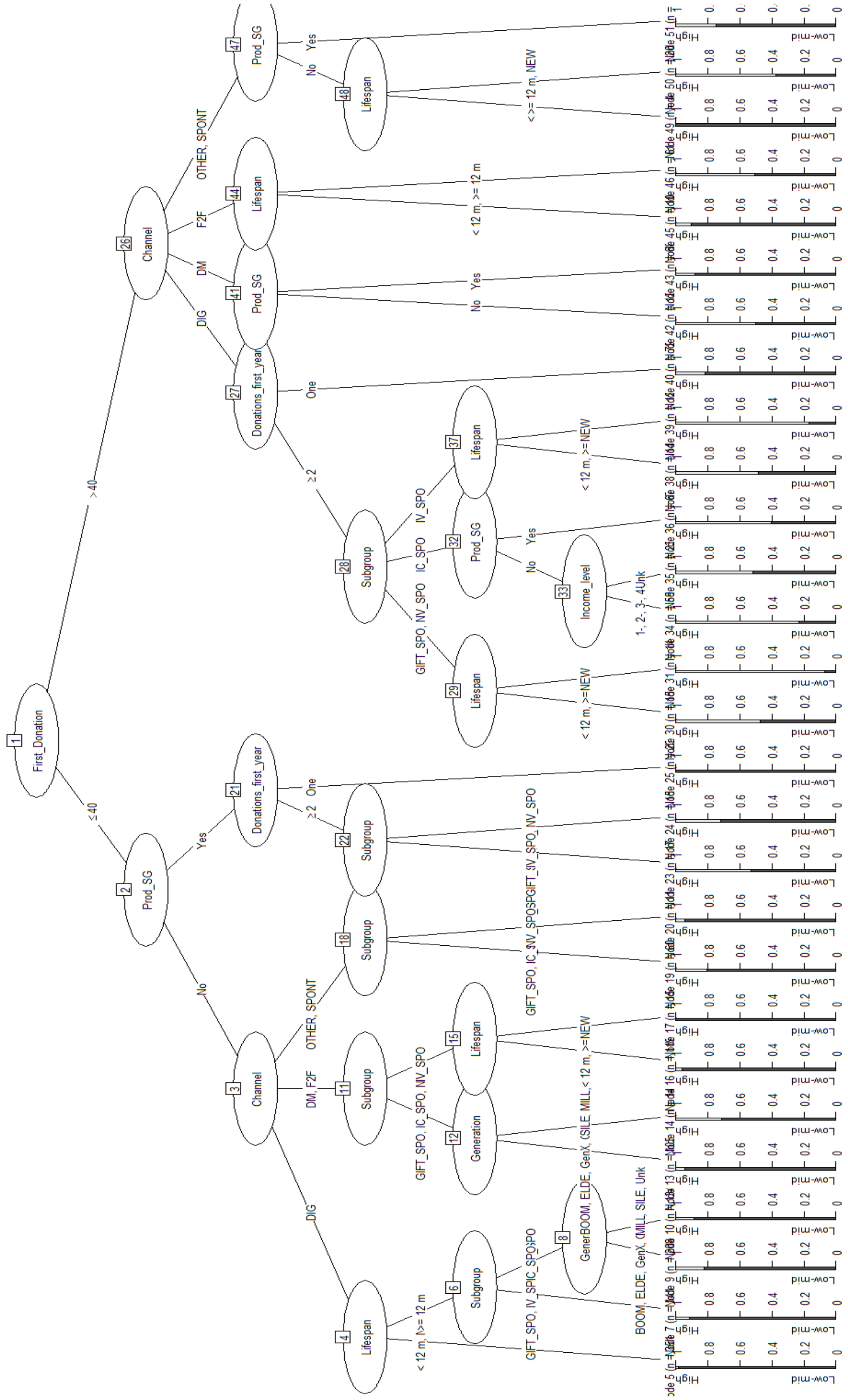

Figure 5. CHAID Model for Sponsors

Table 16. *High-value nodes from the model for Sponsors*

Node #	Count	Error	Variables							Avg AADLTV
			First Donation	Channel	Lifespan	Prod_SG	Donations first year	Subgroup	Income level	
30	156	47%	> 40	DIG	< 12 m, >= 12 m		>= 2	GIFT_SPO, NV_SPO		436.7
31	14	7%	> 40	DIG	New		>= 2	GIFT_SPO, NV_SPO		499.6
34	590	23%	> 40	DIG		No	>= 2	IC_SPO	1-, 2-, 3-, 4-, 5-	469.1
36	67	40%	> 40	DIG		Yes	>= 2	IC_SPO		428.4
38	148	49%	> 40	DIG	< 12 m, >= 12 m		>= 2	IV_SPO		447.3
39	12	17%	> 40	DIG	New		>= 2	IV_SPO		505.6
42	12	50%	> 40	DM		No				411.5
50	203	37%	> 40	OTHER, SPONT	>= 12, New	No				439.6

For the analysis, some high-value nodes were subsequently regrouped, in the cases where only one variable differentiates the nodes and the grouping ensures that all categories of that variable are covered. That is the case for nodes 30 and 31, for example, where the nodes are only different in Lifespan; however, these two nodes cover all three categories, creating one single group for analysis. A similar case occurs in nodes 38 and 39.

There are 1,202 sponsors in the eight nodes which represent 55% of the sponsors in the AADLTV class "High". Five groups of high-value sponsors arose from these nodes:

- Initial donation over € 40
    - Via Digital
        - with more than one donation in the first year
            - Gift, National Village and International Village sponsors
            - International Child sponsors
                - without single donations, where income is known (levels 1 to 5)
                - with single donations
    - Via Direct Mail without Single donations

- Via Spontaneous gift or other channels that stayed loyal for over 12 months, or have not yet achieved 12 months as donors

The majority of sponsors in high-value nodes have as attributes a First donation over € 40, were engaged via Digital and have two or more donations in the first year. Although differentiated nodes are identified that consider other factors, these attributes are generalised within all subgroups. For sponsors coming via Spontaneous Gifts and Other, a longer lifespan (over 12 months) is a determinant factor of high lifetime value. In the case of sponsors via Direct Mail, the high error rate and the small number of sponsors suggest that the attributes in this node are not particularly common in donors with high value. There is no high-value node from sponsors engaged via Face to Face, the second most used channel to recruit sponsors. These groups have an AADLTV above average, compared to the AADLTV of € 383 for the population of Sponsors.

4.1.2 Evaluation of the model for Sponsors

The evaluation of the model is based on a confusion matrix and the results of the 10-fold cross-validation. Table 17 displays the results, which show an overall accuracy of 83%, which is a good indication of the model performance. However, only 67% of sponsors with "High" AADLTV were correctly classified, suggesting that the capacity of the model for its purpose is acceptable but could be improved, as there is the risk that the model does not faithfully represent the interaction between the dependent and independent variables.

Table 17. *Confusion matrix and accuracy of models for Sponsors*

Confusion Matrix Sponsorship model		Predicted	
		High	Low-mid
Observed	High	808	394
	Low-mid	1,360	8,269

Measures	Accuracy	Pos pred value	Neg pred value
	0.8381	0.6722	0.8588

The 10-fold cross-validation resulted in an accuracy of the training set of 83.8% for predictions on the test data, similar than that of the original model, suggesting that the default parameters used by CHAID are adequate and there is no overfitting.

4.2 CHAID model for Committed Givers

There are 15,615 observations in the category of Committed Givers. Table 18 shows the distribution of the data into variables and categories. The split of the First donation was set using percentiles to divide the set into three similarly sized groups and resulted in three categories: less than € 20, between € 20 and € 50, and greater than € 50.

Table 18. *Variables and categories for Committed Givers*

Code	Count	%
Generation		
ELDE	52	0%
SILE	1,395	9%
BOOM	3,277	21%
GENX	3,353	21%
MILL	6,121	39%
GENZ	1,364	9%
UNK	53	0%
Channel		
DIG	3,151	20%
DM	2,370	15%
F2F	8,906	57%
OTHER	99	1%
SPONT	1,089	7%
Income level		
1-	2,388	15%
2-	2,864	18%
3-	4,354	28%
4-	2,917	19%
5-	2,569	16%
UNK	523	3%

Code	Count	%
Gender		
M	7,519	48%
F	7,472	48%
UNK	624	4%
Donations in the first year		
One	2,177	14%
>= 2	13,438	86%
Lifespan		
>= 12 m	12,951	83%
< 12 m	440	3%
NEW	2,224	14%
First Donation		
<20	6,168	40%
20 to 50	4,650	30%
>50	4,797	31%
Product		
Prod_SG	No: 11624	74%
	Yes: 3991	26%

4.2.1 Results for Committed Givers

According to the selection of the variables, the model formula used to run the CHAID function was the following:

$$AALTV_{Class} \sim Prod\_SG + Gender + Generation + Income\ level + Subgroup + Channel + Lifespan + First\ Gift\ Amount + Donations\ first\ year$$

The resulting model had 33 inner nodes and 44 terminal nodes, from which 11 are high-value terminal nodes (see Table 19)

Figure 6 for the plot and the Appendix for full details of all nodes). The model used eight of the input variables and excluded Income level, not considering the variable as statistically significant.

The first variable that splits the model is the First Donation. There were no high-value nodes for donors with a first donation lower than € 20, which indicates that not many donors in that category have high value. For donors in the € 20 to € 50 category, the second division was according to channels, followed by different factors for each channel, including Donations_first_year (for Digital and Other), Prod_CG (for Direct Mail and Spontaneous gifts) and Generation (for Face to Face). In the group of donors over € 50, the second split was Donations_first_year, followed by Channel and Lifespan, and subsequent splits were made using Prod_SG, Subgroup, Gender and Generation.

Table 19. *High-value nodes from the model for Committed Givers*

Node #	**Count**	**Error**	**Variables**							**Avg AADLTV**
			First Donation	**Channel**	**Lifespan**	**Prod_SG**	**Donations first year**	**Subgroup**	**Gender**	
19	56	45%	> 50	DIG, OTHER	New		>= 2	CG, CG FULL		370.6
29	32	34%	> 50	F2F	< 12m		>= 2		M	190.5
34	42	41%	> 50	SPONT	New, < 12m		>= 2			240.0
52	597	29%	20 to 50	DIG, OTHER	>= 12, < 12m	No	>= 2	CG, CG FULL		224.3
53	91	10%	20 to 50	DIG, OTHER	New	No	>= 2	CG, CG FULL		220.3
54	126	11%	20 to 50	DIG, OTHER		No	>= 2	CG SUPP		263.4
57	91	44%	20 to 50	DIG, OTHER		Yes	>= 2		M, UNK	242.1
60	16	44%	20 to 50	DM		No				199.1
71	30	50%	20 to 50	SPONT		No		CG, CG FULL		169.3

72	75	25%	20 to 50	SPONT		No		CG SUPP		234.4
76	47	49%	20 to 50	SPONT		Yes	>= 2		M	190.8

Different variable groups were determinants of high-value donors in the model for Committed Givers. From behavioural factors, all variables (First Donation, Lifespan, Subgroup and Donations_first_year) were influential in the determination of high-value nodes. In the section of factors related to fundraising methods, the variables used (Channel, Prod_SG) contributed to determining various terminal high-value nodes. From the sociodemographic variables, Gender was the only one detected in the determination of three high-value nodes, with Male being the category observed in such nodes. Income level was not found significant to the model and Generation, although considered in the model, was not a decisive factor to the classification of high-value nodes.

From the terminal nodes marked as “High”, eight correspond to donors with the first donation between € 20 and € 50 and 3 to donation over € 50, forming the seven groups:

- First Donation over € 50 with two or more donations in the first year
  - Via Digital or Other channels in the Subgroups CG and CG FULL and Lifespan New
  - Via Face to Face with lifespan lower than 12 months and male
  - Via a Spontaneous gift with Lifespan New or lower than 12 months
- First donation between € 20 and € 50
  - via Digital or Other channels with two or more donations in the first year
    - Without single donations
    - With single donations, male or unknown gender
  - via a Spontaneous Gift
    - without single donations
    - with single donations, two or more gifts in the first year and male

The average AADLTV for these groups of committed givers are above the average of € 138 for all committed givers.

Most of the committed givers in the high-value nodes had the first donation between € 20 and € 50, were engaged via Digital and Other channel and had two or more donations in year 1. For givers with no single donations in this group, all three subgroups are represented (CG, CG FULL and CG SUPP), and givers with single donations are male. For givers starting with Spontaneous gifts, the majority in these nodes did not make single donations; if they did, similar to the case of Digital and Other channels, the common factor is that they were male, and additionally that they made two or more donations in the first year. It can be highlighted that for Digital and Other channels across nodes, a common trait is two or more donations in the first year, indicating how crucial the loyalty in the first months is for these channels.

For the groups of committed givers with a first donation over € 50 and at least two donations in the first year, a short lifespan is observed for these high-value donors, indicating that even though donor in these groups stay for a shorter period, higher donations drove their value.

4.2.2 Evaluation of the model for Committed Givers

The results of the confusion matrix for the model (Table 20) showed an overall accuracy of 81.7%, with a lower positive predictive value for donors in the "High" AADLTV Class of 71.6%. Results indicated a reliable overall performance of the model, however, with less reliability to classify the high-value donors, which is the main objective.

Table 20. *Confusion matrix and accuracy of models for Committed Givers*

Confusion Matrix Sponsorship model		Predicted	
		High	Low-mid
Observed	High	861	342
	Low-mid	2,515	11,897

Measures	Accuracy	Pos pred value	Neg pred value
	0.8170	0.7157	0.8255

The 10-fold cross-validation resulted in an accuracy of the training set of 81.3% for predictions on the test data, which suggests that the original model with the default parameters is an adequate alternative and there is no overfitting.

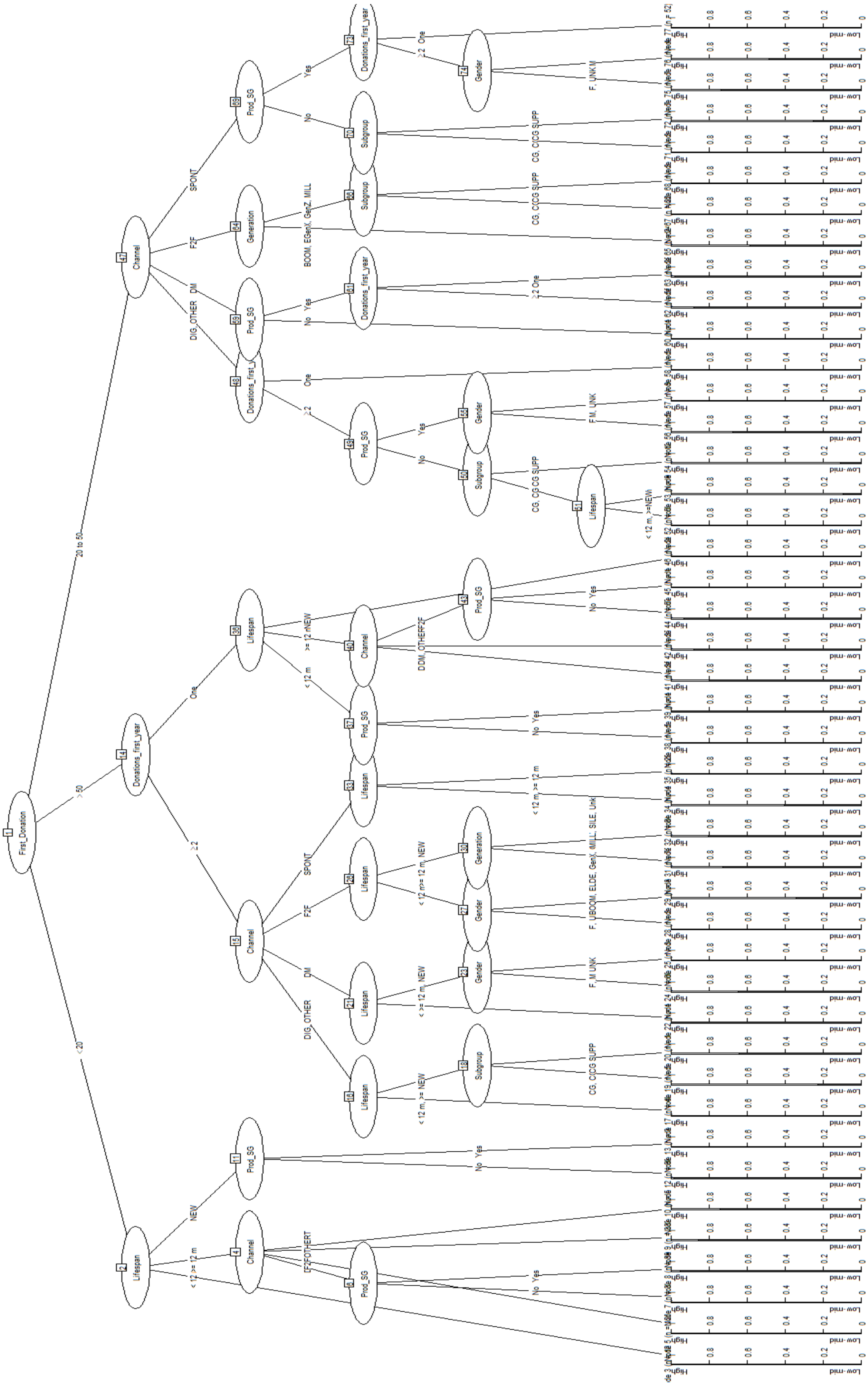

Figure 6. CHAID Model for Committed Givers

4.3 CHAID model for Mid and Major donors

Data were available for 3,921 donors in the category of mid-level and major donors. Table 21 displays the distribution of data and variables. The split of the First donation was set using percentiles to divide the dataset into three groups of similar size, with the following categories: less than € 100, between € 100 and € 1,000, and higher than € 1,000.

Table 21. *Variables and categories for Mid and Major Donors*

Code	Count	%
Gender		
M	2,003	51%
F	1,273	32%
UNK	645	16%
Donations in the first year		
One	1,253	32%
>= 2	2,668	68%
Lifespan		
>= 12 m	2,852	73%
< 12 m	816	21%
NEW	253	6%
First Donation		
<100	1,637	42%
100-1,000	1,279	33%
>1000	1,005	26%
Product		
Prod_SG	No: 3639	93%
	Yes: 282	7%
Prod_CG	No: 3609	92%
	Yes: 312	8%
Prod_Child_Spo	No: 2789	71%
	Yes: 1132	29%
Prod_Vill_Spo	No: 3269	83%
	Yes: 652	17%

Code	Count	%
Generation		
ELDE	25	1%
SILE	600	15%
BOOM	1,563	40%
GENX	1,203	31%
MILL	471	12%
GENZ	14	0%
UNK	45	1%
Channel		
DIG	2,094	53%
DM	355	9%
F2F	86	2%
OTHER	101	3%
SPONT	1,285	33%
Income level		
1-	371	9%
2-	498	13%
3-	997	25%
4-	740	19%
5-	999	25%
UNK	316	8%

4.3.1 Results for Mid and Major donors

The formula used to run the model includes all 12 available variables, as described previously in Table 12, as follows:

$$AALTV_{Class} \sim Prod\_SG + Prod\_CG + Prod\_Child\_Spo + Prod\_Vill\_Spo + Gender + Generation + Income\ level + Subgroup + Channel + Lifespan + First\ Gift\ Amount + Donations\ first\ year$$

The resulting model displayed in

Figure 7 (see the Appendix for full details of all nodes), considered ten variables and excluded Prod_Child_Spo and Gender as non-significant. The tree has 23 inner nodes and 29 terminal nodes, from which 10 are high-value terminal nodes (see Table 22).

The first split was according to the Subgroups, meaning this is the most relevant variable, dividing donors into three groups: one for MID_DON (mid-level donors), one for MJREG_MID (Mid value Spo/CG) and one for the rest of the subgroups. Subgroups in this donor category are heterogeneous and therefore is rational to expect that subgroup plays an essential role in this case. Further splits within subgroups included consideration of other variables such as Lifespan, First donation and Generation.

Table 22 *High-value nodes from the model for Mid and Major donors*

Node #	Count	Error	Variables							Avg AADLTV
			Subgroup	First Donation	Generation	Channel	Lifespan	Prod _SG	Donations first year	
4	131	0%	ACT_MAJ, MID_DON, MJREG_MID		BOOM, GenX, GenZ, MILL, SILE, Unk					840
5	4	25%	ACT_MAJ, MID_DON, MJREG_MID		ELDE					2,012
9	13	0%	TOP _MAJ				< 12 m			46,375
11	95	15%	FOR_MAJ, MID_HIGH, TOP _MAJ	<100, >1000			>= 12m			3,668

Nod e #	Count	Er- ror	Variables							Avg AADLTV
			Subgroup	First Dona- tion	Genera- tion	Chan- nel	Lifespan	Prod _SG	Dona- tions first year	
13	22	18%	FOR_MAJ, MID_HIGH, TOP _MAJ	100- 1,000			>= 12m		>= 2	1,745
15	38	0%	FOR_MAJ, MID_HIGH, TOP _MAJ				New			2,650
16	47	2%	MJREG_HIG H							2,924
33	26	42%	MID_DON				New		>= 2	1,308
40	17	12%	MJREG_MID	>1000				No		3,400
45	234	44%	MJREG_MID	100- 1,000		DIG, SPONT		No	>= 2	734

Behavioural factors were the most relevant to classify high-value donors in this Donor Category. The Subgroup is the most relevant factor in the classification, and other factors are also present with the variables First donation amount, Lifespan and Donations in the first year. In the group of Fundraising Methods, Channel and Prod_SG (the indication of single donations) are relevant to define two of the high-value nodes. Prod_CG and Prod_Vill_Spo were part of the model; however, these were not relevant in the classification of high-value nodes. Socio-demographic factors were not conclusive in the classification of high-value donors. Although Generation was part of the definition of high-value nodes, it was not determinant to differentiative high-value groups of donors. Income level did not appear in the classification of high- value nodes.

The following groups of high-value donors result from the nodes in Table 22:

- Active major donors, Mid value donors and Mid value Sponsor/Committed Givers
- Top Major donors with a lifespan lower than 12 months
- Former active major donors, Mid to high-value donors and Top major donors
    - New (note yet 12 months as a donor)
    - with lifespan over 12 months
        - first donation lower than € 100 or over € 1,000
        - first donation lower between € 100 and € 1,000 and more than two donations in the first year
- Mid to high-value Sponsor/Committed Givers

- Mid value Sponsor/Committed Givers without single donations
    - With the first gift over € 1,000
    - With the first gift between € 100 and € 1,000 via Digital or Other channels and at least two donations in the first year
    - Mid value donors with Lifespan in New who made two or more donations in the first year

The resulting groups reflect the heterogeneity of the Donor Types grouped under the category of Mid-major Donors. The subgroups Active major donors, Mid value donors, Mid value Sponsor/Committed Givers and Mid to high-value Sponsor/Committed Givers are present without almost any other conditions in the nodes, which means most donors in these groups are in these high-value nodes. Mid value Sponsor/Committed Givers in the end nodes do not have single donations. For Former active major donors, Mid to high-value donors and Top major donors the First donation and Lifespan were factors to consider; in these cases, higher first gifts have no other conditions and for lower first donations the regularity (2 or more donations) in the first year is a deciding factor.

4.3.2 Evaluation of the model for Mid and Major donors

The confusion matrix (Table 23) resulted in an overall accuracy of 89.6% for the model, with a positive predictive value of 75.3% for the classification of donors with class “High”. The 10-fold cross-validation performed with an accuracy of 88.5%. These results indicated the adequate performance and reliability of the model and rules out overfitting.

Table 23. *Confusion matrix and accuracy of models for Mid and Major donors*

Confusion Matrix		Predicted	
Sponsorship model		High	Low-mid
Observed	High	572	188
	Low-mid	219	2,942

Measures	Accuracy	Pos pred value	Neg pred value
	0.8962	0.7526	0.9307

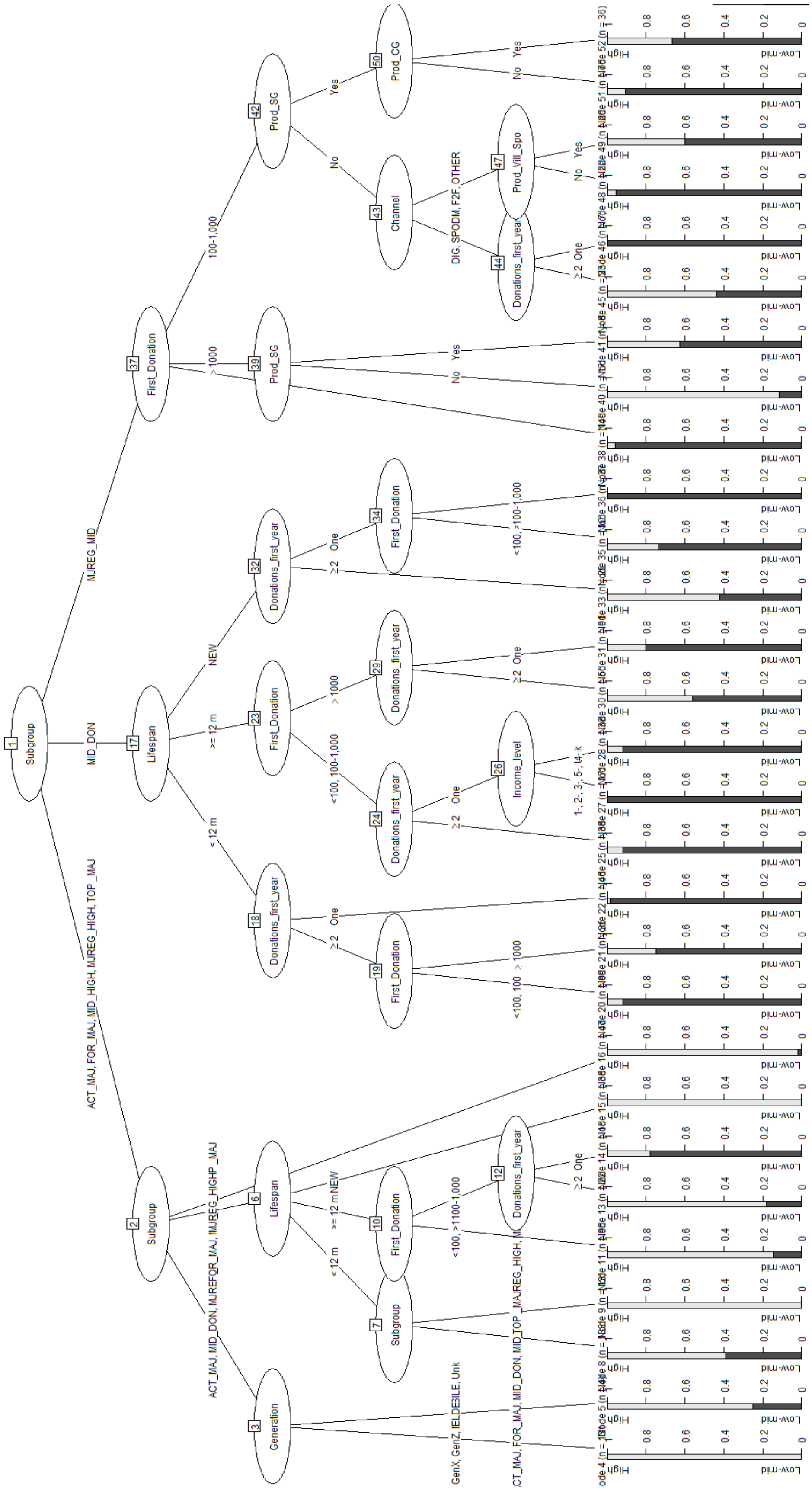

Figure 7. CHAID Model for Mid- and Major Donors

4.4 CHAID model for Single Givers

The dataset for Single Givers contained 261,400 donors classified in the different categories, as shown in Table 24. The categories for First donation was set by approximation using percentiles to divide the dataset into three groups of similar size, with the following categories: less than € 20, between € 20 and € 50 and higher than € 50.

Table 24. *Variables and categories for Single Givers*

Code	Count	%
Generation		
ELDE	3,623	1%
SILE	91,913	35%
BOOM	111,991	43%
GENX	40,306	15%
MILL	9,194	4%
GENZ	464	0%
UNK	3,909	1%
Channel		
DIG	22,184	8%
DM	125,256	48%
OTHER	1,141	0%
SPONT	112,819	43%
Income level		
1-	21,442	8%
2-	29,986	11%
3-	54,006	21%
4-	40,145	15%
5-	38,474	15%
UNK	77,347	30%

Code	Count	%
Gender		
M	100,500	38%
F	129,490	50%
UNK	31,410	12%
Donations in the first year		
One	218,629	84%
>= 2	42,771	16%
Lifespan		
>= 12 m	60,525	23%
< 12 m	165,854	63%
NEW	35,021	13%
First Donation		
<20	77,672	30%
20 to 50	92,736	35%
>50	90,992	35%

4.4.1 Results for Single Givers

The following formula was used to run the model with eight independent variables, as described in Table 12:

$$AALTV_{Class} \sim Gender + Generation + Income\ level + Subgroup + Channel + Lifespan + First\ Gift\ Amount + Donations\ first\ year$$

The initial model had 133 inner nodes and 185 terminal nodes. In this case, given the high number of observations, the CHAID arguments (see Table 13) were modified to select the desired minimum size of the terminal nodes and reduce complexity in the interpretation of the results. Two arguments were modified to refine the model: minsplit and minbucket (see Table 25) to reduce the number of nodes.

Table 25. *CHAID control arguments modified for Single Givers model*

Argument	Default	Value used
minsplit	20	1,000
minbucket	7	1,000

The resulting modified model selected for analyses had 43 inner nodes and 71 terminal nodes. *Figure 8* displays the plot of the model (see the Appendix for full details of all nodes), from which 17 correspond to high-value nodes, as described in Table 26. The model uses all input variables, meaning all are significant to the AADLTV Class, according to the Chi-square tests.

Behavioural variables had the most substantial influence in the AADLTV classes with First donation and Subgroup as the variables used for the first two splits. The results showed, therefore, the importance of the value of the first donations made by donors and the subgroups in the CHAID classification of Single Givers. Furthermore, Lifespan and Donations_first_year were common denominators to form high-value nodes. In the group of Fundraising Methods, Channel is a determinant in the high-value nodes for the new donors. Socio-demographic factors were not so relevant to identify the high-value donors; although the model found Gender, Generation and Income Level significant, these were not determining factors in the definition of high-value nodes and were not present in the leading splits that CHAID identified.

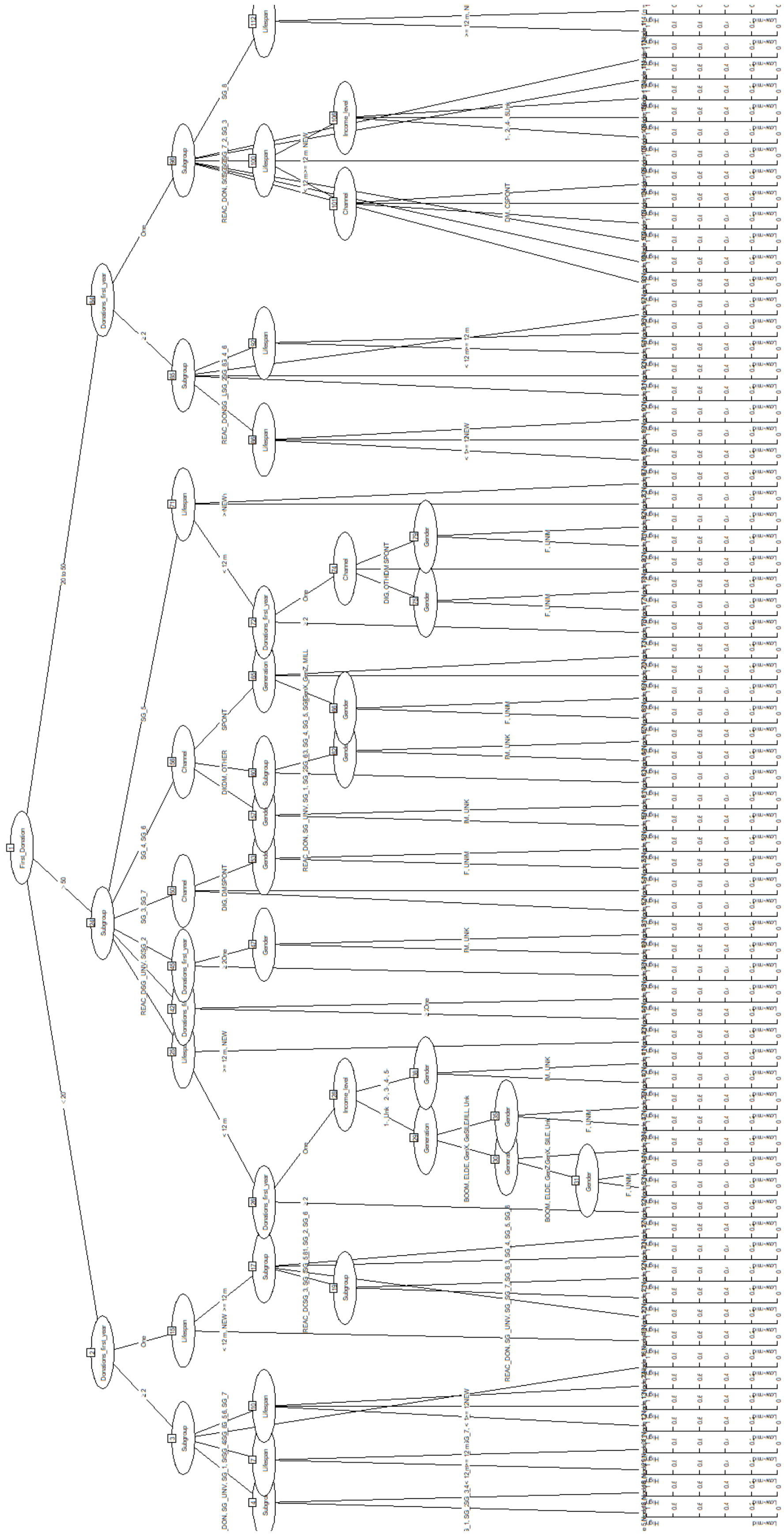

Figure 8. CHAID Model for Single Givers

Table 26. *High-value nodes from the model for Single Givers*

Node #	Count	Error	Variables						Avg AADLTV
			First Donation	Subgroup	Channel	Lifespan	Gender	Donations first year	
27	1,150	41%	> 50	REAC_DON, SG_8		< 12 m		>= 2	70.4
41	2,954	32%	> 50	REAC_DON, SG_8		>= 12m, New			81.0
43	1,413	2%	> 50	SG_UNV, SG_1				>= 2	159.9
44	1,935	7%	> 50	SG_UNV, SG_1				One	122.5
46	3,468	8%	> 50	SG_2				>= 2	107.4
48	1,809	26%	> 50	SG_2			F	One	76.3
49	2,420	21%	> 50	SG_2			M, UNK	One	87.9
73	1,785	5%	> 50	SG_5		< 12m		>= 2	130.5
76	1,339	38%	> 50	SG_5	DIG, OTHER	< 12m	F, UNK	One	71.9
77	1,484	33%	> 50	SG_5	DIG, OTHER	< 12m	M	One	80.0
80	7,587	44%	> 50	SG_5	SPONT	< 12m	F, UNK	One	61.1
81	5,080	39%	> 50	SG_5	SPONT	< 12m	M	One	66.6
82	2,728	4%	> 50	SG_5		>= 12m			138.9
83	9,779	0%	> 50	SG_5		New			106.3
88	1,567	38%	20 to 50	REAC_DON, SG_3, SG_5		>= 12m		>= 2	58.4
89	1,964	11%	20 to 50	REAC_DON, SG_3, SG_5		New			57.8
90	2,105	22%	20 to 50	SG _UNV, SG_1, SG_6				>= 2	78.2

There were no high-value nodes in the group of donors with First donations lower than € 20. There were three high-value nodes in the group of First donations between € 20 and € 50, and the remaining 14 were in the category of First donation higher than € 50, resulting in the following groups of high-value donors:

- First donation amount between € 20 and € 50
    - In subgroups REAC_DON (Reactivated donors), SG_3 (high frequency, low amount), SG_5 (New donors)
        - Lifespan >= 12 months and two or more donations in the first year
        - Lifespan as new

        - In subgroups SG _UNV (new unvalued), SG_1 (high frequency and amount), SG_6 (one-time donor) with two or more donations in the first year

- First donation amount over € 50
    - In subgroups REAC_DON (Reactivated donors), SG_8 (donor with bad postal address)
        - Lifespan < 12 months and two or more donations in the first year
        - Lifespan >12 months or New
    - In subgroups SG_UNV (new unvalued), SG_1 (high frequency and amount)
    - In subgroup SG_2 (low frequency, high amount)
    - In subgroup SG_5 (New donors)
        - Lifespan >= 12 months or New
        - Lifespan < 12
            - With two or more donations in the first year
            - With one donation in the first year
                - Via Digital or Other channels
                - Via Spontaneous gifts
    - In subgroups REAC_DON (Reactivated donors), SG_3 (high frequency low amount), SG_5 (New donors)
        - Lifespan New
        - Lifespan >= 12 months and two or more donations in the first year
    - In subgroups SG_UNV (new unvalued), SG_1 (high frequency and amount) and SG_6 (one-time donors), with two or more donations in the first year

The groups of Single Givers with high value have attributes in common, mostly behavioural. A Subgroup that appears in various nodes is SG_1 (high frequency and amount), indicating as expected that the frequency of donations and higher amounts can drive donor value. Some of the groups contain reactivated donors (donors previously lost that donated after a long period), similarly for medium and high initial donations, indicating that the efforts to regain these donors resulted in a higher value. The presence of subgroup SG_3 (high frequency, low amount) suggests that lower donations can lead to high donor value if donation frequency is high. In contrast, subgroup SG_2 (low frequency, high amount) indicated that high donor values could be achieved with lower frequency if the amounts donated are high. These two factors, amount and frequency, complement each other.

Various high-value nodes refer to donors that are newer or unvalued. These findings need to be addressed carefully, as the lower lifetime opportunity of these donors can skew the results, and the longer-term behaviour may vary, affecting the value of these donor groups. One thing to highlight from this group is the differentiation between channels, in which Digital and Spontaneous gifts are the two channels bringing these new high-value donors.

4.4.2 Evaluation of the model for Single Givers

The confusion matrix for the prediction model (Table 27) results in an accuracy of 90.7% and the positive predictive value for Single Givers in the High AADLTV Class is 78%. These are indicators of reliable performance of the model, the highest among all the four models. The 10-fold cross-validation resulted in an accuracy of 90.9%, indicating that the model has an adequate prediction performance and there is no overfitting.

Table 27. *Confusion matrix and accuracy of models for Single Givers*

Confusion Matrix Sponsorship model		Predicted	
		High	Low-mid
Observed	High	40,561	11,265
	Low-mid	13,028	197,257

Measures	Accuracy	Pos pred value	Neg pred value
	0.9073	0.7826	0.9380

5 Discussion

The analyses of the models in each Donor Category have their particularities and different results. However, there are common points that lead to general findings to respond to the research question about the factors that influence donor value and lead to high-value donors, as well as to the evaluation of the proposed model and its application. This chapter provides a discussion compiling the results for all Donor Categories, limitations and recommendations for SOS CV and future research in the field.

5.1 Methodological considerations

The selection of the CHAID model allowed to address the research questions in a way that other statistical methods would not. The criteria for selecting the CHAID technique is that it creates a predictive model that determines how diverse factors (behavioural, sociodemographic and methodological) merge and interact with each other to find which factors influence the value of donors and what are the common factors within high-value donors, addressing the objectives of the research in an inclusive way. The limitation of using CHAID is the need to categorise the outcome variable, as this limits the possibility to find more precise relations between the variables, such as covariances or correlations.

Internal validity of the study is provided by a comprehensive model that considers various factors found in the theoretical foundation. Limitations to internal validity arise from the assumption of behaviour as past actions, without considering the underlying motivational and perceptional components or the potential influence of organisational practices; however, this consideration deserves a detailed analysis on its own and is therefore not in the scope of the present study.

A strength of the model is the large dataset available to study that constitutes an extensive and representative sample, indicating that the results can be generalised to the total population of donors of SOS CV. However, the results can not be generalised to other settings such as donors in other organisations, or geographically to the country or other countries. For changes in configurations, the study should be replicated using the scenario-specific data that applies.

The limitation related to the dataset is the high number of unknown fields in the socio-demographic variables, which may lead to an underappreciation of these variables and their influence in the results, especially for Single and Mid-major donors.

The evaluation of the models showed that the levels of the overall accuracy of classification are adequate for all the models (over 80%), with a higher degree for Single Givers and Mid-major donors. However, the confusion matrixes showed a lower classification accuracy for the category “High”, which means that the incorrect classification of high-value donors was higher than for the “Low-mid” class. The results of the positive predictive value indicate that, even if the model is acceptable in its overall accuracy, there is room for improvement in the predictive performance of the models in respect to classifying donors in the class High. Future research and applications can refine the variables and respective categories in an attempt to overcome this limitation. The findings of the evaluation show that there is not overfitting in any of the models.

One factor that was present in the literature that the study does not consider is the incorporation of costs in the analysis of lifetime value. As described in the literature review, various authors suggest using a net version of the lifetime value measures, that includes the costs of attracting and maintaining donors, such as costs incurred in the acquisition channels and the cost of ongoing communication with existing donors. The lack of this element implies a limitation of the study to analyse the value in terms of the net revenue generated by donors and comparing between fundraising methods (channels and product), as the costs incurred for each one could potentially result in different results. However, in the absence of structured cost data at a donors level, the gross AADLTV measure is an adequate alternative that provides relevant insights to the research question.

5.2 Results and recommendations

The compilation of results from the four models leads to the conclusion that there are, indeed, factors that influence donor value and common traits within the high-value donors; these attributes are mainly behavioural and to a lesser extent are related to the fundraising methods and donor characteristics.

Behavioural variables are the most influential ones in the determination of groups of donors with a high lifetime value across donor categories. In the four cases, a behavioural factor defined the first split of the CHAID model, and these were the most commonly identified in the definition of high-value nodes. First Donation, in particular, is a variable that is highly influential in the identification of high-value donors across models.

The second set of factors in importance were the Fundraising methods, in which channel played a significant role in the two categories with regular donations (Committed Giving and Sponsorships). The third factor, Donor Characteristics, did not play an essential role as a common characteristic of high-value donors and for specific donor categories, certain variables such as Gender and Income Level were not considered as significant by the model.

The conclusions regarding the factors and their interaction, as well as the importance of behavioural factors in the value of donors, needs careful interpretation. Past behaviour of donors can mask the effect of other factors, being partially driven by past organisational strategies with influence in the results, for example, which donors to target, pricing strategies that define the initial donations or channels that used in the past. In contrast, these are factors that in future can also be controlled by the organisation, and therefore the results of this study provide robust insights and indications for factors that the organisation can consider adapting in the fundraising strategies and efforts. Furthermore, behavioural factors were addressed as an independent variable and analysed in the frame of past behaviour, given that donor value was the outcome of interest in this study. According to the literature review, there are diverse approaches to behaviour that treat it as an outcome variable, influenced by other factors.

Given that the results showed the importance of behavioural factors in the determination of high donor value, a recommendation for further research and practical application is to study giving behaviour comprehensibly as a dependent variable that is influenced by, for example, socio-demographic, motivational or perceptual factors. One possibility is to use theories of behavioural economics, in the frame of decision, applied to techniques for raising funds, to study variables such as the decision to give to charity or how much to donate. Another alternative is to build a comprehensive model of all factors that influence behaviour, similar to the one suggested by Sargeant (1999), applied to new donors of SOS CV and study the evolution over time. A second alternative is a variation of the study in which behaviour is a mediator variable that is influenced by sociodemographics and fundraising methods and that influences donor value directly, rather than as an independent variable.

The following recommendations for SOS CV represent proposals for practical applications of the results: the development of tests where behavioural factors are controlled (e.g. suggested first donation amount) for specific channels, products or target audiences, to compare the results and the projected long-term impact the changes has in the value of donors; the implementation of customised donor journeys to specific donors' groups that allow for feedback loops that update the journey regularly according to past behaviour. Techniques and tools can be found in the fields of data science and machine learning, considering the capabilities and resources of the organisation.

Regarding the results related to fundraising methods, the results differ for individual variables. Channel is a variable that is present in the high-value nodes for all models, although it is not determinant for high-value donors in the case of Mid-major Donors. Digital and Spontaneous Gift are the typical channels for high-value donors, whereas Face to Face, one of the largest channels for acquisition of regular donors, is only present once in the determination of high-value donors, suggesting that a more in-depth analysis of the channel and its performance with regards to donor value could be of interest for the organisation. The Product was a relevant variable in the case of Prod_SG, when related to regular donors (Sponsors and Committed Givers), i.e. to define groups of regular donors is influenced by whether or not they gave single donations on top of their regular commitment, in most cases resulting in high-value groups that did not make single donations.

Results related to socio-demographic variables indicate that these do not have a strong significance when defining groups of high-value donors. Gender played a role in the case of Committed Givers, where specific groups of high-value donors were mostly male; however, for the majority of the nodes, there was no gender differentiation.

6 Conclusion

This Master Thesis aimed to identify the factors that influence the value of donors to SOS CV and which of those factors are common to high-value donors. Based on the statistical analysis on a sample of 292,478 donors to SOS CV using the CHAID methodology on four categories of donors (Sponsors, Committed Givers, Mid-Major Donors and Single Givers), it was concluded that behavioural factors are the more relevant to segment donors according to their value and the determination of high-value donors, whereas fundraising methods and socio-demographic traits are relevant, however, of relative importance for specific groups of donors.

The analysis for the four donor categories showed that behavioural variables such as the subgroup, the lifespan, the number of donations in the first year and, in particular, the amount of the first donation, are the most relevant to start the categorisation of donors according to the donor class and that these are determining factors for high-value donors. From the variables related to Fundraising methods, the channel plays a deciding role in forming high-value nodes, except for Mid-major Donors. Digital and Spontaneous gifts were the channels frequently present in high-value nodes, whereas Face to Face was not. The variable Product was found relevant in selected cases for Single Donations. The socio-demographic factors had only limited influence in high-value nodes, where Gender was present in specific cases.

The CHAID methodology was selected to allow for the analysis of how diverse behavioural factors, sociodemographic traits and fundraising methods interact to form donor segments and identify which of those groups correspond to donors with high-value. The incorporation of costs into the analysis of donor value is suggested for future studies, as it is one of the limitations of the present study. The models built for the four donor categories show high classification accuracy; however, with lower positive predictive value for the class High. It is suggested that variables and categorisation are revised and refined in future research to improve this aspect.

This Master Thesis contributes to research in the area of the lifetime value of individual donors to charity and broadens the scope of the study of factors that influence it. By using a decision tree methodology, this study provides insights as to how various variables interact to explain and predict the value of charity donors, an approach that has not been widely used in the field of donor value and giving behaviour. The study is based on a practical application to the concept of donor lifetime value through a case study, providing a framework that could be replicated in future research for different settings, for example, other countries, charities or groups of donors.

Fields recommended for future research include a more in-depth study in the area of donor behaviour; some of the alternatives in this field are the comprehensive modelling of factors that lead to giving behaviour and the application of behavioural economics in decision making for charitable giving.

References

Abdolvand, N., Albadvi, A., & Koosha, H. (2014). Customer Lifetime Value: Literature Scoping Map, and an Agenda for Future Research. *International Journal of Management Perspective, 1*(3), 41-59. Retrieved June 2, 2019, from https://www.researchgate.net/profile/Neda_Abdolvand/publication/277843944_Customer_Lifetime_Value_Literature_Scoping_Map_and_an_Agenda_for_Future_Research/links/557ebe3308aeb61eae258626.pdf

Aldrich, T. M. (2000). How much are new donors worth? Making donor recruitment investment decisions based on lifetime value analysis. *International Journal of Nonprofit and Voluntary Sector Marketing, 5*(1), 81-89. Retrieved from https://doi-org.uaccess.univie.ac.at/10.1002/nvsm.99

Ambrogetti, F. (2016). *Emotionraising: How to astonish, disturb, seduce and convince the brain to support good causes.* Amazon Digital Services LLC.

Ariely, D., Bracha, A., & Meier, S. (2009). Doing Good or Doing Well? Image Motivation and Monetary Incentives in Behaving Prosocially. *American Economic Review*, 544-55. doi:10.1257/aer.99.1.544

BBC News. (2018). *Oxfam Haiti allegations: How the scandal unfolded.* London: BBC. Retrieved from https://www.bbc.com/news/uk-43112200

Bennett, R. (2006). Predicting the Lifetime Durations of Donors to Charities. *Journal of Nonprofit & Public Sector Marketing, 15*(1-2), 45-67. Retrieved 05 25, 2019, from https://doi.org/10.1300/J054v15n01_03

Berg, L. (2014). Who benefits from behavioural economics? *Economic Analysis and Policy*, 221-232. doi:https://doi.org/10.1016/j.eap.2014.06.001

Chang, C. -T. (2007). Who gives what to charity: characteristics affecting donation behavior. *Social Behavior and Personality: an international journal, 35*(9), 1173-1180. Retrieved from https://search-proquest-com.uaccess.univie.ac.at/docview/209900985/fulltextPDF/FB0993E95DCB4767PQ/1?accountid=14682

Charities Aid Foundation. (2018). *CAF World Giving Index 2018. A global view of giving trends.* Charities Aid Foundation. Retrieved May 12, 2019, from https://www.cafonline.org/docs/default-source/about-us-publications/caf_wgi2018_report_webnopw_2379a_261018.pdf

de Ville, B. (2006). *Decision Trees for Business Intelligence and Data Mining: Using SAS.* Cary, NC: SAS Institute Inc.

de Vries, N. J., Reis, R., & Moscato, P. (2015). Clustering Consumers Based on Trust, Confidence and Giving Behaviour: Data- Driven Model Building for Charitable Involvement in the Australian Not-For-Profit Sector. *PLoS ONE, 10*(4). doi:DOI:10.1371/journal.pone.012213

Díaz-Pérez, F., & Bethencourt-Cejas, M. (2016). CHAID algorithm as an appropriate analytical method for tourismmarket segmentation. *Journal of Destination Marketing & Management*, *5*, 275-282. Retrieved from https://doi.org/10.1016/j.jdmm.2016.01.006

Durango-Cohen, E. J., Torres, R. L., & Durango-Cohen, P. L. (2013). Donor Segmentation: When Summary Statistics Don't Tell the Whole Story. *Journal of Interactive Marketing, 27*, 172-184. Retrieved May 20, 2019, from https://doi.org/10.1016/j.intmar.2013.04.002

Harvey, J., & McCrohan, K. (1998). Fundraising costs societal implications for philanthropies and their supporters. *Business and Society*, *27*(1), 15-22. doi:https://doi.org/10.1177/000765038802700103

Hochman, G., & Ariely, D. (2015). Behavioral Economics. *Emerging Trends in the Social and Behavioral Sciences: An Interdisciplinary, Searchable, and Linkable Resource*. doi:https://doi.org/10.1002/9781118900772.etrds0022

International Fundraising Leadership Forum. (2019). *Peer Review Key Findings 2019.* The Ressource Alliance. Retrieved November 14, 2019, from https://www.slideshare.net/ResourceAlliance/ifl-forum-ifc-2019

Kass, V. G. (1980). An Exploratory Technique for Investigating Large Quantities of. *Journal of the Royal Statistical society, 29*(2), 119-127. Retrieved from https://www4.stat.ncsu.edu/~dickey/analytics/datamine/Reference%20Papers/kass80.pdf

Kuhn, M., Chow, F., & Wickham, H. (2019). rsample: General Resampling Infrastructure. *R package version 0.0.5*. Retrieved from https://CRAN.R-project.org/package=rsample

Kuhn, M., contributions from Wing, J., Weston, S., Williams, A., Keefer, C., Engelhardt, A., . . . Hunt, T. (2019). caret: Classification and Regression Training. *R package version 6.0-84*. Retrieved from https://CRAN.R-project.org/package=caret

Lee, Y.-K., & Chang, C.-T. (2008, March). Intrinsic or Extrinsic? Determinants Affecting Donation Behaviors. *International Journal of Educational Advancement, 8*(1), 13-24. doi:10.1057/ijea.2008.2

Magson, N. (1999). Donors: how much do they give in a lifetime? *International Journal of Nonprofit and Voluntary Sector Marketing, 4*(1), 11-25. doi: https://doi.org/10.1002/nvsm.53

Masters, T. (2000). Deciding to recruit only donors with high lifetime values. *International Journal of Nonprofit and Voluntary Sector Marketing , 5*(3), 241-247. Retrieved May 25, 2019, from https://doi-org.uaccess.univie.ac.at/10.1002/nvsm.116

Milanović, M., & Stamenković, M. (2016). CHAID Decision Tree: Methodological Frame. *Economic Themes, 54*(4), 563-586. doi:DOI 10.1515/ethemes-2016-0029

Milanovic, M., & Stamenkovic, M. (2016). CHAID Decision Tree: Methodological Frame and Application. *Economic Themes, 54*(4), 563-586. doi:10.1515/ethemes-2016-0029

Najev Čačija, L. (2013). Fundraising in the context of nonprofit strategic marketing: Toward a conceptual model. *Management : Journal of Contemporary Management Issues, 8*(1), 59-78. Retrieved from https://hrcak.srce.hr/104238

Nonprofit Tech for Good. (2018). *2018 Global Trends in Giving Report.* Public Interest Registry. Retrieved June 2019, 08, from https://givingreport.ngo/

Noor, A. H., Isa, N. A., Irpan, H. M., Bahrom, H. B., Salleh, A. B., & Ridzuan, A. R. (2015). Characteristic Affecting Charitable Donations Behavior: Empirical Evidence from Malaysia. *Procedia Economics and Finance, 31*, 563-572.

R Core Team. (2019). R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing. Retrieved from https://www.R-project.org/

RStudio Team. (2018). RStudio: Integrated Development Environment for R. Boston, MA: RStudio, Inc. Retrieved from http://www.rstudio.com/

Rupp, C., Kern, S., & Helmig, B. (2014). Segmenting nonprofit stakeholders to enable successful relationship marketing: a review. *International Journal of Nonprofit and Voluntary Sector Marketing, 19*, 76-91. doi:10.1002/nvsm.1492

Sargeant, A. (1999). Charitable Giving: Towards a Model of Donor Behaviour. *Journal of Marketing Management , 15*(4), 215-238. doi:https://doi.org/10.1362/026725799784870351

Sargeant, A. (2001). Using Donor Lifetime Value to Inform Fundraising Strategy. *Nonprofit Management and Leadership, 12*(1), 25-38. doi:10.1002/nml.12103

Sargeant, A., & McKenzie, J. (1999). The lifetime value of donors: Gaining insight through CHAID. *Fund Raising Management;, 30*(1), 22-27. Retrieved May 25, 2019, from https://search-proquest-com.uaccess.univie.ac.at/docview/195880286?accountid=14682

Sargeant, A., Ford, J. B., & West, D. C. (2006). Perceptual determinants of nonprofit giving behavior. *Journal of Business Research, 59*, 155-165. doi:10.1016/j.jbusres.2005.04.006

Schlegelmilch, B. B., Diamantopoulos, A., & Love, A. (1997). Characteristics affecting charitable donations: empirical evidence from Britain. *Journal of Marketing Practice, 3*(1), 14-32. Retrieved from https://search-proquest-com.uaccess.univie.ac.at/docview/211395326?rfr_id=info%3Axri%2Fsid%3Aprimo

Shelley, L., & Polonsky, J. (2002). Do charitable causes need to segmenttheir current donor base on demographicfactors?: An Australian examination. *International Journal of Nonprofitand Voluntary Sector Marketing, 7*(1), 19-29. doi: https://doi-org.uaccess.univie.ac.at/10.1002/nvsm.164

Snipes, R. L., & Oswald, S. L. (2010). Charitable giving to not-for-profit organizations: factors affecting donations to non-profit organizations. *Innovative Marketing*, *6*(1), 73-80. Retrieved 05 13, 2019, from https://businessperspectives.org/images/pdf/applications/publishing/templates/article/assets/3066/im_en_2010_01_Snipes.pdf

SOS Children's Villages International. (2017). *International Annual Report 2016.* Innsbruck: SOS Children's Villages International. Retrieved from https://www.sos-childrensvillages.org/getmedia/b0a2f90f-8193-46a1-90f7-11162046d660/SOS_International-Annual-Report-2016.pdf

SOS Children's Villages International. (2017). *Who we are*. Retrieved May 27, 2019, from SOS Childre's Villages International: https://www.sos-childrensvillages.org/who-we-are

SOS Children's Villages International. (2018). *International Annual Report 2017.* Innsbruck: SOS Children's Villages International. Retrieved from https://www.sos-childrensvillages.org/getmedia/1d7ebfd6-8ed6-43c1-9892-393a0f42355c/SOS_CHILDRENS_VILLAGES_ANNUAL_REPORT_2017_WEB.pdf

SOS Children's Villages International. (2019). *Internal Strategy Survey Results.* Internal report: unpublished.

SOS Children's Villages International. (2019). *International Annual Report 2018.* Innsbruck: SOS Children’s Villages International. Retrieved from https://www.sos-childrensvillages.org/getmedia/7ce55581-adc0-4c7e-97ab-76beda7e94e2/SOS-ANNUAL-REPORT-2018_WEB.pdf

SOS Children's Villages International. (2019). *International Annual Report 2018.* Innsbruck: SOS Children’s Villages International. Retrieved from https://www.sos-childrensvillages.org/getmedia/7ce55581-adc0-4c7e-97ab-76beda7e94e2/SOS-ANNUAL-REPORT-2018_WEB.pdf

SOS Children's Villages International. (2019). *Strategy Survey Results.* Internal report: unpublished.

Spears, L. A. (2002). Persuasive Techniques Used In Fundraising Messages. *J. Technical Writing And Communication, 32*(3), 245-265. Retrieved May 18, 2019, from https://doi.org/10.2190/BE4V-QJNC-Q97H-DFXN

Srnka, K. J., Grohs, R., & Eckler, I. (2003). Increasing Fundraising Efficiency by Segmenting Donors. *Australasian Marketing Journal*, *11*(1), 70-86.

The FoRt Student Project Team. (2015). CHAID: CHi-squared Automated Interaction Detection. *R package version 0.1-2*.

Watson, B. (2015). The origins of international child sponsorship. *Development in practice*, *25*(6), 867-879. doi:10.1080/09614524.2015.1064362

Watson, B., & Clarke, M. (2014). Introduction to Key Issues in Child Sponsorship. *Child Sponsorship*, 1-17. doi:https://doi.org/10.1057/9781137309600_1

World Bank. (2018). *World Development Indicators database.* World Bank. Retrieved May 05, 2019, from https://data.worldbank.org/

Yörük, B. K. (2009). How responsive are charitable donors to requests to give? *Journal of Public Economics, 93*, 1111-1117. doi:10.1016/j.jpubeco.2009.06.001

Yörük, B. K. (2012). Do Charitable Solicitations Matter? A Comparative Analysis of Fundraising Methods. *Fiscal Studies*, *33*(4), 467–487.

Appendix

Detailed CHAID Model results for the four Donor Categories

This appendix contains the details from printing the CHAID model for each Donor Category, aiming at providing clarity and further details of interest for the reader.

CHAID Model for Sponsors

```
Model formula:
AALTV_Class ~ Prod_SG + Gender + Generation + Income_level +
    Subgroup + Channel + Lifespan + First_Donation + Donations_first_year

Fitted party:
[1] root
|   [2] First_Donation <= 40
|   |   [3] Prod_SG in No
|   |   |   [4] Channel in DIG
|   |   |   |   [5] Lifespan < 12 m, NEW: Low-mid (n = 351, err = 1.7%)
|   |   |   |   [6] Lifespan >= 12 m
|   |   |   |   |   [7] Subgroup in GIFT_SPO, IV_SPO, NV_SPO: Low-mid (n =
1413, err = 8.3%)
|   |   |   |   |   [8] Subgroup in IC_SPO
|   |   |   |   |   |   [9] Generation in BOOM, ELDE, GenX, GenZ, SILE, Unk
: Low-mid (n = 2081, err = 17.4%)
|   |   |   |   |   |   [10] Generation in MILL: Low-mid (n = 1345, err = 1
1.4%)
|   |   |   [11] Channel in DM, F2F
|   |   |   |   [12] Subgroup in GIFT_SPO, IC_SPO, NV_SPO
|   |   |   |   |   [13] Generation in BOOM, ELDE, GenX, GenZ, MILL, Unk: L
ow-mid (n = 1032, err = 5.6%)
|   |   |   |   |   [14] Generation in SILE: Low-mid (n = 14, err = 28.6%)
|   |   |   |   [15] Subgroup in IV_SPO
|   |   |   |   |   [16] Lifespan < 12 m, >= 12 m: Low-mid (n = 1194, err =
3.7%)
|   |   |   |   |   [17] Lifespan in NEW: Low-mid (n = 159, err = 0.0%)
|   |   |   [18] Channel in OTHER, SPONT
|   |   |   |   [19] Subgroup in GIFT_SPO, IC_SPO, IV_SPO: Low-mid (n = 607
, err = 19.3%)
|   |   |   |   [20] Subgroup in NV_SPO: Low-mid (n = 112, err = 5.4%)
|   |   [21] Prod_SG in Yes
|   |   |   [22] Donations_first_year >= 2
|   |   |   |   [23] Subgroup in GIFT_SPO, IC_SPO: Low-mid (n = 173, err =
46.8%)
|   |   |   |   [24] Subgroup in IV_SPO, NV_SPO: Low-mid (n = 150, err = 28
.0%)
|   |   |   [25] Donations_first_year in One: Low-mid (n = 23, err = 0.0%)
|   [26] First_Donation > 40
|   |   [27] Channel in DIG
|   |   |   [28] Donations_first_year >= 2
|   |   |   |   [29] Subgroup in GIFT_SPO, NV_SPO
|   |   |   |   |   [30] Lifespan < 12 m, >= 12 m: High (n = 156, err = 47.
4%)
|   |   |   |   |   [31] Lifespan in NEW: High (n = 14, err = 7.1%)
|   |   |   |   [32] Subgroup in IC_SPO
|   |   |   |   |   [33] Prod_SG in No
|   |   |   |   |   |   [34] Income_level in 1-, 2-, 3-, 4-, 5-: High (n =
590, err = 23.1%)
|   |   |   |   |   |   [35] Income_level in Unk: Low-mid (n = 23, err = 47
.8%)
```

```
|   |   |   |   |   [36] Prod_SG in Yes: High (n = 67, err = 40.3%)
|   |   |   |   [37] Subgroup in IV_SPO
|   |   |   |   |   [38] Lifespan < 12 m, >= 12 m: High (n = 148, err = 48.
6%)
|   |   |   |   |   [39] Lifespan in NEW: High (n = 12, err = 16.7%)
|   |   |   [40] Donations_first_year in One: Low-mid (n = 72, err = 18.1%)
|   |   [41] Channel in DM
|   |   |   [42] Prod_SG in No: High (n = 12, err = 50.0%)
|   |   |   [43] Prod_SG in Yes: Low-mid (n = 68, err = 11.8%)
|   |   [44] Channel in F2F
|   |   |   [45] Lifespan < 12 m, NEW: Low-mid (n = 107, err = 9.3%)
|   |   |   [46] Lifespan >= 12 m: Low-mid (n = 617, err = 49.4%)
|   |   [47] Channel in OTHER, SPONT
|   |   |   [48] Prod_SG in No
|   |   |   |   [49] Lifespan < 12 m: Low-mid (n = 4, err = 0.0%)
|   |   |   |   [50] Lifespan >= 12 m, NEW: High (n = 203, err = 37.4%)
|   |   |   [51] Prod_SG in Yes: Low-mid (n = 84, err = 25.0%)

Number of inner nodes:    23
Number of terminal nodes: 28
```

CHAID Model for Committed Givers

```
Model formula:
AALTV_Class ~ Prod_SG + Gender + Generation + Income_level + 
    Subgroup + Channel + First_Donation + Donations_first_year + 
    Lifespan

Fitted party:
[1] root
|   [2] First_Donation <20
|   |   [3] Lifespan < 12 m: Low-mid (n = 128, err = 27.3%)
|   |   [4] Lifespan >= 12 m
|   |   |   [5] Channel in DIG: Low-mid (n = 1205, err = 10.6%)
|   |   |   [6] Channel in DM
|   |   |   |   [7] Prod_SG in No: Low-mid (n = 33, err = 15.2%)
|   |   |   |   [8] Prod_SG in Yes: Low-mid (n = 391, err = 4.3%)
|   |   |   [9] Channel in F2F, SPONT: Low-mid (n = 3339, err = 14.0%)
|   |   |   [10] Channel in OTHER: Low-mid (n = 59, err = 25.4%)
|   |   [11] Lifespan in NEW
|   |   |   [12] Prod_SG in No: Low-mid (n = 980, err = 1.3%)
|   |   |   [13] Prod_SG in Yes: Low-mid (n = 33, err = 6.1%)
|   [14] First_Donation >50
|   |   [15] Donations_first_year >= 2
|   |   |   [16] Channel in DIG, OTHER
|   |   |   |   [17] Lifespan < 12 m, >= 12 m: Low-mid (n = 487, err = 44.6
%)
|   |   |   |   [18] Lifespan in NEW
|   |   |   |   |   [19] Subgroup in CG, CG FULL: High (n = 56, err = 23.2%
)
|   |   |   |   |   [20] Subgroup in CG SUPP: Low-mid (n = 17, err = 35.3%)
|   |   |   [21] Channel in DM
|   |   |   |   [22] Lifespan < 12 m: Low-mid (n = 13, err = 0.0%)
|   |   |   |   [23] Lifespan >= 12 m, NEW
|   |   |   |   |   [24] Gender in F, M: Low-mid (n = 603, err = 35.0%)
|   |   |   |   |   [25] Gender in UNK: Low-mid (n = 39, err = 15.4%)
|   |   |   [26] Channel in F2F
|   |   |   |   [27] Lifespan < 12 m
|   |   |   |   |   [28] Gender in F, UNK: Low-mid (n = 30, err = 26.7%)
|   |   |   |   |   [29] Gender in M: High (n = 32, err = 34.4%)
|   |   |   |   [30] Lifespan >= 12 m, NEW
|   |   |   |   |   [31] Generation in BOOM, ELDE, GenX, GenZ, SILE, Unk: L
ow-mid (n = 730, err = 26.7%)
|   |   |   |   |   [32] Generation in MILL: Low-mid (n = 840, err = 19.2%)
|   |   |   [33] Channel in SPONT
```

```
|   |   |   |   [34] Lifespan < 12 m, NEW: High (n = 42, err = 40.5%)
|   |   |   |   [35] Lifespan >= 12 m: Low-mid (n = 285, err = 39.3%)
|   |   [36] Donations_first_year in One
|   |   |   [37] Lifespan < 12 m
|   |   |   |   [38] Prod_SG in No: Low-mid (n = 59, err = 42.4%)
|   |   |   |   [39] Prod_SG in Yes: Low-mid (n = 11, err = 0.0%)
|   |   |   [40] Lifespan >= 12 m
|   |   |   |   [41] Channel in DIG, SPONT: Low-mid (n = 350, err = 19.7%)
|   |   |   |   [42] Channel in DM, OTHER: Low-mid (n = 370, err = 11.9%)
|   |   |   |   [43] Channel in F2F
|   |   |   |   |   [44] Prod_SG in No: Low-mid (n = 571, err = 6.5%)
|   |   |   |   |   [45] Prod_SG in Yes: Low-mid (n = 11, err = 27.3%)
|   |   |   [46] Lifespan in NEW: Low-mid (n = 251, err = 19.9%)
|   [47] First_Donation in 20 to 50
|   |   [48] Channel in DIG, OTHER
|   |   |   [49] Donations_first_year >= 2
|   |   |   |   [50] Prod_SG in No
|   |   |   |   |   [51] Subgroup in CG, CG FULL
|   |   |   |   |   |   [52] Lifespan < 12 m, >= 12 m: High (n = 597, err =
29.1%)
|   |   |   |   |   |   [53] Lifespan in NEW: High (n = 91, err = 9.9%)
|   |   |   |   |   [54] Subgroup in CG SUPP: High (n = 126, err = 11.1%)
|   |   |   |   [55] Prod_SG in Yes
|   |   |   |   |   [56] Gender in F: Low-mid (n = 59, err = 32.2%)
|   |   |   |   |   [57] Gender in M, UNK: High (n = 91, err = 44.0%)
|   |   |   [58] Donations_first_year in One: Low-mid (n = 52, err = 1.9%)
|   |   [59] Channel in DM
|   |   |   [60] Prod_SG in No: High (n = 16, err = 43.8%)
|   |   |   [61] Prod_SG in Yes
|   |   |   |   [62] Donations_first_year >= 2: Low-mid (n = 554, err = 11.
9%)
|   |   |   |   [63] Donations_first_year in One: Low-mid (n = 286, err = 1
.7%)
|   |   [64] Channel in F2F
|   |   |   [65] Generation in BOOM, ELDE, SILE, Unk: Low-mid (n = 243, err
= 41.2%)
|   |   |   [66] Generation in GenX, GenZ, MILL
|   |   |   |   [67] Subgroup in CG, CG FULL: Low-mid (n = 2224, err = 20.8
%)
|   |   |   |   [68] Subgroup in CG SUPP: Low-mid (n = 30, err = 46.7%)
|   |   [69] Channel in SPONT
|   |   |   [70] Prod_SG in No
|   |   |   |   [71] Subgroup in CG, CG FULL: High (n = 30, err = 50.0%)
|   |   |   |   [72] Subgroup in CG SUPP: High (n = 75, err = 25.3%)
|   |   |   [73] Prod_SG in Yes
|   |   |   |   [74] Donations_first_year >= 2
|   |   |   |   |   [75] Gender in F, UNK: Low-mid (n = 77, err = 26.0%)
|   |   |   |   |   [76] Gender in M: High (n = 47, err = 48.9%)
|   |   |   |   [77] Donations_first_year in One: Low-mid (n = 52, err = 1.
9%)

Number of inner nodes:    33
Number of terminal nodes: 44
```

CHAID Model for Mid-major donors

```
Model formula:
AALTV_Class ~ Prod_SG + Prod_CG + Prod_Child_Spo + Prod_Vill_Spo +
    Gender + Generation + Income_level + Subgroup + Channel +
    Lifespan + First_Donation + Donations_first_year

Fitted party:
[1] root
|   [2] Subgroup in ACT_MAJ, FOR_MAJ, MID_HIGH, MJREG_HIGH, TOP _MAJ
|   |   [3] Subgroup in ACT_MAJ, MID_DON, MJREG_MID
```

```
|   |   |   [4] Generation in BOOM, GenX, GenZ, MILL, SILE, Unk: High (n =
131, err = 0.0%)
|   |   |   [5] Generation in ELDE: High (n = 4, err = 25.0%)
|   |   [6] Subgroup in FOR_MAJ, MID_HIGH, TOP _MAJ
|   |   |   [7] Lifespan < 12 m
|   |   |   |   [8] Subgroup in ACT_MAJ, FOR_MAJ, MID_DON, MID_HIGH, MJREG_
HIGH, MJREG_MID: High (n = 133, err = 39.1%)
|   |   |   |   [9] Subgroup in TOP _MAJ: High (n = 13, err = 0.0%)
|   |   |   [10] Lifespan >= 12 m
|   |   |   |   [11] First_Donation <100, >1000: High (n = 95, err = 14.7%)
|   |   |   |   [12] First_Donation in 100-1,000
|   |   |   |   |   [13] Donations_first_year >= 2: High (n = 22, err = 18.
2%)
|   |   |   |   |   [14] Donations_first_year in One: Low-mid (n = 18, err
= 22.2%)
|   |   |   [15] Lifespan in NEW: High (n = 38, err = 0.0%)
|   |   [16] Subgroup in MJREG_HIGH: High (n = 47, err = 2.1%)
|   [17] Subgroup in MID_DON
|   |   [18] Lifespan < 12 m
|   |   |   [19] Donations_first_year >= 2
|   |   |   |   [20] First_Donation <100, 100-1,000: Low-mid (n = 99, err =
8.1%)
|   |   |   |   [21] First_Donation >1000: Low-mid (n = 36, err = 25.0%)
|   |   |   [22] Donations_first_year in One: Low-mid (n = 463, err = 1.5%)
|   |   [23] Lifespan >= 12 m
|   |   |   [24] First_Donation <100, 100-1,000
|   |   |   |   [25] Donations_first_year >= 2: Low-mid (n = 384, err = 7.8
%)
|   |   |   |   [26] Donations_first_year in One
|   |   |   |   |   [27] Income_level in 1-, 2-, 3-, 5-, Unk: Low-mid (n =
172, err = 0.0%)
|   |   |   |   |   [28] Income_level in 4-: Low-mid (n = 38, err = 7.9%)
|   |   |   [29] First_Donation >1000
|   |   |   |   [30] Donations_first_year >= 2: Low-mid (n = 50, err = 44.0
%)
|   |   |   |   [31] Donations_first_year in One: Low-mid (n = 91, err = 19
.8%)
|   |   [32] Lifespan in NEW
|   |   |   [33] Donations_first_year >= 2: High (n = 26, err = 42.3%)
|   |   |   [34] Donations_first_year in One
|   |   |   |   [35] First_Donation <100, >1000: Low-mid (n = 107, err = 26
.2%)
|   |   |   |   [36] First_Donation in 100-1,000: Low-mid (n = 37, err = 0.
0%)
|   [37] Subgroup in MJREG_MID
|   |   [38] First_Donation <100: Low-mid (n = 1497, err = 3.9%)
|   |   [39] First_Donation >1000
|   |   |   [40] Prod_SG in No: High (n = 17, err = 11.8%)
|   |   |   [41] Prod_SG in Yes: Low-mid (n = 8, err = 37.5%)
|   |   [42] First_Donation in 100-1,000
|   |   |   [43] Prod_SG in No
|   |   |   |   [44] Channel in DIG, SPONT
|   |   |   |   |   [45] Donations_first_year >= 2: High (n = 234, err = 44
.0%)
|   |   |   |   |   [46] Donations_first_year in One: Low-mid (n = 7, err =
0.0%)
|   |   |   |   [47] Channel in DM, F2F, OTHER
|   |   |   |   |   [48] Prod_Vill_Spo in No: Low-mid (n = 22, err = 4.5%)
|   |   |   |   |   [49] Prod_Vill_Spo in Yes: Low-mid (n = 20, err = 40.0%
)
|   |   |   [50] Prod_SG in Yes
|   |   |   |   [51] Prod_CG in No: Low-mid (n = 76, err = 9.2%)
|   |   |   |   [52] Prod_CG in Yes: Low-mid (n = 36, err = 33.3%)

Number of inner nodes:    23
Number of terminal nodes: 29
```

CHAID Model for Single Givers

```
Model formula:
AALTV_Class ~ Gender + Generation + Income_level + Subgroup +
    Channel + Lifespan + First_Donation + Donations_first_year

Fitted party:
[1] root
|   [2] First_Donation <20
|   |   [3] Donations_first_year >= 2
|   |   |   [4] Subgroup in REAC_DON, SG _UNV, SG_1, SG_2, SG_3, SG_6, SG_7
|   |   |   |   [5] Subgroup in REAC_DON, SG _UNV, SG_1, SG_2, SG_4, SG_5,
SG_6, SG_7, SG_8: High (n = 1259, err = 50.0%)
|   |   |   |   [6] Subgroup in SG_3: Low-mid (n = 1626, err = 2.4%)
|   |   |   [7] Subgroup in SG_4
|   |   |   |   [8] Lifespan < 12 m, NEW: Low-mid (n = 2139, err = 0.0%)
|   |   |   |   [9] Lifespan >= 12 m: Low-mid (n = 3826, err = 1.0%)
|   |   |   [10] Subgroup in SG_5
|   |   |   |   [11] Lifespan < 12 m: Low-mid (n = 1298, err = 2.5%)
|   |   |   |   [12] Lifespan >= 12 m: Low-mid (n = 1381, err = 23.4%)
|   |   |   |   [13] Lifespan in NEW: Low-mid (n = 2222, err = 13.5%)
|   |   |   [14] Subgroup in SG_8: Low-mid (n = 1995, err = 8.2%)
|   |   [15] Donations_first_year in One
|   |   |   [16] Lifespan < 12 m, NEW: Low-mid (n = 48941, err = 0.0%)
|   |   |   [17] Lifespan >= 12 m
|   |   |   |   [18] Subgroup in REAC_DON, SG _UNV, SG_1, SG_2, SG_6: Low-m
id (n = 1119, err = 12.7%)
|   |   |   |   [19] Subgroup in SG_3, SG_7, SG_8
|   |   |   |   |   [20] Subgroup in REAC_DON, SG _UNV, SG_1, SG_2, SG_3, S
G_4, SG_5, SG_6: Low-mid (n = 1334, err = 2.4%)
|   |   |   |   |   [21] Subgroup in SG_7, SG_8: Low-mid (n = 1020, err = 0
.9%)
|   |   |   |   [22] Subgroup in SG_4: Low-mid (n = 7861, err = 0.1%)
|   |   |   |   [23] Subgroup in SG_5: Low-mid (n = 1829, err = 1.4%)
|   [24] First_Donation >50
|   |   [25] Subgroup in REAC_DON, SG_8
|   |   |   [26] Lifespan < 12 m
|   |   |   |   [27] Donations_first_year >= 2: High (n = 1150, err = 40.9%
)
|   |   |   |   [28] Donations_first_year in One
|   |   |   |   |   [29] Income_level in 1-, Unk
|   |   |   |   |   |   [30] Generation in BOOM, ELDE, GenX, GenZ, MILL, Un
k
|   |   |   |   |   |   |   [31] Generation in BOOM, ELDE, GenZ, MILL, SILE
, Unk
|   |   |   |   |   |   |   |   [32] Gender in F, UNK: Low-mid (n = 5732, e
rr = 16.1%)
|   |   |   |   |   |   |   |   [33] Gender in M: Low-mid (n = 4026, err = 
18.0%)
|   |   |   |   |   |   |   [34] Generation in GenX: Low-mid (n = 2200, err
= 21.6%)
|   |   |   |   |   |   [35] Generation in SILE
|   |   |   |   |   |   |   [36] Gender in F, UNK: Low-mid (n = 4130, err =
12.9%)
|   |   |   |   |   |   |   [37] Gender in M: Low-mid (n = 2752, err = 15.8
%)
|   |   |   |   |   [38] Income_level in 2-, 3-, 4-, 5-
|   |   |   |   |   |   [39] Gender in F: Low-mid (n = 1302, err = 22.5%)
|   |   |   |   |   |   [40] Gender in M, UNK: Low-mid (n = 1995, err = 30.
3%)
|   |   |   [41] Lifespan >= 12 m, NEW: High (n = 2954, err = 32.1%)
|   |   [42] Subgroup in SG _UNV, SG_1
|   |   |   [43] Donations_first_year >= 2: High (n = 1413, err = 2.4%)
|   |   |   [44] Donations_first_year in One: High (n = 1935, err = 7.2%)
|   |   [45] Subgroup in SG_2
|   |   |   [46] Donations_first_year >= 2: High (n = 3468, err = 7.7%)
```

```
|   |   |   [47] Donations_first_year in One
|   |   |   |   [48] Gender in F: High (n = 1809, err = 25.6%)
|   |   |   |   [49] Gender in M, UNK: High (n = 2420, err = 21.2%)
|   |   [50] Subgroup in SG_3, SG_7
|   |   |   [51] Channel in DIG, OTHER: Low-mid (n = 1089, err = 18.7%)
|   |   |   [52] Channel in DM: Low-mid (n = 1512, err = 6.9%)
|   |   |   [53] Channel in SPONT
|   |   |   |   [54] Gender in F, UNK: Low-mid (n = 2409, err = 10.2%)
|   |   |   |   [55] Gender in M: Low-mid (n = 1839, err = 13.4%)
|   |   [56] Subgroup in SG_4, SG_6
|   |   |   [57] Channel in DIG
|   |   |   |   [58] Gender in F: Low-mid (n = 1000, err = 22.7%)
|   |   |   |   [59] Gender in M, UNK: Low-mid (n = 1966, err = 27.9%)
|   |   |   [60] Channel in DM, OTHER
|   |   |   |   [61] Subgroup in REAC_DON, SG _UNV, SG_1, SG_2, SG_3, SG_4,
SG_5, SG_7, SG_8: Low-mid (n = 1261, err = 18.8%)
|   |   |   |   [62] Subgroup in SG_6
|   |   |   |   |   [63] Gender in F: Low-mid (n = 1897, err = 6.1%)
|   |   |   |   |   [64] Gender in M, UNK: Low-mid (n = 1552, err = 10.9%)
|   |   |   [65] Channel in SPONT
|   |   |   |   [66] Generation in BOOM
|   |   |   |   |   [67] Gender in F, UNK: Low-mid (n = 1975, err = 16.1%)
|   |   |   |   |   [68] Gender in M: Low-mid (n = 1503, err = 21.2%)
|   |   |   |   [69] Generation in ELDE, SILE, Unk: Low-mid (n = 2550, err
= 13.4%)
|   |   |   |   [70] Generation in GenX, GenZ, MILL: Low-mid (n = 1464, err
= 23.2%)
|   |   [71] Subgroup in SG_5
|   |   |   [72] Lifespan < 12 m
|   |   |   |   [73] Donations_first_year >= 2: High (n = 1785, err = 4.6%)
|   |   |   |   [74] Donations_first_year in One
|   |   |   |   |   [75] Channel in DIG, OTHER
|   |   |   |   |   |   [76] Gender in F, UNK: High (n = 1339, err = 37.9%)
|   |   |   |   |   |   [77] Gender in M: High (n = 1484, err = 32.5%)
|   |   |   |   |   [78] Channel in DM: Low-mid (n = 2195, err = 37.4%)
|   |   |   |   |   [79] Channel in SPONT
|   |   |   |   |   |   [80] Gender in F, UNK: High (n = 7587, err = 43.8%)
|   |   |   |   |   |   [81] Gender in M: High (n = 5080, err = 39.4%)
|   |   |   [82] Lifespan >= 12 m: High (n = 2728, err = 4.2%)
|   |   |   [83] Lifespan in NEW: High (n = 9779, err = 0.0%)
|   [84] First_Donation in 20 to 50
|   |   [85] Donations_first_year >= 2
|   |   |   [86] Subgroup in REAC_DON, SG_3, SG_5
|   |   |   |   [87] Lifespan < 12 m: Low-mid (n = 1470, err = 23.6%)
|   |   |   |   [88] Lifespan >= 12 m: High (n = 1567, err = 38.4%)
|   |   |   |   [89] Lifespan in NEW: High (n = 1964, err = 11.4%)
|   |   |   [90] Subgroup in SG _UNV, SG_1, SG_6: High (n = 2105, err = 22.
4%)
|   |   |   [91] Subgroup in SG_2, SG_7: Low-mid (n = 2236, err = 33.7%)
|   |   |   [92] Subgroup in SG_4
|   |   |   |   [93] Lifespan < 12 m, NEW: Low-mid (n = 1798, err = 0.8%)
|   |   |   |   [94] Lifespan >= 12 m: Low-mid (n = 2050, err = 8.2%)
|   |   |   [95] Subgroup in SG_8: Low-mid (n = 1817, err = 19.1%)
|   |   [96] Donations_first_year in One
|   |   |   [97] Subgroup in REAC_DON, SG _UNV, SG_2, SG_3: Low-mid (n = 32
08, err = 14.2%)
|   |   |   [98] Subgroup in SG_1: Low-mid (n = 1336, err = 47.5%)
|   |   |   [99] Subgroup in SG_4: Low-mid (n = 5024, err = 1.1%)
|   |   |   [100] Subgroup in SG_5
|   |   |   |   [101] Lifespan < 12 m
|   |   |   |   |   [102] Channel in DIG: Low-mid (n = 1296, err = 0.5%)
|   |   |   |   |   [103] Channel in DM, OTHER: Low-mid (n = 3747, err = 0.
0%)
|   |   |   |   |   [104] Channel in SPONT: Low-mid (n = 9259, err = 0.2%)
|   |   |   |   [105] Lifespan >= 12 m: Low-mid (n = 1958, err = 18.2%)
|   |   |   |   [106] Lifespan in NEW
```

```
|   |   |   |   |   [107] Income_level in 1-, 2-, 3-: Low-mid (n = 4153, er
r = 2.8%)
|   |   |   |   |   [108] Income_level in 4-, 5-: Low-mid (n = 3048, err =
3.7%)
|   |   |   |   |   [109] Income_level in Unk: Low-mid (n = 3123, err = 6.0
%)
|   |   |   [110] Subgroup in SG_6: Low-mid (n = 13112, err = 0.1%)
|   |   |   [111] Subgroup in SG_7: Low-mid (n = 7389, err = 0.0%)
|   |   |   [112] Subgroup in SG_8
|   |   |   |   [113] Lifespan < 12 m: Low-mid (n = 19843, err = 0.0%)
|   |   |   |   [114] Lifespan >= 12 m, NEW: Low-mid (n = 1478, err = 7.7%)

Number of inner nodes:    43
Number of terminal nodes: 71
```