Università
Ca'Foscari
Venezia

Master's Degree

in Management

Final Thesis

The Big Five Personality Traits and purchase behaviour

A quantitative study

on second-hand clothes consumers

Supervisor
Ch. Prof. Lisa Crosato

Graduand
Ilaria
Cannazza
Matricolation number
890083

Academic Year
2002 / 2003

Acknowledgements

At the end of this five-year journey, I would like to express my sincere gratitude to all the people who have been there along the way. First and foremost, I want to give a big thank you to my thesis advisor, Professor Crosato. She has patiently and devotedly guided me throughout the process of writing my master's thesis. Thanks to her valuable advice and suggestions, I had the opportunity to expand my knowledge and increase my expertise in this sector.

I also want to express my appreciation to all the professors who have helped me to discover the various aspects of this field of study. It is thanks to them that I started to love some subjects and I was able to choose my own path. I am grateful to the University of Venice for giving me the chance to spend a year abroad, and to the EM Business School and its professors for their warm welcome and assistance. Thanks to these various and new experiences, my journey is even richer and more stimulating. I would like to extend my gratitude to all the participants who patiently responded to my thesis. Without their contribution, this work would not have been possible.

A special and immense thank you goes to my parents, who were the first to encourage me to pursue this degree and to explore new horizons away from home. They have always been by my side, supporting me and believing in my abilities.

Another huge thanks goes to Leo, who has supported me for these two years, especially in the moments of difficulties. He has consistently encouraged me to never give up, particularly when I felt inadequate. Thank you for being my constant pillar of support, someone I could rely on day and night, whether it was before exams or after their results. You have always pushed me to prove my capabilities and part of this degree is also yours.

Of course, I cannot forget all my friends across Italy and Europe. Between them I wanted to thank my loyal friends from Lecce, that never forgotten about me in these years in which I was far from them. Then there are all the friends who have become indispensable to me during my time in Padua and have been flatmates, classmates, travel and adventure companions and everything else someone can dream of. This long path has brought me to Strasbourg, where while suffering for the distance from all my friends, some people made their space into my

heart. What I really hope is that these friendships will not simply remain a temporary part of my journey, but will last in time with all the memories connected to them.

This moment represents the end of a long and incredible journey and while I’m happy to move on new adventures, I’m also scared of what will come next. But then I think how much I was terrified before starting the university and how much these five years meant to me in the end, and I can’t wait to continue. I feel strong because I know that you will be there for me and I will be there for you, and that we will choose different paths but remaining in the lives of each other.

Summary

Introduction

According to the theory of the Big Five Personality Traits, the traits defined as neuroticism, extraversion, openness to experience, agreeableness and conscientiousness exist in every individual in various degrees and their interaction gives everyone a unique thought process and behavioural model. This research merges this theory with the classification of motivations to buy second-hand clothes, which include critical, economic, hedonic/recreational and fashion motives. The aim is to create a broader paradigm, which can explain the drivers and the reasons for the expansion of this shopping trend and whether the positive attitudes toward this sector are determined by differences in the purchase motivations of the individuals and by a combination of personality traits and psychographic characteristics. In particular, this research starts focusing on the role of the Big Five Personality traits as explanatory variables of the differences in the motivations of the individuals to purchase.

This study fills a gap in the existent literature, which never included both the Big Five personality traits and the various motivations to purchase second-hand clothes, as a way to explain the attitude of consumers toward this trending way of purchasing. This research can represent the basis for some future marketing campaigns and strategies for the various channels, which want to invest in this expanding sector. The second-hand sector is growing at an incredible rate and it can represent a valid way to reduce the impact of the fashion sector on the environment, which is every year more threatened by intensive mass production. Furthermore, as young generations are becoming more interested in this way of purchasing, studying the characteristics and drivers of the most affectionate consumers is a way to fasten the positive change.

The research is structured across six different chapters, where in the first one there is a review of the literature about the fast fashion paradigm and the problems associated with it. Subsequently, new emerging sustainable fashion trends are presented, in particular the phenomenon of second-hand purchasing. Through the previous literature, its history is explained along with the consumer’s willingness to pay. Furthermore, an explanation is provided regarding the characteristics of different motivations for purchasing second-hand clothes, as well as the characteristics of various retail channels.

The second chapter provides a summary of the existing literature on personality, with a particular focus on the Theory of Traits. Additionally, it explains the traditional methods used to measure these traits across individuals. The majority of the attention is given to the Big Five Personality Traits theory, providing an overview of the history, characteristics, and descriptions of these five traits. Furthermore, this chapter explores how these traits influence individuals' consumption behavior.

The subsequent chapters describe the product of the research and the methodologies, such as multiple linear regression, logistic regression and variable selection methods, used to conduct the analyses through the software RStudio [1]. In particular, the third chapter describe the data collection, realized through an online questionnaire distributed to second-hand clothes consumers. Before conducting the analysis of the 109 completed and usable answers, the various variables were analysed, in order to gain insights about their distribution and their cross-correlations.

The fourth chapter analyses how consumers' Big Five Personality Traits can potentially influence their critical, economic, hedonic, and fashionability motivations to purchase second-hand clothes. In the fifth chapter, the research examines if factors beyond personality traits can explain consumers' attitudes towards second-hand purchases. It explores whether variables considered in the study can account for the variability in Buying Behavior. Moreover, it investigates whether these factors also influence the frequency of individuals' second-hand clothing purchases. Moving to the sixth chapter, the analysis focuses on the possibility of categorizing consumers into distinct independent groups based on their values of purchasing motivations. The aim is to uncover similarities and differences in consumer behavior towards this emerging sector. This method proves to be the most effective in describing the relationships within the data, in a coherent way with the previous studies on the topic.

Among the results of the analysis, only some of them are coherent with the ones from previous literature, which are summarised in the first two chapters. The clustering analysis is the most effective method in describing the relationships within the data. Thanks to it, it was possible to divide the consumers into two groups and describe those who have positive motivations to

[1] The theory about the different analyses is obtained from James et al. (2021)

purchase second-hand clothes through their personality traits and their psychographic characteristics.

Chapter 1: Second-hand

1.1 The fast fashion paradigm

Our society is characterised by rapid changes and modifications and the retail sector has lived some of these transformations, in particular, the apparel sector transformed itself from a push system, in which the designers decide trends and collections, to a pull system where the power is given to the consumer, that is able to dictate times and demands. (Barnes & Lea-Greenwood, 2010) This led to the creation of new concepts, such as just-in-time (JIT) delivery and quick response (QR), to habilitate companies to follow these expectations and requests of consumers in a fast way (Barnes and Lea-Greenwood, 2010). From these changes in the sector, it is born the concept of fast fashion, which expresses a shift in the previous paradigm. (Pookulangara & Shephard, 2013)

This term was coined in the 1990s by the New York Times, defining it as “a business strategy that aims to reduce the processes involved in the buying cycle and the lead times for getting new fashion products into stores, in order to satisfy consumer demand at its peak” (Barnes & Lea-Greenwood, 2006, p. 259, cited in Koay et al., 2023)
The goal for the brands is to reduce the lead time, the amount of time that passes from the ideation to the sale of a product (Watson e Yan, 2013) and to create fashion items which follow the latest trends (Barnes & LeaGreenwood, 2010)
To drive the consumption and purchases of clothes, fast fashion retailers started to introduce more frequently new collections, sometimes on a weekly basis, which lead to faster selling cycles and a change in consumers purchasing habits, which become of smaller entities but more frequent. (Pookulangara & Shephard, 2013)
This business model develops in the consumer a sense of urgency, due to the continuous change of offers, and pushes him to make compulsive and impulsive purchases. (Niinimäki et al., 2020)
Retailers modified their supply chain, to be able to create and diffuse new collections that follow the new requests, in a lead time of around one month. (Watson e Yan,2013) These changes are followed by those in the stores, which are characterised by increasing electronic communications, frequent deliveries and minimal reductions. (Watson e Yan,2013)
This new business model of fast fashion is also based on reducing the life cycle of the garments to a month or less, resulting in fashion overconsumption of low-quality products. This brings

more profits for the growing number of companies that are starting to follow this business model, (Koay et al., 2023) defined as throwaway fashion. (Pookulangara & Shephard, 2013)

A consequence of this new fashion paradigm is that following the trends and so buying more frequently, consumers expect lower prices. (Pookulangara & Shephard, 2013) So following their requests, the prices of garments dropped in recent years, thanks to the raw materials and the labour force of developing countries such as Vietnam, Bangladesh and Pakistan that are being exploited for this continuous production. (Koay et al., 2023)
This model pushed a continuous increase of the clothing demand, which is growing by around 2% every year, with a consequent increase of the quota per capital textile production, that over the period 1975-2018, more than doubled reaching 13 kg per year. Also, global consumption is estimated to be 62 million tonnes per year and will reach about 102 million tonnes in 2030. Today compared to 20 years ago, production is doubled and become more efficient, a factor that allows driving the prices down. This process can be considered an unvirtuous circle because low prices will push more compulsive consumption and waste, which can be shown by the decrease of 36% in the average garment use time with respect to 2005. (Niinimäki et al., 2020)

With time this fast fashion model has become very successful, overcoming the traditional and historical fashion retailers, (Niinimäki et al., 2020) with a market value worldwide in 2009 equal to $22bn and with an estimated value of $40bn in 2030, data that representative of its exponential growth. (Statista, 2022). Despite this exponential growth, as explained in the following chapters, fast fashion will be overcome by second-hand apparel, whose market value was only $10bn in 2009, but is projected to reach $84bn in 2030. (Statista, 2022)

1.1.1 Problems associated with the fast-fashion model

This model is not sustainable in the long term for the environment, because it is based on continuous infinite growth but world resources are finite. These premises led the fashion industry to be the second largest polluting industry on the planet, with increasing damage that follows its growth. (Lin & Chen, 2022)
The fast fashion business model is characterised by a vertical disintegration and global dispersion of its complex supply chain. A consequence of the expansion of the supply chain is the exploitation of developing countries or the Global South that become the largest producers of textiles and clothing, due to their low labour costs. (Niinimäki et al., 2020) Around 70% of the European Union’s textile and apparel imports come from Asian countries where most of the

time the workers are lacking adequate social security and security in the working place, being forced to work more than allowed for low wages. (D'ambrogio, 2014, cited in Şener et al., 2019) These countries are dealing with all the environmental consequences of products that are consumed by the developed countries or Global North, where the headquarters and the creative offices of the brands. This fragmentation and globalisation of the supply chain, all over the world, lead to a consequent increase of industries that are part of the long process from agriculture to manufacturing, logistics and retail. Due to each step of the chain being in a different country, it is more difficult to control and assess the environmental impacts of companies and products, decreasing the transparency of clothes production. (Niinimäki et al., 2020)

The fashion industry represents a key environmental global threat, at the moment this industry produces approximately 8-10% of the global carbon emissions and it is also a major consumer of water, being responsible for around 20% of the industrial water pollution. Looking at a single piece of clothing, such as a t-shirt or a pair of jeans, it is possible to see that the huge water footprint of these products comes from 90% of the cotton cultivation that is exported to poor countries.

It is important to remember that every year this sector is responsible for the introduction of 190000 tonnes of microplastic into the ocean (35% of the total) and 60% of the global fibre production is destined for fashion production. (Niinimäki et al., 2020)

Another hazard from this polluting industry is their use of chemicals in the production process, more than 15000 different types of chemicals with a concentration that can reach 466g per kilo of textile, as shown in an analysis of a European textile-finishing company. These chemicals also have hazardous properties that can represent a big problem for human health, due to our direct contact with them. (Niinimäki et al., 2020)

A natural consequence of the sector of throwaway fashion, characterised by fast trends and fast purchases of low-quality products, is the increase in textile waste, that at the moment is around 92 million tonnes per year or the exportation of these old garments to developing countries.

Both the short clothing lifetimes and the increased number of purchases by individuals had led to an increase in post-consumer textile waste, which now represents 22% of the mixed waste worldwide. To understand the dimension of this phenomenon, the average USA and UK consumers waste 30 kg of clothes every year. The recycling rate of these garments is low, with only 15% of them collected separately and less than 1% of the total production recycled in a closed loop, so in similar quality garments. The waste is not just after the consumption of the pieces of clothing, because according to some research, 15% of the fabrics used in the production are wasted. This waste, produced during manufacturing, is called production waste. Another type of pre-consumer waste is called deadstock, consisting of all the new pieces of clothing that are not sold or that are returned by customers that become waste. According to Ecotextile News, in 2016, only 30% of all the clothes in the European Union is sold at the full retail price, the rest is sold at a discount price or not sold at all. (Niinimäki et al., 2020)

1.1.2 New trends

Due to these profound effects on the environment, the fashion industry in recent years is dealing with increasing investigations into its operations. (Niinimäki et al., 2020)
The environmental and social impacts of this sector increased consumers' worries and started to shape consumers' attitudes, pushing them to request more sustainable products that respect ecological and social principles, (Lin & Chen, 2022) as a way to take into consideration the non-financial cost of this sector (Niinimäki et al., 2020), This movement of "ethical consumerism", involve consumers that are willing to pay a premium price for those products that respect the environment and the society (Pookulangara et al., 2011 cited by Pookulangara & Shephard, 2013) and that press companies to change their production. (Pookulangara & Shephard, 2013)

Even if brands, trying to catch this segment of consumers, are committing to minimise their impacts, this is not enough because consumers are often not able to understand the impact of their buying decisions and still are guided by benefits and motivations that are functional, financial or aesthetic. This results in low participation in the consumption of more sustainable products. (Lin & Chen, 2022) The best for society and for the environment is a degrowth of the fashion industry, a planned economic contraction, driven by a reduction of the volume of production. (Lin & Chen, 2022) The challenge is to define the fair share of production and consumption due to social, cultural and psychological factors, like the future of developing countries that currently are living thanks to mass production. What is needed is a balance, between slowing production and better quality products with longer lifetimes.

1.1.3 Slow fashion

This shift in consumer behaviour toward more sustainable clothing consumption is highlighting the importance to come back to the "slow" fashion, which is declined in fewer purchases by consumers and the production of garments with longer lifetimes. (Niinimäki et al., 2020) The slow fashion movement is a philosophy of attentiveness that takes into consideration the stakeholder's needs and the impact of the garments on consumers, workers and the environment (Pookulangara & Shephard, 2013)

This can be considered a change of paradigm, where the focus for companies moves from increasing profits and sales, through fast manufacturing and low product quality, to a more sustainable production driven by quality and better materials. (Niinimäki et al., 2020) So companies design small lines of high-quality and regional products at fair labour conditions incorporating the business concepts like social responsibility, sustainability and transparency while maintaining the profitability of the company. (Fletcher, 2010, cited in Pookulangara & Shephard, 2013)
In this paradigm, consumers are encouraged to consider their purchase in a holistic manner, where the design, production and consumption of the product, have a positive impact on the world around, with its people, cultures and environment. (Thomas, 2008, cited in Pookulangara & Shephard, 2013)

This approach gives the consumers the possibility to be fashionable, in a sustainable and ethical way (Pookulangara & Shephard, 2013), thanks to a process that is not time-based but quality-based and that focuses on products that are durable and of timeless style. (Fletcher, 2007, cited in Pookulangara & Shephard, 2013) In this way, it is possible to minimise and mitigate the environmental impact of clothes and increase the long-term sustainability of this sector. (Niinimäki et al., 2020)

The purchase has an influence on how individuals feel and behave and it can be seen that purchasing a sustainable product can improve an individual image in a social group with a feeling of exclusivity and differentiation from others. (Meyer, 2001, cited in Şener et al., 2019) The results of the study by Şener et al. (2019) highlight that the perceived customer value influences the purchase of slow fashion products and the willingness to pay more for those.

1.1.4 Circular economy

Another approach to increase the environmental sustainability of the fashion sector is trying to reduce textile waste, and keeping materials and products in the system. This can be done by

implementing the model of circular economy, defined as the opposite of the current linear system that can be declined in take, make and dispose of.
This circular system follows 3 different approaches to be able to close the loop: narrowing (efficiency), closing (recycling) and slowing (reusing). (Niinimäki et al., 2020) "Circular products" are those that are produced with recycled materials that help to close the production loop or those with an extended life cycle, so or more durable or resold. (Pretner et al., 2021)

This new model wants to spread the concept of eco-efficiency by increasing the use of existing products and decreasing the consumption of new ones. (Niinimäki et al., 2020) There are different ways of consuming that search for new ways to close the loops of products and between them, there is the collaborative consumption and the access-based one. They are peer-to-peer ways of renting and sharing items between consumers, and they are starting to increase their popularity in the fashion sector for the research of designer clothing or clothes for special occasions. (Niinimäki et al., 2020) The sharing economy sector, which is based on exchanges and sharing, and leasing and renting, is growing and in Europe is estimated to be worth 28 billion. (Niinimäki et al., 2020)
Companies tend to follow the demand of consumers and due to this rising interest in alternative forms of consumption, they also started to implement some collaborative business models. Especially in the luxury sector, they introduced strategies to increase the lifespan of their products, like repair services and second-hand sales. It is important to consider that the benefits of these new forms of consumption are mitigated by the additional transportation of used clothes. (Niinimäki et al., 2020)

1.2 Second-hand

Inside the circular economy, it is possible to identify the second-hand sector, which in recent years has been growing at high speed with its dimension doubled in a few years, gaining the attention of consumers and companies. This explosion of importance is due to a change of paradigm for this sector, which is no longer perceived as a way to purchase economic and used clothes. Now it has become a global fashion trend that is able to bring with itself some environmental qualities and values, in fact, scholars believe that SHC consumption has positive impacts on society. (Herjanto et al., 2016)

The second-hand sector is defined as "the acquisition of used objects through often specific modes and places of exchange". (Roux & Guiot, 2008, p. 66) Even if it is part of the circular sector, it has its characteristics and particularities (Pretner et al., 2021) and it can be analysed on its own.

The purchase of second-hand is considered a transaction that is advantageous to the buyer, due to the attractive and low price of the object of the exchange. This second-hand object, differentiates itself from a new equivalent one, due to the previous possession by another owner. (Roux & Guiot, 2008) Second-hand purchases can be linked to the concept of smart shopping, defined by Mano and Elliot (1997), as the desire to overcome the conventional markets and purchase small-price products that others don't want anymore. This provokes a sort of self-gratification and a sense of belonging to a group of experts in alternative purchasing practices. (Mano & Elliot, 1997)

For second-hand clothes, we intend pre-owned and used clothes that can be given to family, friends or other consumers through private or market exchanges. In the first case, these pieces of clothing are able to change owners without the payment of a price or a deal of money. This research focuses on market exchanges, characterised by an exchange of money between individuals, that corresponds to the price of the item. (Laitala & Klepp, 2018) The interactions that lead to the exchange can happen in different locations and platforms like second-hand and vintage stores, charity thrift stores, flea markets or online websites and platforms. (Laitala & Klepp, 2018)

1.2.1 Second-hand willingness to pay

According to previous research, consumers are willing to pay a premium price to purchase products that have environmentally friendly attributes or are considered green. (Pretner et al., 2021) But examining the willingness to pay for circular products, Pretner et al. (2021) found that second-hand products are perceived as having inferior quality with respect to comparable new ones. For this reason, the consumers are not ready to pay a premium price for their purchase and also expect a lower one with respect to an equivalent product. (Pretner et al., 2021)

What can cause this difference in willingness to pay, is the strong perception of contamination for second-hand clothes. (Pretner et al., 2021) This concept concerns the interactions that happen when something circulates and it is defined by Baxter et al. (2017, cited in Laitala &

Klepp, 2018) as the positive or negative modification of the perception of an object, after being contaminated by someone or something, real or imagined.

Real contaminants are objective and measurable elements that influence the nature of the piece of clothing, while imagined contaminants are beliefs or mental associations about characteristics of the objects that give a negative or positive value to them. According to Baxter et al. (2017, cited in Laitala & Klepp, 2018) to be able to extend the lifespan of products it is important to eliminate the negative contaminations, with concrete actions for the real ones and with logic for the imaginary ones.

Also, the willingness to pay can be low because sometimes it is not always evident the environmental benefits of buying used garments compared to new ones. (Pretner et al., 2021) In fact, the consumer's willingness to pay for the product increases when there are some pieces of information about its environmental virtues, and even more when a third party or an ecolabel verifies it. This verified information has a stronger impact on low-concerned environmental consumers, it helps to overcome their initial doubts, provide information that the consumers are not conscious of and help them recognise the benefits of the purchase. These statements can decrease the contamination perception, increasing the quality level of the product thanks to the work of this third-party subject. (Pretner et al., 2021) Even with the presence of an external certification, the willingness to pay is always lower than for an equivalent new product. The imposition of a "premium price" for second-hand clothes is not a winning strategy, even if this is in contrast with the literature, that highlights a positive effect of environmental concern on the final WTP and purchasing intention. (Pretner et al., 2021)

1.2.2 Second-hand history

The phenomenon of second-hand shopping can be divided into different periods, where during the eighteenth and nineteenth centuries there was its rise and growth and during the twentieth century this way of purchasing was stigmatised and just attributed to lower social levels. (Ferraro et al., 2016) During these years, second-hand clothing was seen as old products that were worn-out, dirty and stinking (Zaman et al., 2019) and so they were considered undesirable (DeLong et al., 2005).

The change in the consumer's perceptions started in the 1990s, an important time for second-hand clothing consumption, with the trading of second-hand clothes starting to become more popular all over the world. (Herjanto et al., 2016)

Then after the Global Financial Crisis in 2008, second-hand shopping was a way for some consumers to overcome a period of financial difficulties and in the next years due to growing concern about the sustainability of products and clothes, there was another increase of the attention given to this sector. (Ferraro et al., 2016)

It is just in recent times that society's consideration has changed, (Ferraro et al., 2016) with now consumers that attribute a high value to these products, which can be even equal or superior to new ones on the mass market, characterised by low-quality and lack in ethics. (Zaman et al., 2019) This sector is growing in importance because for many customers it is an alternative option to conventional products and retail options in the mass consumer market. This exponential growth can be seen through the graph of Statista (2022) which showed the second-hand apparel market value worldwide from 2021 to 2027, which in 2021 was equal to 138 billions of dollars and is projected to increase to 351 billion in 2027. Due to the high growth of this channel, there are profound transformations in the retail sector, where traditional retailers and fast fashion ones are trying to implement new strategies to not lose consumers. (Ferraro et al., 2016)

Also in recent times, consumers started to consider second-hand clothes as something fashionable. These fashionable clothes are labelled as vintage and have a higher price, to reflect the higher perceived value. (Cervellon et al., 2012). This value can originate from their scarcity and age, because these clothes can be vintage designer pieces, due to their design style or period, due to they are unworn. (Ferraro et al., 2016) As a result, consumers are appealed by the authenticity and uniqueness of these pieces of clothing (Guiot & Roux, 2010) and value them positively, spending energy to find them.

These shifts in the paradigm had led second-hand to become an increasing trend, especially in Europe and USA, transforming the way people dispose of their old articles and pushing them to resell, recover and recycle. This new concept of shopping is constituted by a product and a sales dimension, one concerned on non-buying new items and the other that is about the types of shops one goes to. (Guiot & Roux, 2010)

The growth of the sector of second-hand goods leads to an increase in the number of second-hand retail channels, which quickly become strong competitors of traditional retail outlets. The spread of these alternative retail channels can be explained by the existence of some expectations of the different types of consumers, which are not satisfied by the existing possibilities on the market. In fact, the second-hand sector is able to satisfy some consumers' needs that go beyond the economic advantages, like recreational benefits from flea markets, and garage sales, in which individuals are able to find unusual items that are not present in the conventional mass shops. (Guiot & Roux, 2010)

1.3 Motivations to the purchase of second-hand clothes

Second-hand shopping, the acquisition of second-hand objects by the consumer, through distinctive methods and places of exchange with respect to traditional ones, is driven by different motivations. (Guiot & Roux, 2010) They are able to influence and determine the actions and behaviours of individuals, orientated by their wishes, desires, emotions, feelings and beliefs. (Guiot & Roux, 2010)

“Motivation” can be defined as a set of goals directed towards specific gratification and satisfaction (McGuire, 1974, cited in Padmavathy et al., 2019) or as the justification by the consumer for behaving in a specific way. (Laitala & Klepp, 2018) According to Mucchielli (1992, p. 29), motivations can be defined as “the totality of irrational determinants of human conduct” that “impel the individual to act” and “direct his behaviours”. (cited in Roux & Guiot, 2008)

In the past years, researchers studied the second-hand clothes sector focusing only on the economic values of this sector and its economic savings that can be very important for some categories of individuals. With time scholars started to investigate other motivations that are able to drive this new form of purchasing, including more complex and different factors that can influence the consumer’s perception of second-hand clothes, apart from financial needs.

The reasons for the growing interest of the population in this sector don’t lie just in the economic and financial convenience of it but also in the sales channels. They satisfy some consumers' needs, like rummaging through the products to find something, searching for surprising and unusual products, hunting for unique products that are not present in traditional shops, and the possibility to bargain and have a discussion with the seller. (Roux & Guiot, 2008)

For these reasons, sometimes it makes no sense to compare the second-hand purchase with an equivalent new product, in fact behind the simple transaction lies an affective dimension, that is born from the history of the product or the interaction with the seller during the purchase. (Roux & Guiot, 2008)

Experiential motivations linked to the nature of the offering

Looking at the consumer motivations to buy second-hand products concerned with the nature of the offering, it is possible to distinguish different variables that influence the purchase: the originality of the products, the nostalgic pleasure of finding objects from the past, self-expression through the restoration, repair, or transformation of things and the congruence between the object and the consumer. (Guiot & Roux, 2010)

The consumer can be interested in the products for their authentic, unique or rare composition or because they can represent their personality or because thanks to them it is possible to push creativity in their combinations and use, giving the object a creative potential different from the original one. (Sherry, 1990) It is also possible that these objects can evoke emotions or memories of the past, incorporating their history and provoking the feeling of nostalgia in the passionate consumer. (Guiot & Roux, 2010)

Experiential motivations linked to channel characteristics

The channel characteristics and the context of the purchase of second-hand objects are important because they can provoke recreational sensations in the consumer through social contact, stimulation, and treasure-hunting values. The context in which people buy second-hand doesn't include just physical shops but also locations in which it is possible to stroll between the different stands and experience amusement in which no one is forced to buy anything. (Guiot & Roux, 2010) If these locations are distant or are available just in certain periods, the research becomes something similar to a journey to discover something new. In this context, the interaction between the seller and the buyer is also frequent and can be something that is researched by both of them, as a way to discover more about the history of the object and its values. Sometimes this interaction and research of the object can derive the consumer a profound link with it. (Guiot & Roux, 2010)

1.3.1 Studies about Motivations

These motivations to purchase include the individual's values and beliefs, social pressure and the desire for self-expression. They are multifaceted and are compounded then more than a factor. (Herjanto et al., 2016)

According to Westbrook and Black (1985), shopping motivations are formed by three different parts: the desire to acquire a product, the wish to satisfy some personal links not connected with the item and the goal of achieving specific ends. (Westbrook & Black, 1985) According to this theory, it is possible to hypothesise that consumers shop second-hand also because they like the more informal atmosphere and the context in which they acquire these goods, which sometimes are unique and original. (Sherry, 1990)

Stone et al. (1996) defined five dimensions of consumer motivations for the acquisition of products: non-planned and impulsive purchase, the atmosphere of the location, exploration and treasure-hunting, social interaction and the research for products of quality.

Bardhi and Arnould (2005) focusing on second-hand shopping, identified a duality of economic and hedonic motivations of the consumer, in a double dimension constituted by both the product and the channel of purchasing. These motivations for the purchase should not be seen as opposite poles, but as dimensions that are complex and that are interdependent in their influence on the individual. (Bardhi & Arnould, 2005)

Most second-hand transactions are made in an informal way and location, so it is difficult to evaluate the motivations that push consumers because there is a lot of variability in this way of purchasing. (Roux & Guiot, 2008) Different scholars tried to categorise these different sets of motivations, in particular, it is relevant the work of Guiot and Roux (2010), they created a classification of the motivations following the motivation-based approach.

Through their scientific research, they proposed a reliable and valid scale for measuring the motivations for the purchase of second-hand goods. This scale is divided into a tripartite structure (critical, economic and recreational motivations) that declines itself in a second-order hierarchical structure, cataloguing the different motivations for the purchase of second-hand products through different distribution channels. Motivations that determine the individual's behaviour in this sector can be articulated as general motives that are reflected in some specific dimensions. (Guiot & Roux, 2010)

Ferraro et al. (2016) introduced the additional motivation of fashionability to Guiot and Roux's (2010) scale. They were the first researchers to recognise the role of this factor in influencing the purchase of second-hand clothes, also due to the recent second-hand retail market developments. (Ferraro et al., 2016)

According to previous research, these combined motivations are able to predict consumer behaviour in the sector of second-hand, the frequency of buying second-hand clothes and the number of second-hand channels visited, confirming the links between purchasers' overall motivations and their actual behaviour. (Guiot & Roux, 2010) According to Guiot and Roux (2010), economic and recreational motivations are intrinsically intertwined, Ferraro et al. (2016) extended that critical, and fashion motivations are also intertwined with economic and recreational motivations for some consumers. This interest in the fashionability of second-hand clothes is an opportunity for retailers to create value for consumers through fashion, focusing on offering products that are curated and follow fashion trends. (Ferraro et al., 2016)

I will sum up the more important definitions provided by different authors about this complex set of categories of motivations.

1.3.2 Economic motivations

In past studies, the economic dimension was considered the main reason for the purchase of second-hand goods. The importance of these economic-saving benefits for consumers comes from the fact that individuals have different financial resources that have to be divided into proportions for each area of their personal life. (Williams & Paddock, 2003, cited in Guiot & Roux, 2010) The economic convenience of second-hand objects can decrease the budget allocation pressure on individuals, in a way that allows them to satisfy their primary needs and buy clothes without renouncing other items. (Guiot & Roux, 2010). This aspect can be essential for low-income consumers that in this way of purchasing can find a way to regain breath. (Ferraro et al., 2016).

In fact, it is possible to observe that people with lower income or financial availability tend to direct themselves toward consumption alternatives of conventional channels, such as the second-hand market. (Herjanto et al., 2016) In their research, (Grasso et al., 2000) find that it is not the personal income of the individual, but the collective household income, that influences purchasing decisions.
Another factor that can influence the expenditure is the individual spending control mechanism, frugality, which limits unnecessary purchases of unuseful goods to be able to focus on long-term goals. These people, with higher levels of frugality, tend to have high values of economic consciousness and they tend to be less oriented toward materialistic and compulsive behaviours. (Herjanto et al., 2016)

The economic motivations for the purchase of second-hand clothes can be declined into some further categories which are the role of price and the fair price. (Guiot & Roux, 2010)

Price sensitivity or price consciousness is defined by Lichtenstein, Ridgway and Netemeyer as "the degree to which the consumer focuses exclusively on paying low prices". (1993, p. 235, cited in Roux & Guiot, 2008) What can happen is that due to their strong attention to price, consumers tend to focus just on some possible negotiations and haggling without evaluating the objects they are buying. (Bardhi & Arnold, 2005)

Fair price can be defined as the denial to pay a premium price for an object that is new compared to a used one. (Roux & Guiot, 2008) According to the authors Roux and Guiot (2008), for an individual, the possibility to access lower-priced second-hand clothes brings an economic value to the purchase but also a psychological satisfaction, that derives from the price gratification and the opportunity to bargain hunting, a dimension connected with power. (Herjanto et al., 2016)

To confirm the theory, Christiansen and Snepenger (2005), state that often individuals' experts in the purchase of second-hand products have a below-average income or financial ability but focusing on the economic motivations of the purchase is not enough to describe the heterogeneity of this channel and context. (cited in Laitala & Klepp, 2018)

1.3.3 Hedonic/recreational motivations

In the literature, different hedonic and recreational motivations were identified as being able to motivate the purchase of second-hand products. (Laitala & Klepp, 2018)

Recreational motivations for acquiring second-hand clothes originate from some missing situational factors in the conventional channels (Guiot & Roux, 2010) and derive from the characteristics and the appeal of the second-hand purchasing channels. In fact, they often are characterised by the possibility to shop outdoors, various offerings, visual stimulation due to the big quantity of goods, the possibility to interact with others, the feeling of affiliation and also the fast hunting searching for valuable pieces. (Sherry, 1990)

Recreational motivations can be declined in some situational factors, that can be summed up in further categories: treasure hunting, originality, nostalgic and social contact. (Roux & Guiot, 2008)

An important unique factor that distinguishes these channels is the possibility of treasure hunting and searching and finding authentic clothes or valuable ones. (Ferraro et al., 2016). Most of the time consumers enjoy the research between a multitude of clothes of that rare or unique piece. When individuals go into these second-hand channels to search for something, they feel the adrenaline of the hunt and this experience provokes positive emotions and satisfaction. (Herjanto et al., 2016) This concept is linked with browning behaviour, defined by Bloch and Richins (1983, p. 389) as “the examination of a store’s merchandise for recreational or informational purposes without a current intent to buy” that can be stimulated by the characteristics of this sector. (cited in Roux & Guiot, 2008) Also, Lombart (2004) describes it as some leisure activities with mostly recreational purposes. (cited in Roux & Guiot, 2008) So it is possible to hypothesise that the recreational motivations can specify in behaviours like wandering around just to look and to specifically buy something for enjoyment. (Roux & Guiot, 2008) The atmosphere of these stores is so different from the traditional shops that for the consumers it feels like being in a museum to explore and discover. (DeLong et al., 2005)

Second-hand clothes can be appreciated for their rarity, originality, and their geographical region. Consumers may search for something that is unique or different from what is present in the mass market stores, as a way to differentiate in an original and entertaining way. Also in that environment is possible for the consumer to satisfy some nostalgic values, that originate from being surrounded by clothes and objects with a past that evokes a certain era. (Guiot & Roux, 2010). This is why second-hand objects are often not compared with new ones, taking into consideration the price, because the purchase can evoke different emotions in the individual. (Ferraro et al., 2016)

What is also important is the social interaction that can originate between the seller and the buyer. It is mostly friendly and it can add value to the purchase, creating a sense of community and fulfilling the socialisation aim of some consumers. (Ferraro et al., 2016)

These hedonic and recreational motivations of the purchase also lead to the expansion of this way of purchasing. (Ferraro et al., 2016). There is a growing segment of second-hand shopping enthusiasts and collectors that love the thrill of the hunt and the possibility to find a rare piece of clothing, or something significant that can help them define themselves. (DeLong et al., 2005)

1.3.4 Critical motivations

Guiot and Roux (2010) included the critical dimension into the classical economic-hedonic duality representation of motivations to purchase (Bardhi & Arnould, 2005) Individuals recently are more concerned about the environmental and societal impact of clothes, generating an eco-movement that fights to reduce the impact of excessive, wasteful and environmentally unsound products. (Ferraro et al., 2016) So purchasing second-hand clothes is linked with political consumption, which Micheletti (2003, p 2) defines as “action by people who make choices of producers and products with the goal of changing objectionable institutional or market practices.” (cited in Laitala & Klepp, 2018) The concept of political consumerism is linked with environmental values, the ethics of individuals and their attitude against waste, and it is particularly important for some sectors like fashion, which are characterised by high waste. (Laitala & Klepp, 2018)

According to (Guiot & Roux, 2010), this critical motivation for the acquisition of second-hand objects can be declined for the consumer in the possibility of avoiding conventional channels, supporting ethical and ecological concerns about recycling and combating waste, and avoiding ostentation through purchases. (Guiot & Roux, 2010)

Some individuals feel they can escape the conventional market channels, being able to consume less and avoid substituting products with new ones, when there is not a real need for the change. They follow values of anti-ostentation, conscious and voluntary avoidance of everything linked with fast fashion and mass consumption, with its negative impacts and homologation. Also, these consumers follow an anti-waste value, trying to extend the lifespan of clothes, using them for a long period, which can be considered an opposite trend compared to the actual throwaway fashion. (Guiot & Roux, 2010) Their desire to put a distance between themselves and the conventional system, (Guiot & Roux, 2010) avoiding large corporations shows the desire to respect the environment and workers. (Ferraro et al., 2016) Their way of purchasing becomes a form of public protest showing the dissatisfaction with the current clothes purchasing system and the waste of resources. (Guiot & Roux, 2010)

This critical motivation is also strongly connected with social aspects and the reputation of the individual, in fact, clothes and styles are a way for consumers to show their identity, personality and personal values. (Laitala & Klepp, 2018) People use their clothes and external look to be easily classified as a part of a specific group of people. So it is possible to use vintage or retro clothing, to show ethical and environmental beliefs that are respected through the distance of the individual from fast fashion and the consumption system and their negative impacts. (Laitala & Klepp, 2018) This behaviour of individuals can be considered a reverse of the Veblen effect, an indifference to the social codes in the system as a way to distinguish themselves from the masses. (Roux & Guiot, 2008)

1.3.5 Fashionability motivations

The fashion motivations to purchase second-hand clothes take their name from the word fashion, defined by Sproles and Burns (1994, p.4) as “a style of consumer product … that is temporarily adopted by a discernible proportion of members”. (cited in Ferraro et al., 2016) This fashion motivation is so based on the desire for fashion-seeking by the consumers. This motivation was introduced recently by Ferraro et al., (2016) because thanks to the change in the way second-hand clothes are considered, now it is fashionable to purchase from second-hand retail stores, while before these clothes were seen in a negative way.

Fashion can be very important because through clothes it is possible to show their inner personality and people use the way they dress to be part of a specific group of people. So, some consumers purchase vintage or retro clothing to highlight a difference from the masses and to follow the current fashion trend. (Laitala & Klepp, 2018) Some specific vintage clothes are very rare, because sometimes they were produced in limited editions, and also they can be very exclusive, because people need to find that piece by browsing around, and cannot just buy that in a store. These characteristics are connected with the authenticity and the originality of the piece, which not just gives an economic value but also a historical one. (Herjanto et al., 2016)

So second-hand consumers look at these pieces of clothing as premium clothes or treasures, because through them they feel part of the history of fashion and they are able to connect with the past. (McColl et al., 2013 cited in Herjanto et al., 2016) Also, if someone wears these vintage pieces is considered someone with a high fashion sense (Quayson, 2010, cited in Herjanto et al., 2016), something that can spread self-confidence and make people feel unique. (Herjanto et al., 2016)

These fashionability motivations are very relevant to consumers and this is highlighted by the research of Ferraro et al., (2016) which shows that most (83%) of second-hand consumers are driven by fashion in their second-hand store purchases.

1.4 Different retail channels

Zaman, Park H., Kim and Park S. H. (2019) stated in their research that it is possible to describe second-hand clothing shoppers with six different consumer orientations which are frugality, style consciousness, ecological consciousness, dematerialism, nostalgia proneness, and fashion consciousness. (Zaman et al., 2019)

According to McGuire, (1968, cited in Zaman et al., 2019) orientations of consumers can be used to determine their attitudes and behaviours, because they reflect their values and beliefs. Consumers' inner world pushes them to behave in a determined way and so these orientations can be used to segment second-hand clothing consumers. (Zaman et al., 2019)

In the context of second-hand, there are different types of retailers with different profiles in consumer orientations, so in the research, it is not possible to just focus on one. These retail groups can be classified into thrift stores, which sell clothes from local donations, consignment stores that are for-profit businesses, online stores (Zaman et al., 2019) and second-hand markets.

The authors theorised that every type of store, due to its different retail mix, can serve a specific segment of consumers according to their orientations and the items that they search for. (Zaman et al., 2019) The results of the research demonstrated that the three different types of retailers should be considered separately, due to the different profiles of their consumers. (Zaman et al., 2019)

In particular, among the different consumer orientations, frugality and style consciousness are significant for all shopper groups. Looking at the results of this research, the high level of frugality shows that this variable is the best predictor of second-hand shopping (Yan et al., 2015). At the same time, it is important for all second-hand stores to have some long-lasting pieces of timeless styles (Ferraro et al., 2016) to be able to satisfy the needs of style-conscious consumers. (Zaman et al., 2019) These different attitudes of consumers are very important for the different shops because it is possible to use the variety in consumer orientation to better target and satisfy actual and potential needs, developing new store strategies and new merchandise. (Zaman et al., 2019)

1.4.1 Thrift Stores

Thrift stores, also defined as charity shops are those shops that have a non-profit purpose in their activity. The clothes and items are donated by people or shops that want to contribute or liquidate their products and the sales in their retail outlets are used for charitable purposes, such as helping people in need. Due to the wide variety of thrift stores existing, the mission activities are varied and can vastly differ from one another. These shops, a part of charitable fundraising, raise awareness for their mission and provide clothes for low-income shoppers. (Mitchell & Montgomery, 2010)

According to the results of the research (Zaman et al., 2019) thrift shoppers tend to be characterized by significantly high values of dematerialism, which measure the resistance to overconsuming. So thrift stores need to have and display garments that fulfil the minimalist lifestyle and values, like basics and practical clothing, to follow the characteristics of their consumers. They should not forget to vary their offer with the presence of some unique pieces, for those consumers that search for exclusive styles. They can also promote their image of a place for the community, in their mission of offering affordable clothes that can be purchased by everyone. (Zaman et al., 2019)

1.4.2 Consignment shoppers

On the opposite line of charity stores, consignment stores can be defined as for-profit organizations, which are selling clothes provided by a third-party seller, often people vending their clothes. These shops are privately owned and widely differ from one another. Due to the profit purpose, the owners try to give them a unique style and the products are of high quality and highly 'curated'. (Evans et al., 2022) They tend to be independent and so it is their offer, dimension, style and price range. In this category, it is possible to find the vintage shops that sell those clothes of a particular style or era, that are positively valued by consumers.

For these reasons, consignment shoppers tend to be characterized by significantly high values of nostalgia proneness and fashion consciousness compared to thrift store consumers. Some strategies to attract their consumers can be to focus more on vintage clothing, play some retro music and use vintage themes in their shops, to increase the value of nostalgia proneness in their customers, that are mostly baby boomers and seniors, characterized by higher values of this variable. In their strategy it is important to remember to vary their offer, displaying also fashion and in-demand clothes, to attract and retain fashion-conscious consumers, the younger generations. (Zaman et al., 2019)

Also, consignment stores are more prevalently frequented by females (64%) than male consumers (36%) so the owners can develop some marketing strategies to satisfy this target. (Zaman et al., 2019)

1.4.3 Second-hand markets, flea markets

There is not wide research on second-hand markets or flea markets, due to their highly specific characteristics and their variability. (Sherry, 1990) In recent times, this sector is growing in

importance, due to the growth of second-hand consumption, while in the past it has been considered at the Bottom of the Pyramid. (Petrescu & Bhatli, 2013)
These markets tend to be extended as a shopping mall, where in the latter consumers buy utilitarian goods at fixed prices (Petrescu & Bhatli, 2013), and in flea markets the purchase follows long research and negotiation, so the price is not fixed. (Sherry, 1990). These second-hand markets represent an informal marketplace and economy, that in the past was mostly frequented by low-income shoppers, for the possibility to find low-cost goods, but now also other categories are becoming close to this sector. This is due to the possibility of having a new shopping experience, being surrounded by a varied offer of products and cultural immersion perceived by the consumer. (Petrescu & Bhatli, 2013)

1.4.4 Online stores

Online stores can be defined as e-commerce platforms (Evans et al., 2022) that divide into internet websites and internet communities, where the firsts are the online version of existing second-hand shops and the latter are virtual communities of users. These communities give everyone the possibility to sell and buy second-hand clothes without an intermediary and so the possibility to search, buy, sell and interact with others through the platform. (Padmavathy et al., 2019). In this context, the adoption of new emerging platforms such as Vinted, Vestiarie Collective and Depop is growing at fast rate between young generations. (Statista, 2022)

Due to the growth of online second-hand peer-to-peer/customer-to-customer (P2P/ C2C) e-commerce, that provides additional business opportunities to e-commerce companies, (Padmavathy et al., 2019) conceptualized, developed and validated a scale (OSSM) to measure online second-hand shopping motivation. In fact, the motivations of consumers for online shopping differ from the ones in an offline environment, even if there are some aspects in common. (Ferraro et al., 2016). An example can be the need for social interaction, which is more important for consumers that use offline channels. In particular, the new generations are moving their purchasing habits towards online channels, considered more comfortable and fast. (Xu et al., 2014)

Online shoppers tend to be characterized by significantly high values of ecological consciousness, nostalgia proneness and fashion consciousness. Looking at these results, some strategies to promote a more sustainable practice from these online websites can be to promote more sustainable packaging and discounts on bulk, to follow the desires of their consumers.

They can also develop an ecological mission that can be displayed on their website. (Zaman et al., 2019)

Due to the nostalgic inclination of their consumers, they can include on their websites, some sections for highlighting through storytelling, the authenticity and craftsmanship of their original designer products. (Zaman et al., 2019)

It is also important for them to target their fashion-conscious consumers, having seasonable fashion clothes. They can show, through some inspirations on their website, how consumers can combine the pieces and how can look together, to push them to buy the whole outfit. A strategy can also be to work with social media influencers, that can share positive feedback on wearing second-hand clothes. (Zaman et al., 2019)

Chapter 2: The Big Five Personality traits

2.1 Personality Theory

“Personality is a central concept in the social sciences because it speaks about people: who they are, how they come to be, and where they are hedging in their lives.” (Piedmont, 1998, pag. 1) The psychological characteristics give the individual a way to fulfil his needs and pursue his goals, furthermore they influence how we perceive and interpret the outside world. These characteristics help us to define ourselves as unique individuals, orient us toward the world and help us to adapt to it. (Piedmont, 1998)

“Personality can be defined as the intrinsic organization of an individual’s mental world that is stable over time and consistent over situations.” (Piedmont, 1998, pag. 2) Personality is a structured system, located within the person, through which individuals organize themselves. This system is not derived from the external environment and it is stable and consistent over life and different contexts, so the individual perception of the world is the same over time. (Piedmont, 1998) At the same time, personality can be defined as a dynamic structure, as something that develops over time, becoming more specific and unique through the course of life. In fact, the personality is constituted of two different aspects, genotype and phenotype. The genotype is the internal organization of latent dimensions of the individual, which guide him and the phenotype is the manifestation in a defined moment of these individual's characteristics combined. (Piedmont, 1998) The phenotypes, through the influence of culture, context and situations, can change and evolve with time, but the genotypes, which are the root of the behaviour, will always be the same. (Piedmont, 1998)
Over the years, psychologists tried to define the different dimensions and characteristics of the individual's personality (Gohary & Hanzaee, 2014) and how it is able to influence human behaviour. (Ajzen, 2005)

2.1.1 Theory of trait

Inside personality psychology, the development of the concept of trait had multiple explanations and multiple personality traits had been specified. The personality, a temporally stable, cross-situational difference between individuals, can be defined as a characteristic which has a strong influence on a wide range of trait-relevant responses. (Ajzen, 2005) Where traits of individuals are consistent and recurrent patterns of behaviours, which determine how someone act and react in the various situations and are able to differentiate individuals, providing an empirical generalization of some categories of behaviour. (McCrae & Costa Jr, 2008)

According to Allport (1937), different traits in different degrees are part of every individual and it is their interaction which gives everyone his unique behavioural and thought process. It is possible to state that, it is the sum of personality traits which determines the composition of someone's behavioural characteristics (Tsao & Chang, 2010) so, it is possible to do the inverse process, to look at the behaviour as manifestations of the personality. (Ajzen, 2005) Traits can be defined as the personality characteristics which are heritable and which individuals acquire genetically. Due to their origin, they are considered as stable over the lifespan of individuals. (Eysenck, 1990)

With time it is born an agreement about the existence of five super traits, which can enclose all the variability in individual’s personality. (Gohary & Hanzaee, 2014)

2.1.2 How to measure traits

These traits, which combined determine the personality of someone, are latent variables which is not possible to observe directly. The only way to study them is analyse the behaviours which they influence, asking the person himself, people close to him/her or an external observer. (Ajzen, 2005)

Even if, observing the behaviours is the best way to gain insights about the personality, this method is expensive and time-consuming. So, for practical reasons, are often used self-reports of behaviour or reports provided to people close to the individual. (Ajzen, 2005) Most of the time, these self-reports are questionnaires and multiple-choice responses, where the individual is asked to express his agreement or disagreement with a statement about his/her behaviour. (Ajzen, 2005) Noone has full knowledge about his trait profile and people search for a way of specifying it, for example looking at its manifestations through actions and experience. Individuals have a wider knowledge about their behaviours, their private thoughts, feelings, and desires than people close to them. For these reasons, they can have different opinions about their personality traits. (McCrae & Costa Jr, 2008)

These self-reports can produce a large variety of responses and they are not evaluative toward the individual, they just describe the tendency to have a certain behaviour in a specific context. (Ajzen, 2005) The responses, which reflect a specific personality trait, are generic and are not targeted to a specific group of individuals. This questionnaire approach has its focus just on the individual and it is useful to differentiate and to classificate of people into different personality traits groups (Ajzen, 2005)

There are very long lists of existent traits and most of them are redundant, but identifying the underlying dimensions, it is possible to summarize them. (Costa Jr & McCrae, 1992) There are various theories about the classification and the importance of traits, because a multitude of researchers is trying to define the main traits, which are able to describe the personality of individuals.

In particular, Eysenck (1990) formulated a three-factor theory, where the three dominant traits were Psychoticism (P), Extraversion (E) and Neuroticism (N). For being able to formulate this theory, he studied the genetic and the environmental contribution of individual differences, focusing on the studying of twins. (Eysenck, 1990).
I use in my research the theory of the Big Five Personality Traits, according to which exist five personality traits, which are the evolution and the convergence of all the previous theories on personality.

2.2 The Big Five Personality Traits

There is a convergence of studies which indicate the existence of five basic dimensions of personality, which are able to explain most of the variance in personality, even if with small differences in their definition. (Costa Jr & McCrae, 1992)

The Theory of the Big Five or the Five Factor Model of Personality is considered the best theory to explain the variability between individuals and their temperamental differences. This theory can represent a way to search for links between the personality and the behaviours of individuals. (Wiggins, 1996) These Big Five personality traits are able to describe comprehensively and parsimoniously the numerous and the most important phenotypic differences between individuals, through the use of a set of highly replicable dimensions.

(Wiggins, 1996) This model of personality attributes is derived from lexical analysis and it is a descriptive model, which illustrates the differences in behaviours which originate from the personality, rather than an explanatory model which gives a reason for these discrepancies in the behaviour of the individuals. (Wiggins, 1996) The Big Five are considered a dimension of the perceived personality of the individual because it relies on the individual's perception of their own aggregates of laypersons. (Wiggins, 1996)

These "Big-Five" factors have traditionally been numbered and identified as (I) Extraversion, (II) Agreeableness, (III) Conscientiousness, (IV) Emotional Stability (vs. Neuroticism), and (V) Openness. (McCrae & Costa, 1987)

The Big Five are a quite perfect comprehensive organizing structure for most of the personality attributes, in which each factor represents a major concentrations of attributes which are displayed in a continuum. (Wiggins, 1996) Traits are organized hierarchically from narrow and specific to broad and general dispositions, where the Big Five constitute the highest level of the hierarchy. (McCrae & Costa Jr, 2008) So, the Big Five Factors can be considered a broader personality dimensions or facets, where the others facets can be included in the Big Five or are blends of them. (de Raad & Perugini, 2002) The facets are very important, because the spectrum of each of the Big Five can’t be descripted just by the higher level. The overall structure of the Big Five doesn’t change if new facets are added to the factors in the model, but they provide more specificity and stratification which is not possible in the higher levels. (de Raad & Perugini, 2002)

Even if the various studies confirmed the convergence between the basic attributes of the different traits, there may be some differences concerning the descriptions of the factors for each trait, regarding their number, variants and the culture influences. (De Raad & Perugini, 2002)

2.2.1 History of the theory

In his research Cattell (1947) described thirty-five variables, which were indicated as a way to represent the different dimensions of personality. Since Cattell (1947) formulated a description of the personality using English trait terms, there were various alternative studies on other possible configurations. Later, Norman in 1963, after studying these traits, identified a significant subset composed of twenty out of these thirty-five variables, creating the Five-Factor Model.

Wiggins in 1968 identified extraversion and neuroticism as the more recurrent dimensions within the questionnaire approach and defined them as the “Big Two”. (cited in De Raad & Perugini, 2002)   The dimension of Openness to experience was added subsequently by Costa and McCrae (1976 cited in De Raad & Perugini, 2002)

It was Golderberg, some years later, in 1981, the first one to use the term Big Five to define these traits. Goldberg (1990) analysed the Cattel variables with an orthogonal rotational method to check the replicability of only five factors. This structure is stable over the variations in the number of factors rotated and this is a differential characteristic, because usually, factor structures are sensitive to the selection of variables analysed. (Goldberg, 1990) So, Goldberg (1990) confirmed the robustness of this structure, demonstrating that the core variables associated with each factor, if analysed with other variables, continue to have the same meaning and behaviour, within a comprehensive set of 1431 trait adjectives and within a set of 479 commonly used terms, collected through self and peer descriptions. (Goldberg, 1990) So from a sample of English trait adjectives, collected through self and peer descriptions, it is always possible to arrive to a variant of the Big Five model, because it represents all trait terms. (Goldberg, 1990) The subsequent big five are demonstrated as replicable and they can be considered as a general representation of the English trait lexicon which is more representative than other classifications from previous studies. (Goldberg, 1990)

Subsequently, the Five-Factor model of personality had its validity assessed by different authors. McCrae e Costa (1987), through two different data sources (self-reports and peer ratings) and two instruments (adjective factors and questionnaire scales), demonstrated the existence of substantial cross-observer agreement on the five adjective factors. Their results reinforced the usefulness of this five-factor model in personality research. (McCrae & Costa, 1987)

2.2.2 Characteristics of these traits

The trait perspective is based on some assumptions about the human nature and the Five Factor Theory assumes that people are knowable, rational, variable, and proactive. (McCrae & Costa Jr, 2008) The assumption of knowability of individuals considers possible to study their personality through scientific studies. The assumption of rationality concerns people as being capable to understand themselves and others, considering that individuals suffer of bounded rationality, so a rationality which is limited by the imperfections of human thought processes.

Variability asserts that individuals are different between them in psychologically significant ways and proactivity refers to the consequences of human actions, through which individuals search a way to control and influence their lives. (Funder, 1995, cited in McCrae & Costa Jr, 2008)

These Big Five Factors are considered the main model for describing personality, thanks to their characteristics of exhaustiveness in catching the semantics of personality and their use of ordinary language.

Given their robustness to the standards, these different five factors of personality are considered as equal, but they differ between themselves for their descriptive importance and replicability. (Wiggins, 1996) The most easily replicable ones are the first three factors, or Extraversion, Agreeableness and Conscientiousness (Saucier, 1995, cited in Wiggins, 1996) while the fifth one is the weakest one.

Their main characteristics are:

- **Comprehensiveness:** A valid personality model considers all individuals' important differences such as the emotional, interpersonal, experiential and attitudinal ones.
  The psychological approach is based on some leading principles, which ensure the relevance of traits. (De Raad, 2000) In fact, individual differences are endless, but most of these are minor and so they are not codified with specific terms. (Goldberg, 1990) In his study, Goldberg (1981) developed the lexical hypothesis, which guarantees the coverage of the whole trait domain.
  "Those individual differences that are the most significant in the daily transactions of persons with each other will eventually become encoded into their language. The more important in such a difference, the more people will notice it and wish to talk of it, with the result that eventually they will invent a world for it" (Goldberg, 1981, p. 141-142) This definition explains that, the more important individual differences in human interactions, are identified in some or all world's languages through single terms. According to the author these five factors are both necessary and sufficient to cover the space between all these words. (Goldberg, 1990)

- **Heritability:** From the assumptions of the theory, every individual has a common genome and the personality traits are perpetuated as genetic noise. So, these traits are heritable, because they partially originate in the individuals from some genetic influences. This is demonstrated by the research of Jang et al., (1996) which studied the personality traits of 123 pairs of identical twins and 127 pairs of fraternal twins. In this research, the broad genetic influence and so the heritability of each trait, was discovered to be equal to 53% for extraversion, 41% for agreeableness, 44% for conscientiousness, 41% for neuroticism, and 61% for openness to experience. From these results, the genetics influence a wide proportion of the variance of these traits, but the majority of variance still derives from non-shared environmental influences. (Jang et al., 1996)

- **Universality:** The personality of an individual is affected by the context and by the culture in which he is, but the basic structure of personality is the same for everyone. (Costa Jr & McCrae, 1992) In fact, these traits are a function of biology, meaning that the structure of personality is universal. (McCrae & Costa Jr, 2008) This assumption is confirmed by the results of numerous studies, which showed the existence and similarity of these traits in various cultures. (McCrae & Costa Jr, 2008) In their study Weisberg et al., (2011), analysing the gender differences across ten different aspects of the Big Five, concluded that women tend to have higher values of the personality traits of Extraversion, Agreeableness and Neuroticism compared to men. These differences don't depend from the characteristics of the country of the study (Weisberg et al., 2011), in fact, the universality applies also to gender differences, which are not influenced by the different gender roles across cultures. (McCrae & Costa Jr, 2008)

- **Longitudinal stability:** According to the postulated of the FFT of McCrae and Costa Jr. (2008), individual traits differences are not perfectly stable, to be more precise, just around 45% of the variance in the personality traits is stable across the adult lifespan. The personality development is determined by biological maturation, which is mostly around the first third of life and not by life experience. (McCrae & Costa Jr, 2008) So, the personality of individuals, can go though some small changes over their course of life. (Costa Jr & McCrae, 1992) Donnellan and Lucas (2008) studied the cross-sectional age differences in the Big Five personality traits across respondents from 16 to around 80 years old. The results showed that older individuals tend to be less extraverted, less neurotic and less open to experience, but also more agreeable and conscientious compared to younger individuals. (Donnellan & Lucas, 2008) It is also possible to observe some declines in some traits over long periods, but the mean levels are mainly constant. (Costa Jr & McCrae, 1992)

2.2.3 Extraversion

The first pair of traits, defined as Extraversion and Introversion, belong to the interpersonal dimension and it is about the quality of social interactions of the individual. (Tsao & Chang, 2010) These traits are relevant in a variety of contexts and it is the pair of constructs more studied, as highlighted already in 1930 by Guilford and Braly. (cited in De Raad, 2000) These two traits are a multifaceted construct, so they can decline into different situational adaptations and behaviours, depending on the context in which the individual is. What is important to remember is that the interpretation of this trait depends on the different taxonomies which are used to describe it. (De Raad, 2000)

According to Jung (1917), people with high levels of extraversion direct their psychic energy toward the external world, while people with high levels of introversion direct this energy toward their internal world. (cited in De Raad, 2000) This trait, like the others, is not absolute, so people do not divide into two separate groups with opposite characteristics. Instead there is a variability of characteristics on the line that connects the poles of extraversion and introversion. (De Raad, 2000) Keeping in mind this continuo, people characterized by high values of extraversion are often comfortable in socialization contexts, are confident toward what they don't know and they are interested in the external world, people and events. On the other side of the continuo, high introvert's individuals tend to be more interested in their psyche and prefer to be alone and in a safe space, like their own house. (De Raad, 2000)

Considering the Five Factor Model by Norman (1963) someone extroverted is talkative, frank, open, adventurous and sociable whereas introverts are silent, secretive, cautious and reclusive. Extraversion, also defined as gregariousness, can be described as the inclination of individuals to be in companionship and in those situations characterized by social stimulation. So extrovert individuals tend to have social skills, lots of friends, tend to participate in team sports and are members of different clubs. (McCrae & Costa, 2008) Also, the dimension of extraversion encloses a broad group of traits, such as the inclination to live with positive emotions such as happiness and sociability (Costa Jr & McCrae, 1992) and to have personality characteristics like being affectionate and loving, conversational, someone who loves fun and who is passionate. (Costa Jr & McCrae, 1992) To sum up this complex trait, it is possible to state that liking to socialize is an example of the personality trait of extraversion. (Fujiwara & Nagasawa, 2015)

Whereas the opposite trait, defined as introversion or social isolation, is the tendency to be more attired by the internal world than the external one. (Costa Jr & McCrae, 1992) So, individuals characterized by low levels of extraversion tend to be shy, reserved and not strong in socializing in all the different contexts. (Fujiwara & Nagasawa, 2015) They can be loner, quieter and more passive because they do not search for fun or experiences of socialization. (Costa Jr & McCrae, 1992)

2.2.4 Agreeableness

The second of the Big Five Personality Traits is named Agreeableness and it is also concerned with interpersonal relationships. On the contrary of extraversion, it is the personality trait with the smallest history.

In his research, Wiggins (1991) theorized that individuals with high degrees of agreeableness are dominated by a sense of "communion", the membership of a larger community, which can be spiritual or social. People who feel part of this community tend to seek intimacy, a sense of union and solidarity with this larger group. (cited in De Raad, 2000) For this reason, they are characterised by kindness, altruism and cooperativeness and they tend to be empathic toward others and their needs. Furthermore, agreeable individuals have high capabilities of emotional self-regulation, so they tend to experience less strong negative emotions, like anger or fear. (McCrae & Costa, 2008)

According to the American-English five-factor structure, individuals who have high levels of agreeableness, friendly compliance and socialization levels can be defined as soft-hearted, trusting, generous, acquiescent, (Costa Jr & McCrae, 1992) understanding, collaborative, tolerant, and gullible. (McCrae & Costa, 2008) So, people with high levels of agreeableness tend to be less selfish and stubborn and more kind and honest, compared to ones with lower levels of this dimension. (Fujiwara & Nagasawa, 2015)

Due that agreeable people need harmonious relations with others (Zurawicki, 2010, cited in McCrae & Costa, 2008) and are not searching for interpersonal conflict, in a discussion, they tend to comply and yield to others' opinions. They also tend to forgive others' behaviours, which they don't like and they don't use offensive language. (McCrae & Costa, 2008) This can lead them to be victim of harassment by others, (McCrae & Costa, 2008) also because these people are more trusting toward the external environment and see less the danger. (Tsao & Chang, 2010)

Instead, people characterized by high levels of the opposite trait, defined as antagonism, are individuals who can be defined as ruthless, suspicious, critical and cynical. (Costa Jr & McCrae, 1992) They are not trusting and they doubt what they don't know. (Gohary & Hanzaee, 2014) For this they tend to be characterised by hostility, insensitivity and indifference. (Caprara et al., 1994)

2.2.5 Conscientiousness

The third trait of the big five, called conscientiousness, is the more important trait in those situations and contexts in which achievement is considered fundamental, such as educational and working environments. For its individual and societal relevance, conscientiousness is the trait with the highest research growth in recent times. (De Raad, 2000)

Conscientiousness can be described as the aim to accomplish something by the individual. People with high values in this trait are often strongly motivated in reaching their goals, so they can be very systematic in their work, organized, efficient, practical and consistent in their efforts. (De Raad, 2000) This trait can be declined as a will to achieve something and describe those individuals who are hardworking, punctual, ambitious, persevering and diligent. (Costa Jr & McCrae, 1992) People can be very conscientious in their working environment, because they may search a personal satisfaction in these circumstances. This can derive from a difficulty in the interpersonal context, which can in turn originate from their low sensitivity. (Tsao & Chang, 2010)

The opposite trait, defined as Undirectedness, characterizes people who are negligent, lazy, disorganized, late, aimless and quitting. (Costa Jr & McCrae, 1992) Some possible facets to describe this trait are competence, order, dutifulness, achievement striving, self-discipline and deliberation. If the individual has high values of conscientiousness is characterized by high values of these facets, otherwise is described by their opposite. (Caprara et al., 1994)

2.2.6 Neuroticism

The fourth personality trait, Neuroticism or negative affectivity, can be defined as the tendency to experience psychological distress, so individuals with high levels of this trait tend to be emotionally labile. (Gohary & Hanzaee, 2014) This trait has a negative denotation because it implies that people with high values of this personality trait tend to be always or frequently worried, temperamental, self-pitying, self-conscious and emotional. (Costa Jr & McCrae, 1992)

So neurotic individuals tend to be weaker against emotional and continuous sufferance, furthermore they have stronger and longer negative emotional responses. (Tsao & Chang, 2010) According to the study of Costa Jr. & McCrae (1992), with increasing levels of neuroticism, individuals tend to perceive more vividly and more frequent conditions such as vulnerability, depression, and anxiety and they tend to be more impulsive.

These higher levels of anxiety can lead them to suffer from some bodily aches, such as headaches and stomach difficulties. Furthermore, they tend to be hypochondriacs, insecure and inadequate. For these reasons sometimes people with high levels of neuroticism may present some elements or suffer from some psychiatric conditions. (Costa, Jr & McCrae, 1992)

The opposite trait of neuroticism is defined as emotional stability (Goldberg, 1990) and it is its positive pole. It can be evaluated as a resource for individuals, especially in environments which tend to be stressful or demanding. (De Raad, 2000) This emotional stability leads these individuals to be dominant and secure (Caprara et all., 1994). As a consequence, this trait is important in organizational contexts because it characterizes individuals who are serene, even-tempered, flexible and strong, who are considered as a resource for a company. (Costa Jr & McCrae, 1992)

This trait can be useful in analysing someone’s achievement in contexts which involve challenges in learning and education. People can be positioned on an imaginary line, where on one side there are individuals who behave in a maladaptive way and adopt a helpless style when a problem appears and this behaviour can stop them from solving the situation. On the other side, there are people who see possible problems as challenges and are searching them as a way to overcome obstacles. They are able to succeed problems, following strategies which come from their mastery-oriented and adaptive response style. (Dweck, 1986, cited in De Raad, 2000)

2.2.7 Openness to experience

The fourth trait, called Openness to Experience, is substantially related to what Goldberg (1990) defined “Intellect” and Norman (1963) defined “Culture”.

Some characteristics of people with high values of this trait are being very imaginative, creative, original, liberal, (Costa, Jr & McCrae, 1992) intellectually curious and very sensitive to art and beauty. (Gohary & Hanzaee, 2014) Due to their curiosity, they are inclined to consider different points of view and opinions and furthermore they prefer variety. (Tsao & Chang, 2010) This leads them to be less prejudiced and to try new experiences, which they see as an opportunity to learn and discover new things. (Tsao & Chang, 2010) For these reasons, individuals with high values of this trait tend to be open to new ideas, values and beliefs, this can lead them to be less conventional. (Gohary & Hanzaee, 2014)

According to Matzler et al. (2006) individuals open to experience, have richer emotional lives because they go through positive and negative emotions in different contexts. Their curiosity about the inner and the outer world (Gohary & Hanzaee, 2014) can led them to be more behaviourally flexible and able to adapt in various situations. (Costa, Jr & McCrae, 1992)

The opposite pole of this trait, defined as Closedness, describes people who are down to earth, uncreative, conventional, uncurious, conservative in thoughts and actions. (Costa Jr & McCrae, 1992) These people don’t have many interests (Gohary & Hanzaee, 2014) because they are not open to new things and possibilities and are comfortable in their routine.

2.3 Consumption behaviour

The Big Five Personality Traits are accepted as relevant and valid dimensions of the nature of individuals, in different fields of research. (De Raad & Perugini, 2002). The research about this theory is essential because the individual’s purchase motivation, toward different products and sectors, is influenced by their personality and their psychological state. So, part of the variability, between purchase choices and behaviours of consumers, can be explained through their personality attributes. (Tsao & Chang, 2010) This is because, our traits are the sources of our judgments and choices (Gohary & Hanzaee, 2014) and often they reflect someone’s values and preferences. (Tsao & Chang, 2010)

Previous research provides evidence that these traits represent observable differences in the behaviours of individuals, because the behaviour is a result of their interaction with the characteristics adaptations. (McCrae & Costa Jr, 2008) The latter are the individual’s habits, attitudes and skills, which in turn are shaped by the interaction of the personality disposition with the social and physical environment. They differ between contexts, cultures, families and differently from the personality traits, they change in order to adapt to different situations. (McCrae & Costa Jr, 2008)

The study by Gohary and Hanzaee (2014) analyse the relationship between the Big Five personality traits and the different purchase motivations and behaviours of consumers, such as shopping to satisfy hedonic and utilitarian values or compulsive and impulsive shopping behaviour. The personality traits of Conscientiousness, Neuroticism and Openness have a positive relationship with compulsive buying, impulsive buying and utilitarian values, where individuals characterized by high values of the traits of conscientiousness and agreeableness tend to have high utilitarian values and low compulsive and impulsive shopping behaviour. Consumers with higher values of the trait of neuroticism tend to follow impulsive buying behaviours more often, while the trait of openness is negatively correlated with impulsive and compulsive shopping behaviours. This trait is instead a significant predictor of utilitarian values. The trait of extraversion has a significant positive relationship with the hedonic and utilitarian values of shopping, while it has a negative relationship with compulsive shopping and buying.

Matzler et al. (2006), analysing two of the Big Five (Extraversion and Openness to experience) and their relationship with hedonic values, brand affect and loyalty, researched the individual determinants of brand affect. Checking if the individual’s differences across these traits, it is possible to see that these traits are positively related to hedonic product value. So, these traits are able to influence, directly in the case of openness and indirectly in the case of extroversion, the brand affect, which drives attitudinal and purchase loyalty.

According to the study of Turkyilmaz, Erdem, and Uslu (2015), all the Big Five personality traits, combined with the external factor of website quality, can influence the values of impulsive online buying, defined by Beatty and Ferrell (1998, cited in Turkyilmaz et al., 2015) as an unexpected and instantaneous purchase by a consumer, that had no pre-shopping intentions. In particular, the traits of neuroticism and conscientiousness negative influence the values of online impulse buying, while the effect of the traits of agreeableness, extraversion and openness to experience is positive.

In their study, Fujiwara and Nagasawa (2015) analyse the effect of the Big Five personality traits on the development of purchase intentions for some luxury brands for Japanese people.

Their results showed that the personality trait of conscientiousness is not significant in influencing the purchasing choices of individuals, while the trait of neuroticism has a negative influence on the purchasing intentions of cars from the luxury brands. Whereas, following the hypotheses, the personality trait of Openness to Experience is positively related to purchase intentions for luxury goods, which can be seen as unique pieces of high artistic value.

Tsao & Chang (2010) explored the impact of the Big Five personality traits on the purchase behaviour of e-shoppers, consumers who shop online, and how these traits can influence and determine the hedonic and utilitarian purchase motivations of consumers. Three of the Big Five, neuroticism, extraversion and openness to experience, have a positive relationship with the hedonic purchase motivation and so, people characterized by high values of these traits, search for entertainment during their process of online shopping. Furthermore, when people are characterized by high values of the personality traits of Neuroticism, Agreeableness or Openness to Experience, they follow their utilitarian values when they shop online.

Lissitsa and Kol (2021) studied the relationship between the Big Five personality traits and the m-shopping intentions for hedonic products where m-shopping refers to mobile commerce, the activity to purchase products through mobile devices. The results of the research showed the existence of positive correlations between m-shopping intention and the personality traits of extraversion, openness to experience and conscientiousness for some generational cohorts and the existence of negative correlations with the trait of neuroticism.

Furthermore, Hirsh (2010) studied the link between personality characteristics and environmental concern, because people can vary considerably in their attitudes toward environmental issues. It was highlighted that, higher levels of Agreeableness and Openness in the individuals, were linked to greater environmental concern, showing a relationship between the two. Whereas for the traits of Neuroticism and Conscientiousness this relationship, even if still positive, was less strong.

Milfont and Sibley (2012) studied the associations between the Big Five personality traits and the environmental engagement, both at individual and at national level. From their results, the traits of agreeableness, conscientiousness and openness to experience are able to explain the variability of environmental engagement, both at individual and national level.

The study of Sun et al. (2018) explored the effects of individual’s personality traits on consumer’s attitude toward green buying and toward the intention to buy green products.
The results indicated that the personality traits of extraversion, agreeableness, openness to experience and conscientiousness positively affect consumer’s attitude toward green buying. In turn, consumer’s attitude, conscientiousness, openness to experience and extraversion affect consumer’s intention to buy green products positively and significantly.

According to the research of Duong (2022), the avoidance of purchasing products which have negative impacts on the ecosystem and society, known as green consumption (Jaiswal & Kant, 2018, cited in Duong, 2022) is strongly associated with some of the Big Five personality traits. These traits have significantly different effects on green consumption attitudes and on the intentions of pro-behavioural consumption of individuals. In particular, the personality trait of agreeableness is positively associated with attitude towards green products and with green purchase intention. Whereas, the trait of consciousness, even if is positively associated with attitudes towards green products, it is not correlated with green purchase intention. The trait of openness to experience has no correlation with green purchase intention and it is negatively associated with attitudes towards green products. The contrast of this results with the previous research, it is explained by the fact that in Asian countries, where the analysis was conducted, creativity and innovation are restricted.
These previous analyses confirmed the hypothesis that, part of the variability across motivations, attitudes and actual behaviours of consumers can be explained through their unique combination of the various values of the Big Five personality traits. In the following chapters, it will be explored the possibility to explain, the various motivations to purchase second-hand clothes, the attitude toward this sector and the frequency of purchase of consumers, taking in consideration their personality.

Chapter 3: Analysis of the data

In this chapter, we described the procedure of data collection and also the various variables, through the analysis of their bar plots, in order to gain insights about their distribution. Furthermore, we conducted a cross-correlation analysis of all the variables included in the study.

3.1 Data collection

In order to understand the influence of the Big 5 Personality Traits on the motivations for buying second-hand clothes and the variables which are able to explain the Buying Behaviour of consumers who purchase second-hand, I developed an online questionnaire, with structured and closed questions, targeted at individuals who bought a piece of second-hand clothing at least once in their lives. The sample is a non-probability one, where respondents were found through "convenience" methods. The questionnaire was diffused contacting individually possible respondents, through social networks such as Facebook and Instagram and through messaging services such as WhatsApp and Telegram.

In the process of identifying the form and layout of the questionnaire, I divided it into several parts, one for each of the different constructs analyzed. After the introduction, I placed in the following order the filter questions, the descriptive ones, the questions corresponding to the research objectives, with the most sensitive ones at the end and the demographics questions.

The first question is a contingency one that puts an end to the questionnaire for individuals who never purchase second-hand clothes. The other questions are designed to overcome the unwillingness to answer and are multiple-choice questions. I measured the research constructs with multi-item 5 or 7 Likert scales that ranged from "1= strongly disagree" to "5/7= strongly agree" taken from the existing literature on the topic of interest.

Measuring the Big Five Personality Traits has always been challenging, due to the difficulty of using a standard scale. Due to the lack of enough time to analyze these personality traits in detail, with many questions, I decided to use the BFI-10, the short form with 10 items of the Big Five Inventory developed by Rammstedt and John (2007). The respondent is asked to describe himself expressing agreement or disagreement in 10 sentences over a 5 Likert Scale.

To measure the reasons why consumers purchase second-hand clothes, I used the motivations scales reported by Guiot and Roux (2010) to measure the economic, critical and hedonic and recreational motivations which influence the purchase. I adapted these scales making them more specific about clothes and expanding their applicability also on online purchases. From the article of Ferraro et al. (2016), in which they explore the role of fashionability, I borrowed the scale of the fashionability motivation. To make this scale more inclusive and to better describe this dimension I integrated some questions from Sprotles & Kendall (1986) and Lee & Kim (2017), adapting them to the context of the research.

To measure the intention to buy second-hand clothes and second-hand clothes buying behaviour I took inspiration from the scales from Ögel (2022) about the same construct.

The final section of the questionnaire included some socio-demographic questions divided into segments for facilitating their answers about the gender, generation and monthly discretionary income of the respondents and an open question about their nationality.

The bugs in the questionnaire were eliminated through pretesting with three respondents similar to those who are included in the actual survey.

3.2 Data description

For a response rate of 65,2%, out of 167 responses, a total of 109 completed and usable questionnaires were used in the analyses conducted through the software RStudio.
The aim of the research is to create a broader paradigm that can explain the drivers and the reasons for the expansion second-hand clothing purchase and whether the differences in the purchase motivations of the individuals are determined by a combination of personality traits.
Before starting the analysis, it is important to perform some basic business analytics on the different variables. The univariate representation of the variables and the possible cross-correlations between them are useful to identify possible outliers and observe their distribution. So I conducted one-way tabulations for each variable in the study.

3.2.1 Socio-demographic variables

Looking at Figure 3.3 and Figure 3.4, it is possible to state that the sample is not exactly representative of the real population, having a strong gender bias towards females. In fact, 76 of the respondents are female (69,7%), only 29 are male (26,6%) and 4 are binary people or people belonging to other categories (3,7%). This inequality in the gender frequency of individuals can affect the validity of the analysis of aggregated data.

The sample is also not well-balanced in terms of the age of the respondents. The individuals are primarily from Generation Z (84,4%, 92 respondents), less than 10% (8,3%, 9 respondents) are from Generation Y, just 2 respondents are from Generation X (1,8%), and the remaining 6 individuals (5,5%) are from previous generations.

Regarding the higher educational level completed by the respondents, for around half of them (56%, 61 respondents) it is a Bachelor's Degree, while for 22,9% of them is Primary or secondary school and for the remaining 23 individuals (21,1%) is higher than the Bachelor's Degree.

Looking at Figure 3.1 and Figure 3.2, around half of the respondents (49,5%, 54 individuals) belong to the lowest monthly income category of under 500 euros a month. The second biggest category with 22 respondents (20,2%) is the one with a monthly income that is between 500 and 1000 euros per month, 13 individuals (11,9%) have a monthly income between 1000 and 1499 euros and 15 of them (13,8%) have their income instead included between 1499 and 2499 euro. Just 5 of the respondents (4,6%) have a monthly income greater than 2500 euros. Due to the high number of categories, I transformed this categorical variable Monthlyincome into a binary variable with two categories, "Low income" and "High income", where the first one includes the values "Less than €500" and "€500 - €1000", while the latter all the other categories. Looking at the distribution of this new binary variable in Figure 3.3 and 3.4 the category “Low income” has a relative frequency of 36,7% (40 individuals) and the category “High income” of 63,3% (69 individuals).

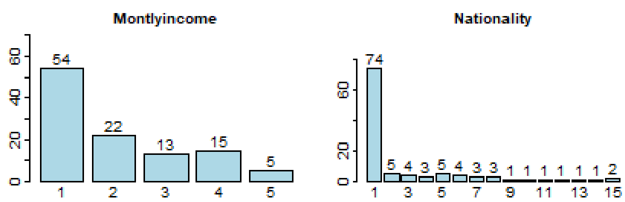

Figure 3.1 Bar Plots of the variables “Monthly income” and “Nationality”

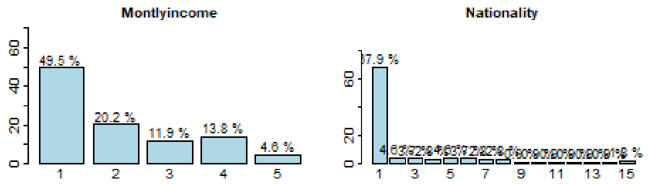

Figure 3.2 Bar Plots of the variables "Monthly income" and "Nationality" with percentuals

From Figure 3.1 and Figure 3.2, the distribution of the nationality differences in the sample are not well-balanced even if there is a lot of variability in the answers. In particular, the distribution is mostly focused, in fact, 74 individuals (67,9%) are Italian and the rest of the individuals are split into fourteen other nationalities with low frequencies, all under 5%.

Due to the high number of categories and their low frequency, I transformed the categorical variable Nationality into a binary one where "Italian" is a category and all the other nationalities form another category renamed "0". From the Figure 3.3 and Figure 3.4, the category "Italian" has a frequency of 67,9% (74 individuals) and the other category of 36,7% (40 individuals).

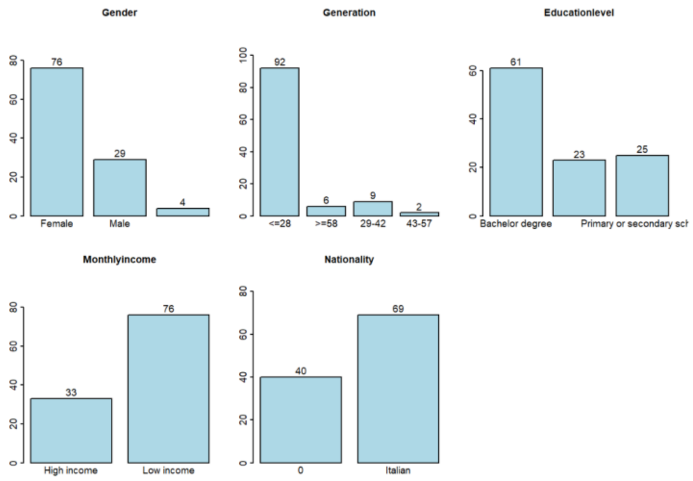

Figure 3.3 Bar Plots of the socio-demographic variables

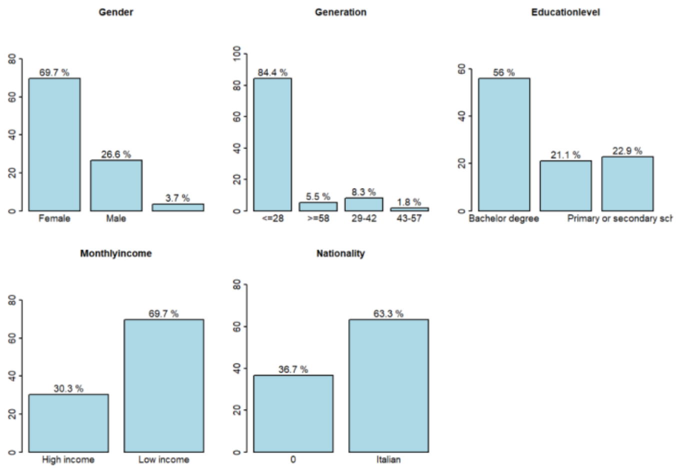

Figure 3.4 Bar Plots of the socio-demographic variables with percentuals

3.2.2 Buying Behaviour

Looking at the Figure 3.5 and Figure 3.6, it is possible to observe that, the two variables which measure the Buying Behaviour of the consumers, have most of their distribution concentered between the values of 4 and 5. These values correspond to favorable considerations of this form of shopping, so this highlights the preference of this sample of consumers toward second-hand clothes purchases.

I create another variable called "BuyingBehaviour" which is the sum of the two variables extracted from the questions. I also create a dichotomous variable, named BB_bin which takes the value 1 if the values of Buyingbehaviour_1 and Buyingbehaviour_2 are equal or bigger than 4, while otherwise, its value will be 0. In this way, it is possible to conduct the analysis using logistic regression with a binary dependent variable, codified in 0 for lower values of Buying Behaviour and in 1 for higher values. This variable has a relative frequency of 22,9% for low values of buying behaviors and 77,1% for high values of buying behaviour.

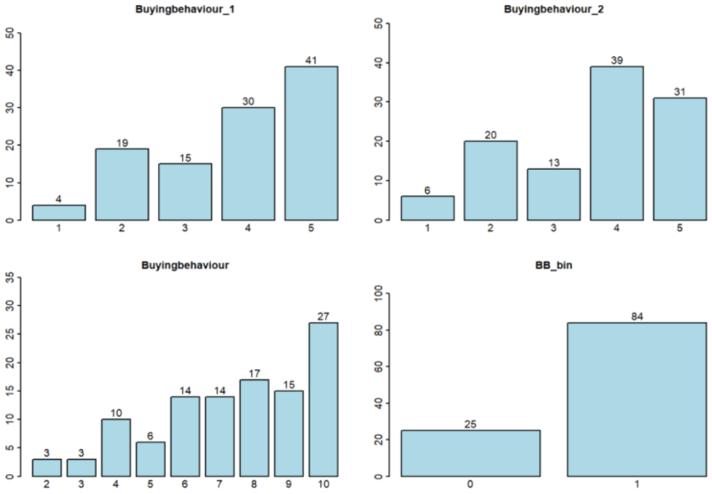

Figure 3.5 Bar plots of the variables of Buying Behaviour

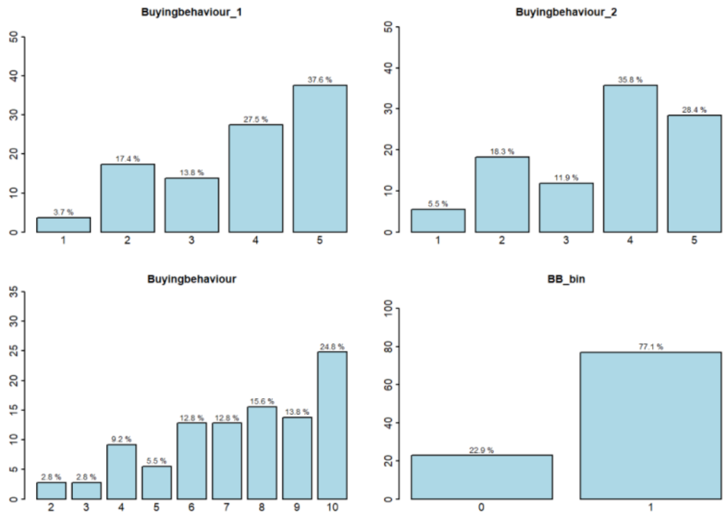

Figure 3.6 Bar plots of the variables of Buying Behaviour with percentuals

3.2.3 Critical motivations

Every aspect under analysis is measured through two different questions which are codified into two distinct variables. I sum these values to create a new variable for every aspect and I sum all the different variables about the single motivations to create variables for some broader categories such as Critical motivations.

The original variables are measured on a 7-Likert scale, and it is possible to see from Figure 3.7 and Figure 3.8, that most of the frequency is between the values of 5 and 7, which correspond to a high agreement from the respondents that they purchase second-hand for critical motivations.

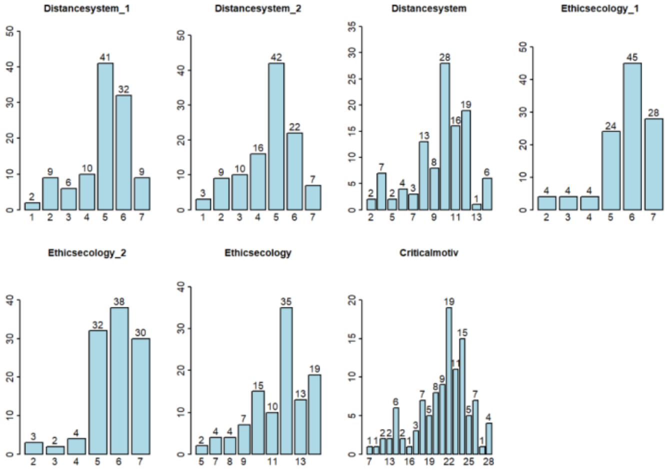

Figure 3.7 Bar plots of the variables of critical motivations

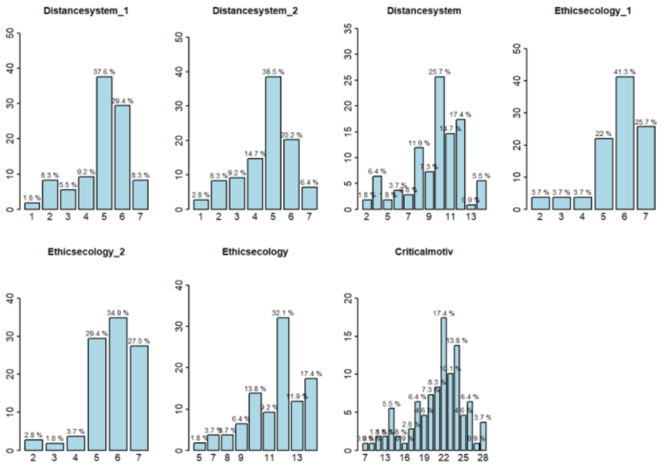

Figure 3.8 Bar plots of the variables of critical motivations with percentuals

3.2.4 Economic motivations

From Figure 3.9 and 3.10, the distribution of frequency tends to be skewered toward the higher values of these variables, in fact, the frequency for the value 1 is under 3% for all the different variables.

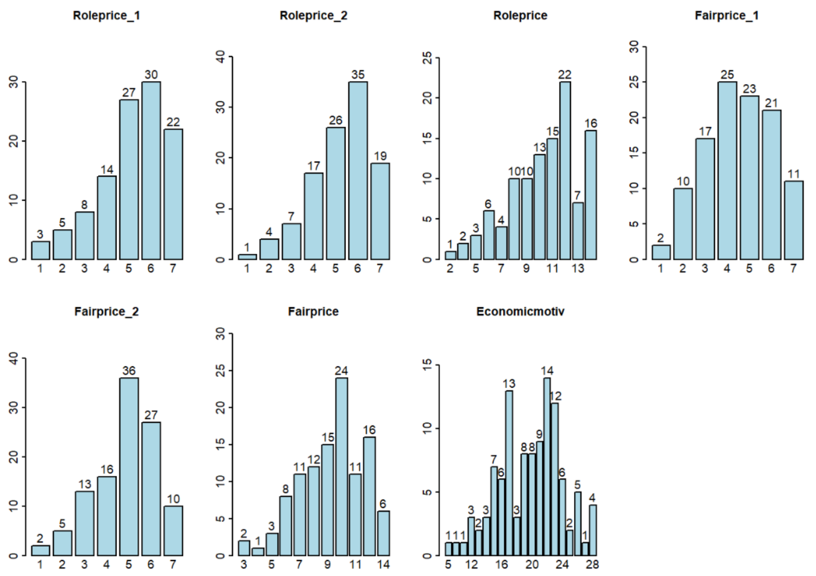

Figure 3.9 Bar plots of the variables of economic motivations

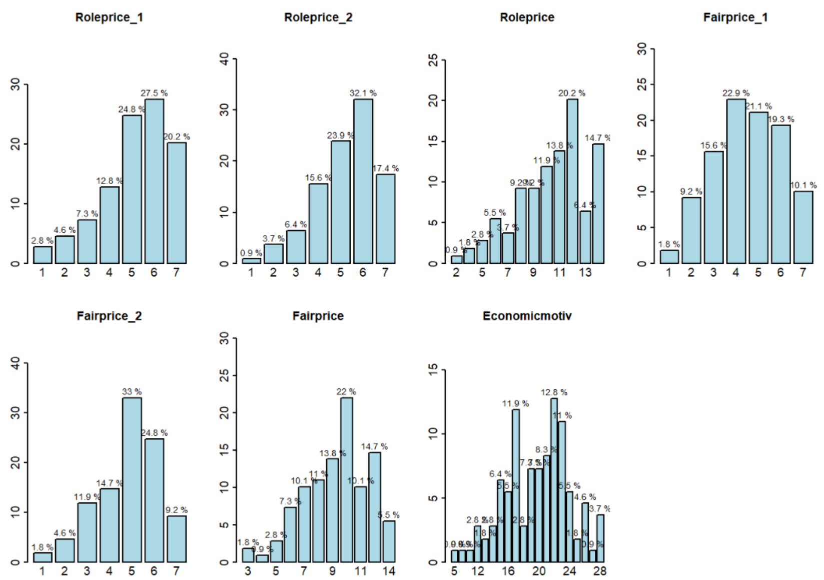

Figure 3.10 Bar plots of the variables of economic motivations with percentuals

3.2.5 Hedonic motivations

From Figure 3.11 and Figure 3.12, the distribution of frequency tends to be skewered toward the higher values for the variables which measure the motivations of treasure hunting, originality and nostalgic as reasons to purchase second-hand clothes. For the variables about the social contact motivation, the distribution is more balanced and it is present the opposite trend compared to the other variables. In this case, the higher value has a very low frequency, while for the others it was the lower value which had a frequency under 3%. Figure 3.11 and Figure 3.12 show that people tend to be driven positively toward the purchase by these hedonic motivations, except for the possibility to have a social interaction with the buyer, which instead negatively influences their purchases.

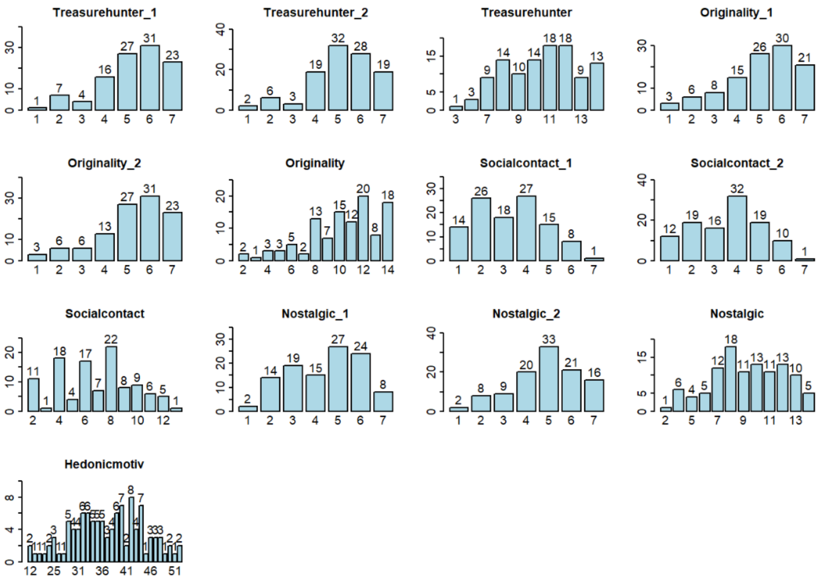

Figure 3.11 Bar plots of the variables of hedonic motivations

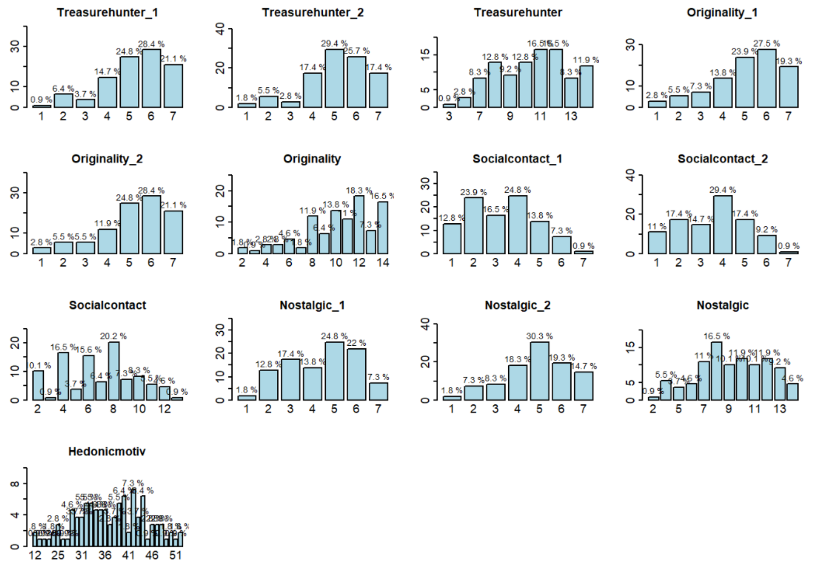

Figure 3.12 Bar plots of the variables of hedonic motivations with percentuals

3.2.6 Fashionability motivations

Looking at Figure 3.13 and Figure 3.14, it is possible to see that the relative distribution of the fashionability motivations to purchase second-hand clothes is mostly between the values of 5 and 6, while for the highest value 7, it is unexpectedly low.

From Figure 3.13 and Figure 3.14, the distribution of the frequency of the purchases, where higher values correspond to a higher frequency, shows that almost no one of the interviewees (0.9%) purchased weekly codified as 5, while most of them purchase quarterly (29,4%) or once a year (32,1%). Furthermore, there is not a unique more frequented channel for the consumers interviewed as highlighted by the spread of the frequency between the different types. It is possible to say that the second-hand stores are the most frequented in this sample with 35,6% of the consumer percentage, while the charity stores are the least frequented with a frequency of 14,7%.

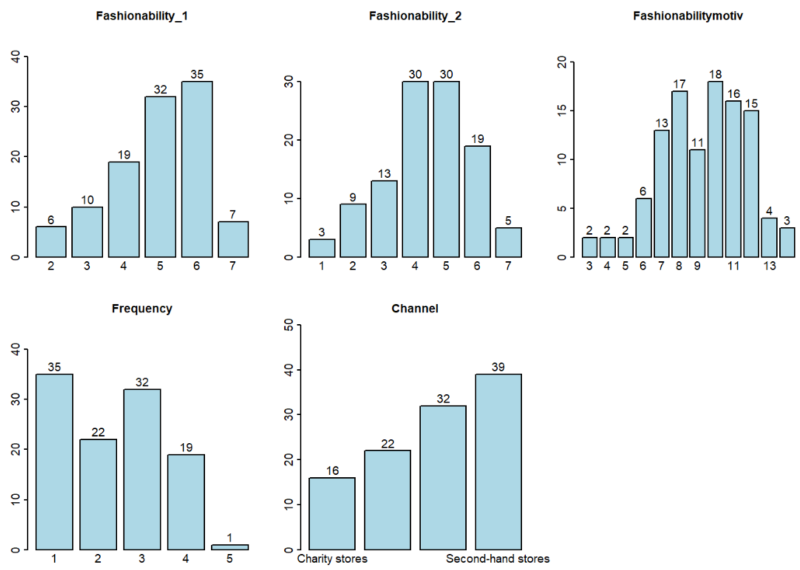

Figure 3.13 Bar plots of the variables of Fashionability motivations, Frequency and Channel

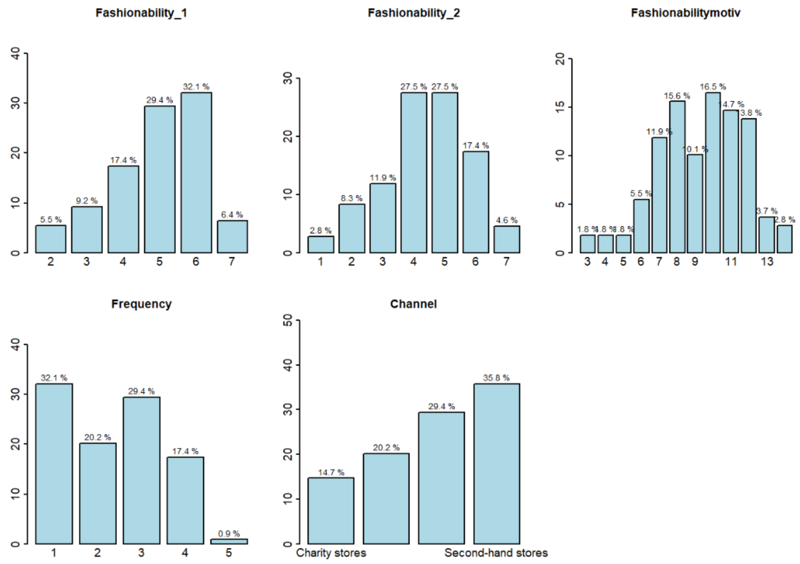

Figure 3.14 Bar plots of the variables of Fashionability motivations, Frequency and Channel with percentuals

3.2.7 The Big Five Personality Traits

Comparing the distribution of frequencies between the two variables which measure each personality trait, showed in Figure 3.15 and Figure 3.16, it is possible to see that, except in the case of neuroticism, the distributions are different or opposite. This can be a problem in the analysis because the distributions for the same traits were supposed to be similar or equal.

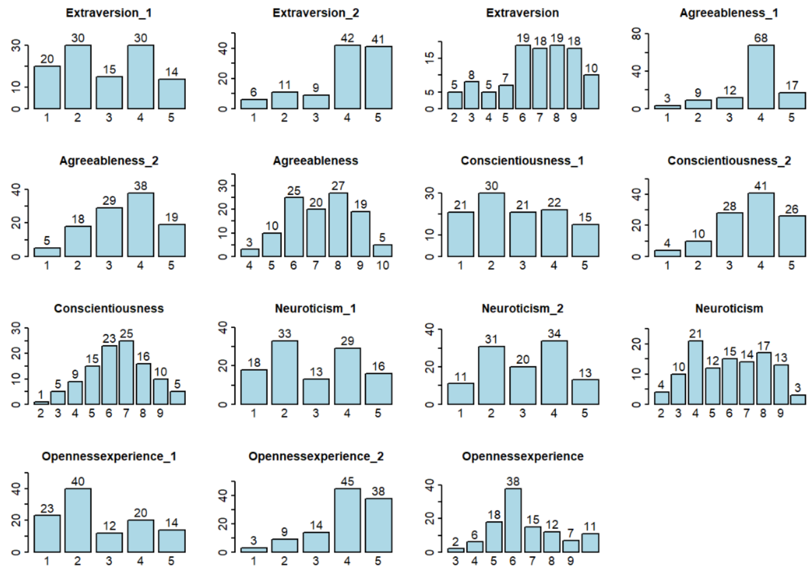

Figure 3.15 Bar plots of the variables of Big Five Personality Traits

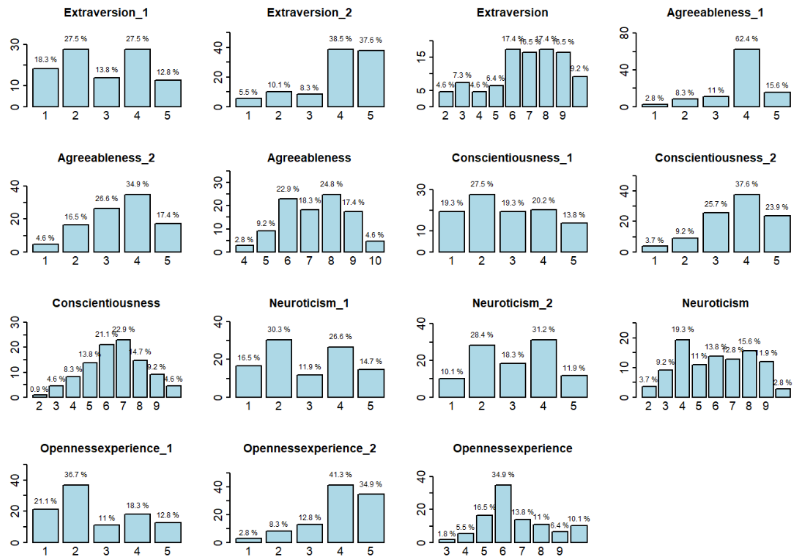

Figure 3.16 Bar plots of the variables of Big Five Personality Traits with percentuals

3.3 Cross-correlations between variables

Cross-correlations between variables are used to check the presence of a significant relationship between the variables.

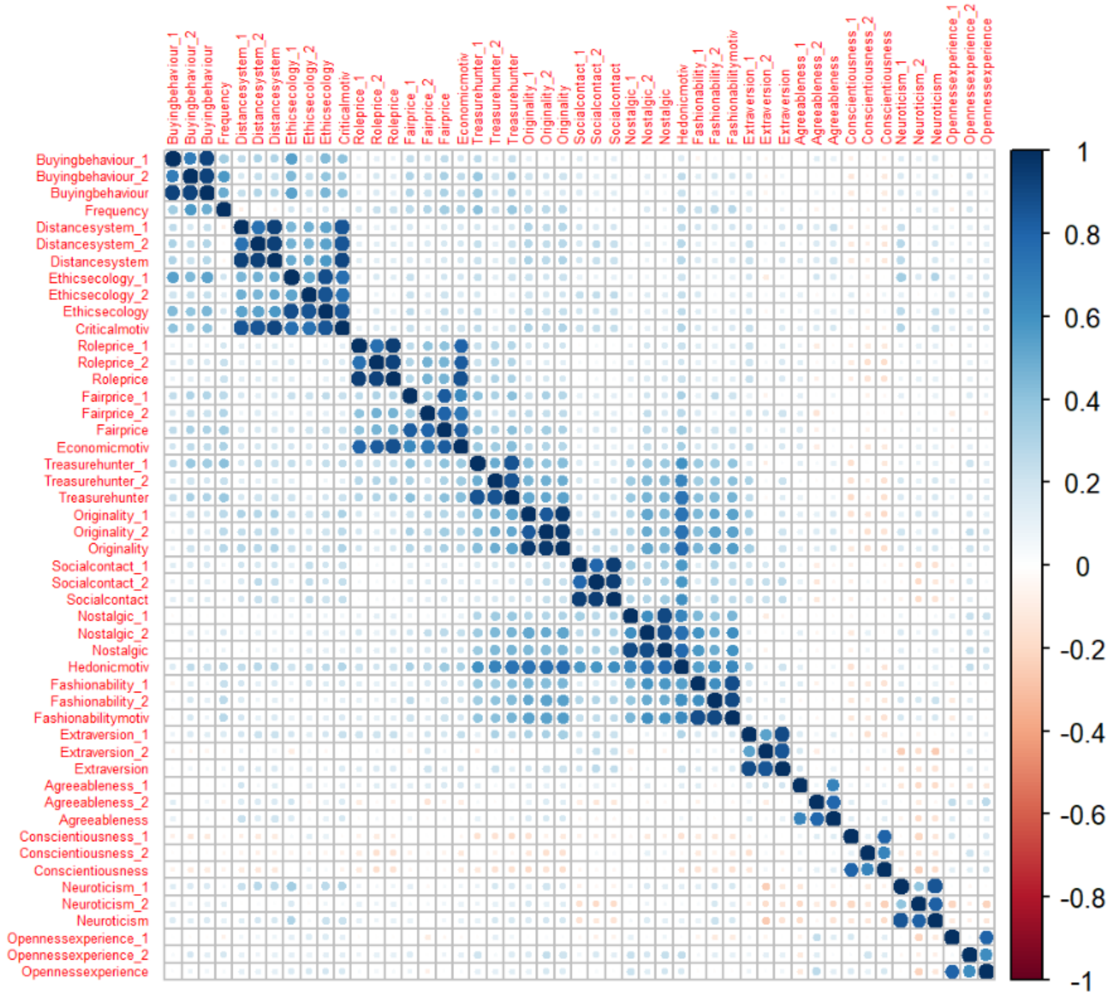

Figure 3.17 Table of cross-correlation between the variables

From Figure 3.17, it is possible to see that the cross-correlations between the variables are not too strong, except for the variables which measure the same aspect and for those created with the sum of other variables.

From the analysis of the correlations of all the numeric variables with Buyingbehaviour_1, it is possible to see the variables with higher correlation are Distancesystem_2, Ethicsecology_1, Roleprice_2, Fairprice_1, Treasurehunter_1, Originality_1, Socialcontact_1, Nostalgic_1, Fashionability_1, Extraversion_1, Agreeableness_2, Conscientiousness_1, Neuroticism_1, Opennessexperience_2. From the analysis of the correlations of the variables with

Buyingbehaviour_2, it is possible to see the variables with higher correlation are Distancesystem_2, Ethicsecology_1, Roleprice_2, Fairprice_1, Treasurehunter_1, Originality_1, Socialcontact_1, Nostalgic_2, Fashionability_2, Extraversion_1, Agreeableness_2, Conscientiousness_1, Neuroticism_1, Opennessexperience_2. The variables Nostalgic and Fashionability can be considered an exception in this analysis because they are the only ones that change. Nostalgic_1 and Fashionability_1 are the ones to have a higher correlation with Buyingbehaviour_1, instead Nostalgic_2 and Fashionability_2 are the ones to have a higher correlation with Buyingbehaviour_2.

From the analysis of the correlations, it is possible to see that the variables with higher correlation with the variable Frequency are Distancesystem_1, Ethicsecology_1, Roleprice_2, Fairprice_2, Treasurehunter_1, Originality_2, Socialcontact_1, Nostalgic_2, Fashionability_2, Extraversion_1, Agreeableness_2, Conscientiousness_2, Neuroticism_2, Opennessexperience_2.

Chapter 4: The influence of personality on motivations

In this chapter, we want to study, through the linear regression analysis, if the Big Five Personality Traits can explain the different motivations to purchase second-hand clothes by consumers, such as critical, economic, hedonic and fashionability motivations. In other words, we want to see if the personality of the individual is able to influence the reasons why individuals purchase in this new way.

4.1 Linear regression theory

Linear regression is a simple way to perform supervised learning and predicting a quantitative response Y on the basis of a single or multiple predictor variable X. In the first case, it is called simple linear regression, while if there are more independent variables it is called Multiple linear regression.

Multiple linear regression is based on some highly restrictive assumptions about the characteristics of the relationship between the predictors and the response. The additivity assumption means that the association between a predictor and the outcome is not influenced and does not depend on the values of the other predictors. The linearity assumption means that the change associated with a one-unit increase in the independent variable is constant for every value of the regressor.

The multiple linear regression model includes multiple predictors, giving each of them a separate slope coefficient.

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \cdots + \beta_p X_p + \epsilon,$$

Assuming to have p distinct predictors where X_j represents the jth predictor and β_j quantifies the association between that variable and the response, β_j can be interpreted as the average effect on Y of a one-unit increase in X_j, holding all other predictors fixed. While β_0 is an unknown constant which represents the intercept of the model. The error term ϵ is assumed to be independent of the regressors and consists of everything which is not included in the model. In fact, it is possible that the relationship between the variables is not linear or that there are other variables which influence the response variable.

In the multiple linear regression, it is also possible to have some independent variables which are qualitative predictors with two or more levels. If the qualitative predictor or factor has only two levels it is possible to create an indicator or dummy variable, which has two possible numerical values, most of the time the values of 0 and 1. The coefficient β_0 is the average value of the dependent variable for individuals in the class 0, while the sum of $\beta_0 + \beta_1$ is the average value for individuals in the other class, so β_1 can be interpreted as the average difference in the value of the dependent variable for the two groups of individuals.

When the qualitative predictors have more than two levels it is needed to create n-1 dummy variables, where n is the number of levels of the variable. The level which is not codified into a dummy variable is considered the baseline for comparison.

Estimating the Regression Coefficients

The regression coefficients β_0, β_1,..., β_p are unknown and must be estimated using the least squares approach, so choosing them in order to minimize the sum of squared residuals (RSS), where a residual is a difference between the observed value and the response value predicted by the multiple linear models.

$$\begin{aligned}\mathrm{RSS} &= \sum_{i=1}^{n}(y_i - \hat{y}_i)^2 \\ &= \sum_{i=1}^{n}(y_i - \hat{\beta}_0 - \hat{\beta}_1 x_{i1} - \hat{\beta}_2 x_{i2} - \cdots - \hat{\beta}_p x_{ip})^2.\end{aligned}$$

In other words, these estimated values are calculated in order to reduce the difference between the prediction and actual values of the data. The multiple least squares regression coefficient estimates $\hat{\beta}_0$, $\hat{\beta}_1$,..., $\hat{\beta}_p$, can be used to make predictions using the formula

$$\hat{y} = \hat{\beta}_0 + \hat{\beta}_1 x_1 + \hat{\beta}_2 x_2 + \cdots + \hat{\beta}_p x_p.$$

The first step in the multiple linear regression analysis is to check if there is a relationship between the response variable and the predictors. It is formulated a null hypothesis H_0: $\beta_1 = \beta_2 = \cdots = \beta_p = 0$ which means that there is no relationship between the regressors and the output and an alternative hypothesis H_a: at least one β_j is non-zero, which means that there is a relationship.

The hypothesis test is achieved through the F-statistic, which will have a value close to 1 where there is no relationship between the dependent variable and all the predictors and the null hypothesis is accepted. Instead when there is a relationship and the alternative hypothesis is accepted the value of the F-statistic will be bigger than 1 showing that at least one predictor must be related to the dependent variable.

$$F = \frac{(\mathrm{TSS} - \mathrm{RSS})/p}{\mathrm{RSS}/(n - p - 1)},$$

The value of the F-statistic that is necessary to refuse H_0 depends on the number of observations, in particular when n is large, an F-statistic little larger than 1 might still provide evidence against the null hypothesis, while when n is small it is needed a larger value.

The F-statistic follows an F-distribution when the null hypothesis is true and the errors ϵ_i have a normal distribution, so it is possible to calculate the p-value associated with it. It is important to first check this statistic because by chance about 5 % of the p-values associated with the independent variables will be below 0.05, even if these variables are not really significant in explaining the dependent variable. So it is not possible to just look at the p-values of the single variables to decide if there is a relationship, instead the F-statistic is not influenced by the number of predictors.

If looking at the p-value of this statistic, it is possible to conclude that at least one of the predictors is related to the response variable, the next step in the analysis is to look at individual p-values of the variables to check which ones of them are significant. This process called variable selection determines which predictors are significant and associated with Y. To test this, it is needed to determine if the estimate, ad example for β_1, is sufficiently far from zero and this depends on the accuracy of $\hat{\beta}_1$ or on its standard error. If the standard error $SE(\hat{\beta}_1)$ is small, it is possible that small values of $\hat{\beta}_1$ give evidence for this value to be different from 0. Instead, if $SE(\hat{\beta}_1)$ is large, $\hat{\beta}_1$ value must be large in absolute value too. If the p-value is smaller than 0,05 it is possible to say that there exists an association between this independent variable and the response variable.

Fit of the model

It is possible to check the fit of the model to the data through the residual standard error or RSE and the fraction of variance explained or R-square.

The residual standard error is an estimate of the standard deviation of ϵ, the error term associated with each observation. It is the average amount of the deviation of the dependent variable from the true regression line. If the predictions calculated through the model are close to the real observations, the model fits the data well and the residual standard error is small. Instead, it is large if the predictions are far from the real observations.

$$\mathrm{RSE} = \sqrt{\frac{1}{n-p-1}\mathrm{RSS}},$$

The R-square explains the proportion of variance of the dependent variable explained by the model and has a value between 0 and 1.

$$R^2 = \frac{\mathrm{TSS} - \mathrm{RSS}}{\mathrm{TSS}} = 1 - \frac{\mathrm{RSS}}{\mathrm{TSS}}$$

TSS measures the total variance in the response variable before the regression is performed, while RSS is the amount of variability which is not explained by the regression. The difference between these two terms is the amount of variability of the dependent variable explained by linear regression. This value is divided by TSS to create a measure of the proportion of this variability. If this statistic has a value close to 1, it means that this regression explains a big proportion of the variability of the dependent variable through the independent ones, while for values close to 0 it is true the opposite.

The value of R-square increases when more variables are added to the model, even if they are not significant because adding a variable decreases the residual sum of squares. For solving this problem in the interpretation, it is possible to look at the adjusted R-square.

The adjusted R-square value doesn’t increase with the number of regressors and if its value is close to 1, the model in consideration is able to explain a high percentage of the variance of the dependent variable. For a least squares model with d variables, the adjusted R2 statistic is calculated as Adjusted $R2 = 1 - RSS/(n - d - 1)\ TSS/(n - 1)$. This statistic is useful for selecting among various models which contain different numbers of variables.

After the fit of the model, it is possible to predict the response variable on the basis of a set of values for each of the predictors, even if this prediction is uncertain.

To analyze the data, I start conducting some multiple regression analysis to check my research hypothesis that the Big Five Personality Traits are able to influence the different motivations to purchase second-hand clothes of individuals.

4.2 Critical motivations analysis

I perform a multiple linear regression to check if the Big Five Personality Traits are able to explain the critical motivations of individuals. The socio-demographic characteristics are added as control variables.

Through the lm() function, I fit a multiple linear regression model. The basic syntax is lm(y ~ $x_1, x_2 \ldots x_p$, data), where y is the response, $x_1, x_2, \ldots x_p$ are the predictors, and data is the data set in which these variables are kept.

```
Call:
lm(formula = Criticalmotiv ~ Extraversion_1 + Agreeableness_2 +
    Conscientiousness_2 + Neuroticism_1 + Opennessexperience_2 +
    Gender + Generation + Educationlevel + Monthlyincome + Nationality,
    data = datasetb)

Residuals:
    Min      1Q  Median      3Q     Max
-11.138  -2.090   0.660   2.089   7.494

Coefficients:
                                           Estimate Std. Error t value Pr(>|t|)
(Intercept)                                 15.2044     2.7133   5.604 2.09e-07 ***
Extraversion_1                               0.6097     0.2834   2.151  0.03402 *
Agreeableness_2                              0.2134     0.3594   0.594  0.55406
Conscientiousness_2                         -0.2536     0.3742  -0.678  0.49961
Neuroticism_1                                0.7367     0.2953   2.495  0.01435 *
Opennessexperience_2                         0.5059     0.3696   1.369  0.17432
GenderMale                                  -0.5581     0.8738  -0.639  0.52461
GenderNon-binary / third gender              3.2208     2.0063   1.605  0.11178
Generation>=58                               0.8301     1.9725   0.421  0.67485
Generation29-42                              0.8188     1.5535   0.527  0.59937
Generation43-57                              5.7814     3.0008   1.927  0.05705 .
EducationlevelHigher than bachelor degree   -2.1508     1.0576  -2.034  0.04481 *
EducationlevelPrimary or secondary school   -2.6260     0.9426  -2.786  0.00646 **
MonthlyincomeLow income                      1.5914     1.0686   1.489  0.13976
NationalityItalian                          -0.2303     0.8518  -0.270  0.78752
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 3.752 on 94 degrees of freedom
Multiple R-squared:  0.2972,    Adjusted R-squared:  0.1925
F-statistic: 2.839 on 14 and 94 DF,  p-value: 0.001386
```

Figure 4.1 Output of the linear regression model n. 1, dependent variable: Criticalmotiv

The first step in the analysis is to look at the F-statistic to check if it is possible to refuse the null hypothesis. This value is bigger than 1 and the p-value is smaller than 0,05, so we state that the model does explain significantly the critical motivation.

Looking at the coefficients of the significant variables it is possible to state that (keeping constant all the other variables):

-For every increase of one point in the value of Extraversion_1, the Critical motivation increases by 0.609 points.
-For every increase of one point in the value of Neuroticism_1, the Critical motivation increases by 0.736 points.

Since the independent variable Education has three categories (“Primary or secondary school”, “Bachelor degree”, “Higher than bachelor degree”) it is automatically divided into 2 binary variables and one category is left as the reference to compare the results.

From Figure 4.2 the binary variables are

- Higher than bachelor degree = 1 (if the person has a level of education higher than a bachelor's degree) and 0 (if the person doesn’t).
- Primary or secondary school = 1 (if the person has primary or secondary school as the highest level of education) and 0 (if the person doesn’t). The variable Bachelor's Degree is coded implicitly: if the score of the two variables is 0, then the person has the Bachelor's Degree as its higher level of education.

```
contrasts(factor(datasetb$Education))

##                              Higher than bachelor degree
## Bachelor degree                                        0
## Higher than bachelor degree                            1
## Primary or secondary school                            0
##                              Primary or secondary school
## Bachelor degree                                        0
## Higher than bachelor degree                            0
## Primary or secondary school                            1
```

Figure 4.2 Categories of the variable Education

In this model, both Higher than bachelor's degree and Primary or secondary school have a negative and significant coefficient with a p-value under 0.05. This means that individuals in these categories have smaller Critical motivations compared to individuals in the reference group. In particular, individuals whose higher level of education is Primary or secondary school have a -2.626 lower value on Critical motivations compared to individuals with a Bachelor's Degree. A similar result was for individuals who have a higher level than a bachelor's degree of education and have a -2.151 lower value on Critical motivations compared to individuals with a Bachelor's Degree.

The model has an adjusted R-squared of 0.1925, this means that this model is able to explain the variability of the variable Critical motivations for around 20%. So results show that these variables predict Critical Motivations in 19,25% of its variance.

4.3 Economic motivations analysis

```
Call:
lm(formula = Economicmotiv ~ Extraversion_1 + Agreeableness_1 +
    Conscientiousness_2 + Neuroticism_2 + Opennessexperience_2 +
    Gender + Generation + Educationlevel + Monthlyincome + Nationality,
    data = datasetb)

Residuals:
    Min      1Q  Median      3Q     Max
-10.719  -3.112   0.327   2.819   8.120

Coefficients:
                                           Estimate Std. Error t value Pr(>|t|)
(Intercept)                               11.771823   3.592815   3.276  0.00147 **
Extraversion_1                             0.687960   0.328033   2.097  0.03866 *
Agreeableness_1                            0.704588   0.486440   1.448  0.15082
Conscientiousness_2                       -0.458884   0.427437  -1.074  0.28576
Neuroticism_2                              0.461264   0.355481   1.298  0.19761
Opennessexperience_2                       0.563337   0.423245   1.331  0.18641
GenderMale                                 1.472308   0.967489   1.522  0.13142
GenderNon-binary / third gender            4.404055   2.287261   1.925  0.05719 .
Generation>=58                             0.477354   2.242991   0.213  0.83193
Generation29-42                            1.528396   1.778798   0.859  0.39240
Generation43-57                            2.409537   3.383923   0.712  0.47819
EducationlevelHigher than bachelor degree -0.461625   1.172605  -0.394  0.69471
EducationlevelPrimary or secondary school  0.874875   1.072085   0.816  0.41653
MonthlyincomeLow income                    0.702934   1.184267   0.594  0.55423
NationalityItalian                        -0.006962   0.944460  -0.007  0.99413
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 4.255 on 94 degrees of freedom
Multiple R-squared:  0.1782,    Adjusted R-squared:  0.0558
F-statistic: 1.456 on 14 and 94 DF,  p-value: 0.1436
```

Figure 4.3 Output of the linear regression model n. 2, dependent variable: Economicmotiv

The first step in the analysis is to look at the F-statistic to check if it is possible to refuse the null hypothesis. This value is slightly bigger than 1, but due to the small number of observations, it is important to also check the p-value of the F-statistic. This p-value is bigger than 0,05, so we accept the null hypothesis and we state that the model does not explain significantly the Economic motivation.

4.4 Hedonic motivations analysis

```
Call:
lm(formula = Hedonicmotiv ~ Extraversion_1 + Agreeableness_1 +
    Conscientiousness_1 + Neuroticism_1 + Opennessexperience_2 +
    Gender + Generation + Educationlevel + Monthlyincome + Nationality,
    data = datasetb)

Residuals:
     Min       1Q   Median       3Q      Max
-16.2902  -5.0740  -0.0838   4.6787  20.3644

Coefficients:
                                     Estimate Std. Error t value Pr(>|t|)
(Intercept)                           20.5399     5.6207   3.654 0.000424 ***
Extraversion_1                         1.4190     0.5745   2.470 0.015307 *
Agreeableness_1                        1.3953     0.8497   1.642 0.103926
Conscientiousness_1                   -1.0426     0.5715  -1.824 0.071270 .
Neuroticism_1                          0.9377     0.5953   1.575 0.118537
Opennessexperience_2                   1.9278     0.7316   2.635 0.009838 **
GenderMale                            -1.4050     1.7338  -0.810 0.419789
GenderNon-binary / third gender       -1.8760     4.0716  -0.461 0.646037
Generation>=58                         1.6965     3.9022   0.435 0.664745
Generation29-42                        1.4450     3.1312   0.461 0.645520
Generation43-57                       -1.1508     6.0491  -0.190 0.849533
EducationlevelHigher than bachelor degree  1.9180  2.0966   0.915 0.362619
EducationlevelPrimary or secondary school -0.0786  1.9024  -0.041 0.967131
MonthlyincomeLow income                0.9998     2.1298   0.469 0.639849
NationalityItalian                    -2.2619     1.6867  -1.341 0.183144
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 7.548 on 94 degrees of freedom
Multiple R-squared:  0.2308,    Adjusted R-squared:  0.1162
F-statistic: 2.014 on 14 and 94 DF,  p-value: 0.02454
```

Figure 4.4 Output of the linear regression model n. 3, dependent variable: Hedonicmotiv

The first step in the analysis is to look at the F-statistic to check if it is possible to refuse the null hypothesis. This value is bigger than 1 and the p-value is smaller than 0,05, so we state that the model does explain significantly the hedonic motivation.

In this model, only the independent variables of Extraversion_1 and Opennessexperience_2 are significant in explaining the variability of the variable Hedonic motivations to purchase second-hand clothes by individuals. Looking at their coefficients it is possible to state that (keeping constant all the other variables), for every increase of one point in the value of Extraversion_1, the hedonic motivation increases by 1.419 points. For every increase of one point in the value of Opennessexperience_2, the Hedonic motivation increases by 1.927 points.

The model has an adjusted R-squared of 0.1162, which means that this model is able to explain the variability of the variable Hedonic motivations for a very low value, around 11,6%. This shows that the model is not really good at explaining this complex motivation.

4.5 Fashionability motivations analysis

```
Call:
lm(formula = Fashionabilitymotiv ~ Extraversion_1 + Agreeableness_1 +
    Conscientiousness_1 + Neuroticism_1 + Opennessexperience_2 +
    Gender + Generation + Educationlevel + Monthlyincome + Nationality,
    data = datasetb)

Residuals:
    Min      1Q  Median      3Q     Max
-5.9658 -1.2769  0.1297  1.3178  5.4715

Coefficients:
                                          Estimate Std. Error t value Pr(>|t|)
(Intercept)                                6.63578    1.77230   3.744 0.000312 ***
Extraversion_1                             0.26210    0.18114   1.447 0.151234
Agreeableness_1                            0.21123    0.26793   0.788 0.432451
Conscientiousness_1                       -0.22610    0.18019  -1.255 0.212672
Neuroticism_1                              0.32625    0.18770   1.738 0.085456 .
Opennessexperience_2                       0.23926    0.23069   1.037 0.302310
GenderMale                                -0.87940    0.54670  -1.609 0.111068
GenderNon-binary / third gender           -1.73950    1.28384  -1.355 0.178689
Generation>=58                             0.96964    1.23043   0.788 0.432649
Generation29-42                            0.29132    0.98731   0.295 0.768597
Generation43-57                            0.87952    1.90740   0.461 0.645784
EducationlevelHigher than bachelor degree  0.30007    0.66108   0.454 0.650944
EducationlevelPrimary or secondary school -0.08747    0.59987  -0.146 0.884385
MonthlyincomeLow income                    0.34076    0.67156   0.507 0.613058
NationalityItalian                        -0.34672    0.53185  -0.652 0.516050
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 2.38 on 94 degrees of freedom
Multiple R-squared:  0.1435,    Adjusted R-squared:  0.01596
F-statistic: 1.125 on 14 and 94 DF,  p-value: 0.3469
```

Figure 4.5 Output of the linear regression model n. 4, dependent variable: Fashionabilitymotiv

The first step in the analysis is to look at the F-statistic to check if it is possible to refuse the null hypothesis. This value is slightly bigger than 1, but due to the small number of observations, it is important to also check the p-value of the F-statistic. This p-value is bigger than 0,05, so we accept the null hypothesis and we state that the model does not explain significantly the fashionability motivation.

4.6 Interpretation of results

From the results of the different multiple regression models, it is possible to see that just some of the Big Five Personality Traits are able to explain the variability in the different consumer's motivations to purchase second-hand clothes.

In particular, more extroverted individuals tend to have higher values of Critical and Hedonic motivations to purchase second-hand.

Extroverted individuals enjoy socialization and being an active part of the community, so people with high extroversion want to support others and have a positive impact. (Sun et al., 2018) According to some previous research, the personality traits of extraversion, positively affect consumer attitudes toward green buying (Sun et al., 2018), a term to define the buying behaviours of the customer, directed to green and sustainable products, that can be considered positive for the natural environment. (Duong, 2022)

Milfont and Sibley (2012) reported that extraversion is significantly related to environmental engagement and according to the results of the research of Duong (2022), the effects of extraversion on the intention to engage in pro-environmental behaviour are only significant among women. In accordance with these previous researches, looking at the results, it is possible to state that, extroverted individuals tend to purchase second-hand clothes following critical motivations such as distancing themselves from the consumeristic society and fighting to avoid waste, for their wish to have a positive impact on a community that they feel part of.

Consumers who are more extroverted are also positively driven to purchase second-hand clothes for Hedonic motivations. In fact, according to the research of Matzler et al. (2006) extroverted consumers, who have a high degree of socialization and like their interaction with others, can be more inclined toward the hedonic motivations of consumption. This is because they like to share their experiences and to listen to the opinion of other people regarding their choices of purchases. These attitudes are essential in the second-hand market because the exchange is based on curiosity about the history of originals and unique objects. Most of the time, the buying interaction that leads to the purchase, is a dialogue between two individuals, where the seller tries to convince the buyer about the quality and the convenient price of the piece of clothing. Furthermore, according to the research of Gohary and Hanzaee (2014), individuals with high levels of extraversion have a significant positive relationship with hedonic values of shopping, because they tend to focus more on the positive aspects and emotions they perceive. It is possible to hypothesize that extroverted individuals may enjoy some specific aspects of second-hand shopping, such as the treasure-hunting research of a valuable piece in a context full of possibilities. Another characteristic of second-hand shopping is the unpredictability of the offer, which can push impulse buying for the fear of not find again a rare piece of clothing. Where Beatty and Ferrell (1998) defined impulsive buying as an unexpected and instantaneous purchase by a consumer, that had no pre-shopping intentions either to purchase that product category or to satisfy a buying task. According to the research of Turkyilmaz et al. (2015), the trait of extraversion is positively able to explain online impulse buying, and this can be extended to the impulsiveness of second-hand purchases, driven by hedonic motivations.

From the analysis, it emerges that consumers who are more neurotic are positively driven to purchase second-hand clothes following Critical motivations.
Neuroticism or negative affectivity can be defined as the tendency to experience psychological distress (Gohary & Hanzaee, 2014), so people with high values of this trait tend to be always or frequently worried. (Costa & McCrae, 1992) This can be the reason why individuals, characterized by high values of this trait, tend to value critical motivations as something which drives positively their purchase of second-hand clothes. It is possible that, compared to emotionally stable individuals, they worry more about the environment and they feel more pressure about the impact their purchases have on an already fragile ecosystem. So, it is possible that they purchase second-hand to feel less guilty and to reduce in any possible way their negative impact on the ecosystem, to feel less stressed about the future. This result is coherent with the study of Hirsh (2010), who unexpectedly highlighted that neurotic individuals have significantly higher levels of environmental concern.

According to the results of the analysis, individuals with higher values of the personality trait Openness to experience are positively driven to purchase second-hand clothes following their Hedonic motivations.
These results are coherent with the research of Matzler et al. (2006), according to which consumers with higher values for this trait can be more influenced by their hedonic motivations in the shopping environment. Hedonic motivations are related to the gain of gratification through shopping (Tsao & Chang, 2010) and through the research of something unique that can give individuals some strong emotional stimuli. In the context of second-hand shopping, these stimuli can derive from the possibility of treasure hunting, which drives the curiosity of

individuals. Individuals who are open to experiences are always searching for new experiences and so the second-hand sector can successfully attract these people thanks to the configuration of the channels, that always permit new discoveries. Also, Gohary and Hanzaee (2014), report that these consumers perceive the hedonic values of products more respect individuals that are defined as Closedness people.

Furthermore, people with high values of this trait tend to be interested in art. According to the results of the research of Fujiwara and Nagasawa (2015), individuals open to experience have higher purchase intentions for luxury goods, for their unique artistic value. It is possible that they can have positive evaluations for vintage pieces of clothing, considered unique and original pieces from another period, satisfying in this way the motivations of originality and nostalgic pleasure.

As for the trait of extroversion, according to the research of Turkyilmaz et al. (2015) the personality trait of openness to experience is positively able to explain online impulse buying, results that can be extended to the impulsiveness of second-hand purchases, driven by hedonic motivations.

Chapter 5: A comprehensive description of attitudes of individuals

In this chapter, we want to study, through the logistic analysis, the possibility to explain the attitude of consumers toward the second-hand purchase, seeing if the variability of the Buying Behaviour is explained by the other factors included in the research. Furthermore, we want to study what is able to explain the frequency of purchase of second-hand clothes of individuals.

5.1 Logistic Regression Theory

The linear regression model assumes that the response variable Y is quantitative, but in many situations, it can be qualitative or categorical. The approach for predicting qualitative responses is a process called classification. Predicting a categorical response can be referred to as "classifying" that observation, assigning the observation to a category or class. The problem with linear regression is that the predictions are not sensible, in fact, for small values it is possible to have a negative probability and for big values, it is possible to have a probability bigger than one, unless the range of X is limited. This is in contrast with the definition of probability which must be a value between 0 and 1. The methods used for classification, predict the probability that the observation belongs to each of the categories, as the basis for making the classification.

The classifier is built through just a group of observations and this classifier must perform not just on this training data but also on other observations not used to train it.

5.1.1 Multiple Logistic Regression

To avoid this, it is possible to use logistic regression, a classification method useful for qualitative response values. This regression models the probability that the outcome belongs to a category, so model p(Y) between the values 0 and 1 for every value of X probabilities. The logistic function produces an S-shaped curve that produces sensible prediction, close but never below 0 and close but never higher than 1 and better captures the range of probabilities.

Multiple logistic regression is used when a binary response is influenced by multiple p predictors $X = (X_1,...,X_p)$ and the probability of Y can be defined as $p(X)=Pr(Y=1|X)$.

$$p(X) = \frac{e^{\beta_0+\beta_1 X_1+\cdots+\beta_p X_p}}{1 + e^{\beta_0+\beta_1 X_1+\cdots+\beta_p X_p}}.$$

$$\log\left(\frac{p(X)}{1-p(X)}\right) = \beta_0 + \beta_1 X_1 + \cdots + \beta_p X_p,$$

The quantity $p(X)/[1-p(X)]$ is called the odds and can range between the values of 0 and ∞, where values of the odds close to 0 indicate very low probability and close to ∞ very high.

The logarithm of this value is called the log odds or logit and it is linear in X.

The coefficient β_1 has a different interpretation compared to the linear model, in fact in the logistic model an increase of X of 1 unit changes the log odds by β_1 or multiplies the odds by $e^{\wedge\beta 1}$. Due to the fact that the relationship between p(X) and X is not linear, the amount of change of p(X) for a one-unit increase of the dependent variable is influenced by the current value of X. Even if the magnitude of change is not always equal, for a positive β_1 there will always be an increase in p(X), while for a negative β_1, there will be a decrease of p(X).

After the estimation of the coefficients, it is possible to calculate the probability of Y, for every value of the regressors, which can be quantitative or qualitative.

<u>Estimating the Regression Coefficients</u>

The coefficients of the model are unknown and they are estimated using just a part of observations, called a train. In order to fit the model it is used the method of maximum likelihood, which searches for estimates of the coefficients such that the predicted probability is as close as possible to the observed results. So the estimates of β_0, β_1,..., β_p are chosen to maximize the mathematical equation called a likelihood function.

The accuracy of these coefficient estimates can be calculated through their standard errors with the z-statistic. The z-statistic associated with β_1 is equal to $\hat{\beta}_1/SE(\hat{\beta}_1)$. For a large absolute value of this statistic and a low value of p-value, it is possible to reject the null hypothesis and state that the probability of Y is not influenced by X. The estimated intercept β_0 is typically not of interest, in fact, it is used to adjust the average fitted probabilities.

5.2 Shrinkage Methods

In the regression models, the decision about which regressors include and which do not is one of high importance. In the linear models, the subset selection is performed through the least squares method. An alternative way to fit a model with all the p predictions can be to constrain or regularize their coefficient estimates or to shrink them toward zero. This can be a way to reduce the variance of the model, which can be high if all the p predictors are included in the regression.

The ridge regression and the lasso regression are the best-known techniques for shrinking the regression coefficients toward zero, where the first method includes all p predictors in the final model and the second one imposes some of them equal to 0. The lasso regression overcomes the disadvantage of ridge regression when the number of regressors is high. It does so by shrinking the coefficients estimated towards zero and forcing some of them to be equal to zero when the tuning parameter λ is sufficiently large. The lasso coefficient minimizes the quantity:

$$\sum_{i=1}^{n}\left(y_i\left(\beta_0+\sum_{j=1}^{p}x_ij\beta_j\right)-\log\left[1+\exp\left(\beta_0+\sum_{j=1}^{p}x_ij\beta_j\right)\right]\right)+\lambda\sum_{j=1}^{p}|\beta_j|$$

So, the purpose when performing lasso is to find the set of coefficient estimates with the smallest logistic log-likelihood function.

The lasso, implicitly assuming that some coefficients are equal to zero, performs a variable selection. As a result, models generated from the lasso are simpler and easier to interpret because they have fewer regressors. The lasso operates better where there is just a small number of predictors who have significant coefficients and the others are small or equal to zero.

The lasso results depend on the values of the tuning parameter λ, in fact when λ is equal to 0, the lasso gives the minimum log-likelihood. When λ is sufficiently large the lasso gives the null model with just the intercept, because all the coefficient estimates are shrunk to the 0 value. From this, it is derived that selecting the right value for λ is essential because the lasso can produce models with every number of variables, whereas increasing the value of lambda decreases the variance but increases the bias. The tuning parameter λ can be selected through cross-validation, selecting a grid of λ values and calculating the cross-validation error for each of them. Then the model is fitted again using all the available observations and the value of λ with the minimum cross-validation error. For the best value of lambda, there will be some predictors, called signal variables, which are related to the outcome and unrelated predictors, called noise variables, whose coefficients will be zero. The model is fitted again, using only the relevant regressors, which have coefficients different from zero, to have a better description and reduce the total variance of the model.

5.3 First model

In order to better analyze the variables which are able to describe and predict the variable Buying Behaviour of consumers who purchase second-hand clothes, I will perform some regressions with the binary outcome BB_bin. This variable has values equal to 1 when the buying behavior of the individuals is high and positive, while has values equal to 0 in the opposite case. I will conduct the logistic regression just on a part of the whole dataset. In particular, I will only consider the numeric variables of the different motivations and the personality traits, which are more correlated with the dependent variable Buyingbehaviour_1. This choice can be explained by the fact that each aspect, so each motivation and personality trait, was measured through two different and independent questions, which were codified into two different variables. I will include in the model also the variables Frequency and Channel and the categorical variables which are concerned with the socio-demographic characteristics of the sample.

I will repeat this analysis considering instead the variables which are more correlated with the dependent variable Buyingbehaviour_2 and I will confront the two models to choose the one that is more able to explain the binary variable BB_bin

Test and Train

To avoid incurring overfitting problems in the model, it is important to split the dataset into two different parts, train and test, where the first one will be used for the estimation and the second for the testing of the validity of predictions. In fact, it is best to test the model on a sample which is not included in the estimation process. In this particular case, I split the dataset into the train, ⅔ of the number of observations and the test, the remaining part.

Using the function glmnet() it is possible to fit the lasso regularization path, in particular in the formula it is set alpha=1 for performing the Lasso regression and family= “binomial” for indicating the logistic regression. The best model (out) is the one with type.measure="mae" due to the fact that the model with type.measure = "class" has a lower R2. In fact, the best model has an Adjusted R-squared=0.19, while using type.measure = "mae", the best model has an Adjusted R-squared= 0.524

```
## Call:  cv.glmnet(x = x[train, ], y = y[train], type.measure = "mae",
nfolds = 10, alpha = 1, family = "binomial", standardize = T)
##
## Measure: Mean Absolute Error
##
##       Lambda Index Measure      SE Nonzero
## min 0.01064    33  0.4958 0.07214      18
## 1se 0.09042    10  0.5663 0.05874       2
```

Figure 5.1 Results of cross-validation n.1

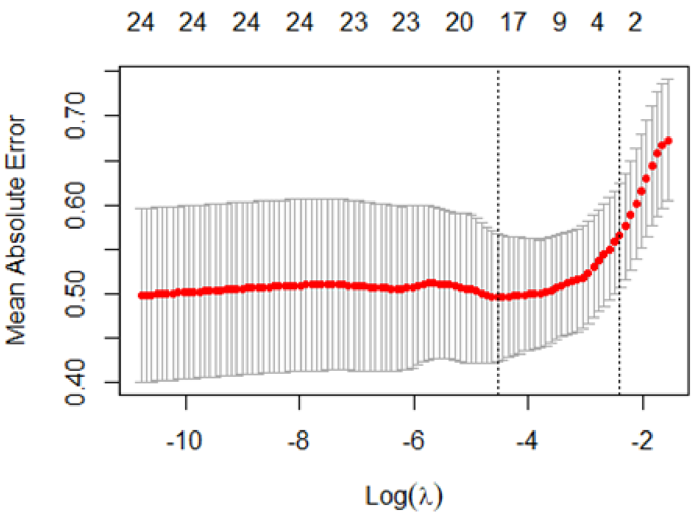

Figure 5.2 Plot of the cross-validation error according to the log of lambda n.1

The dashed vertical lines in Figure 5.2 indicate the log of the optimal value of lambda. The value of lambda 0.01064119 is the one where the cross-validation error is the smallest. The small value indicates that the optimal fit involves only a small amount of shrinkage.

<u>Assessing the classification performance of the model</u>

After having estimated the model on the train set, it is possible to predict and assess the classification performance of the model on the test. So I calculate the probabilities of the variables being predicted in the right category, transforming these predictions into a binary result and plotting them in a confusion matrix to check the results.

Through the confusion matrix, it is possible to check if the model correctly classifies or predicts if a consumer, for which are known just the independent variables, is an individual with low or high Buying Behaviour.

The confusion matrix contains

- the fraction of cases in which the model correctly predicts (or classifies) the respondents with high values of Buying Behaviour
- the fraction of cases in which the model wrongly predicts (or classifies) the respondents with high values of Buying Behaviour
- the fraction of cases in which the model correctly predicts (or classifies) the respondents with low values of Buying Behaviour
- the fraction of cases in which the model wrongly predicts (or classifies) the respondents with low values of Buying Behaviour

```
##         y.test
## bin.pred         0         1
##        0 0.5000000 0.1111111
##        1 0.5000000 0.8888889
```

Figure 5.3 Confusion matrix n. 1

The confusion matrix allows to measure several rates of correct classification and the value of the misclassification error.

- The sensitivity is the percentage of observations belonging to the class "1" correctly classified and in this model, it is equal to 0.8888889

- The specificity is the percentage of observations belonging to the class "0" correctly classified and in this model, it is equal to 0.5

- The accuracy resumes good classification cases from the confusion matrix, independently from the class. So it is a measure of the good of the general prediction of the model and in this model, it is equal to 0.7837838

- The misclassification error is equal to 1- accuracy, so it calculates the global percentage of misclassified observations independently from the class and in this model, it is equal to 0.2162162
  These rates of correct classification don't support the validity of the model for prediction purposes. In fact, specificity is equal to 50%, corresponding to casualty.

Post Lasso OLS regression

I create a data frame in which I include just the subsets of regressors chosen by the lasso method, with coefficients different from zero.

```
Call:
glm(formula = y ~ ., family = binomial(link = "logit"), data = x_GLM)

Deviance Residuals: 
     Min        1Q    Median        3Q       Max  
-2.35685   0.01347   0.13220   0.30301   2.00042  

Coefficients:
                                             Estimate Std. Error z value Pr(>|z|)   
(Intercept)                                 -10.44066    4.38802  -2.379  0.01734 * 
Frequency                                     1.58745    0.48251   3.290  0.00100 **
ChannelInternet.websites..internet.communities -3.45418  1.32723  -2.603  0.00925 **
Ethicsecology_1                               1.78907    0.54598   3.277  0.00105 **
Fairprice_1                                   0.28654    0.26295   1.090  0.27585   
Treasurehunter_1                              0.16940    0.30436   0.557  0.57782   
Nostalgic_1                                  -0.06642    0.30866  -0.215  0.82963   
Fashionability_1                              0.50275    0.42979   1.170  0.24210   
Extraversion_1                               -0.81541    0.39916  -2.043  0.04107 * 
Conscientiousness_1                          -0.15184    0.34847  -0.436  0.66304   
Opennessexperience_2                          0.05162    0.36477   0.142  0.88747   
GenderMale                                    2.16193    1.18018   1.832  0.06697 . 
GenderNon.binary...third.gender              -2.35349    1.67140  -1.408  0.15910   
Generation29.42                              -0.61965    1.48968  -0.416  0.67744   
Generation43.57                              -8.66718    3.02354  -2.867  0.00415 **
EducationlevelHigher.than.bachelor.degree    -1.64967    1.09416  -1.508  0.13163   
EducationlevelPrimary.or.secondary.school    -0.05918    1.11446  -0.053  0.95765   
MonthlyincomeLow.income                      -3.12082    1.30098  -2.399  0.01645 * 
NationalityItalian                            0.91831    0.83310   1.102  0.27034   
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for binomial family taken to be 1)

    Null deviance: 117.393  on 108  degrees of freedom
Residual deviance:  55.925  on  90  degrees of freedom
AIC: 93.925

Number of Fisher Scoring iterations: 7
```

Figure 5.4 Output of the lasso regression model n. 1, dependent variable: BB_bin

In this model not all the regressors even if their coefficients are different from zero, are significant in describing the buying behaviour. In order to have insights into how well this model fits the data, it is possible to look at the value of the Akaike information criterion (AIC). AIC is computed considering both the number of the independent variables considered in the model and the maximum likelihood estimate of the model, so how well this model is able to replicate the data. It can be utilized to evaluate different models and decide the one which is the best representation of the observations. In this case, the AIC of the model is equal to 93,925.

Assess the goodness of the model

In order to assess the goodness of fit of the model or the difference between observed values and predicted values, it is possible to calculate the likelihood ratio test.

This model can be retained because the likelihood ratio test has a p-value smaller than 5%, this very low value means that the model is accurate. In the case this value was bigger than 5%, the model would not pass the hypothesis test, meaning that the model is not able to explain the dependent variable in a better way that the null model, which contains just the intercept.

The value of the R-square statistic gives information about how well the model explains the dependent variable, but in logistic models, it is not possible to calculate this statistic. There are some alternative statistics called pseudo-R-squares which quantify the model fit. McFadden's pseudo R-squared quantifies the improvement over the null model and explains variance, but it gives no information about the correlation between the predicted and observed values. McFadden’s R squared measure is defined as

$$R^2_{\text{McFadden}} = 1 - \frac{log(L_c)}{log(L_{\text{null}})}$$

where Lc is the (maximized) likelihood value of the fitted model, and Lnull is the likelihood value for the null model, which includes only the intercept. The log-likelihood value is always negative for logistic regression. The values of this statistic can vary between 0 and 1, where if the value is close to zero the model has no predictive power and if it is close to 1 the model has a good predictive ability. High values of pseudo-R-square can be also influenced by other factors like overfitting, and so do not necessarily mean that the model is valid.

To calculate this statistic, it is needed to calculate the null model, which is composed just of the intercept.

The McFadden’s pseudo R-squared for this model is 0.524, it quantifies the explained variance and this value is quite high.

5.4 Second model

I repeat the same analysis with another model in which I will only consider the numeric variables of the different motivations and the personality traits, which are more correlated with the dependent variable Buyingbehaviour_2. I will include in the model also the variables

Frequency and Channel and the categorical variables which are concerned with the socio-demographic characteristics of the sample.

<u>Test and Train setting</u>

To avoid incurring overfitting problems in the model, also, in this case, it is important to split the dataset into two different parts, train and test, with the same procedure of the first model

Again the best model (out) is the one with type.measure="mae" because of the model with type.measure="class" has a lower pseudo-R-squared. In fact, the best model has pseudo-R-squared=0.19, while with type.measure="mae", the best model has pseudo-R-squared= 0.53

```
out

##
## Call:  cv.glmnet(x = x.[train, ], y = y[train], type.measure = "mae",
nfolds = 10, alpha = 1, family = "binomial", standardize = T)
##
## Measure: Mean Absolute Error
##
##       Lambda Index Measure      SE Nonzero
## min 0.01282    31  0.4856 0.06928      15
## 1se 0.07507    12  0.5506 0.05796       3
```

Figure 5.5 Results of cross-validation n.2

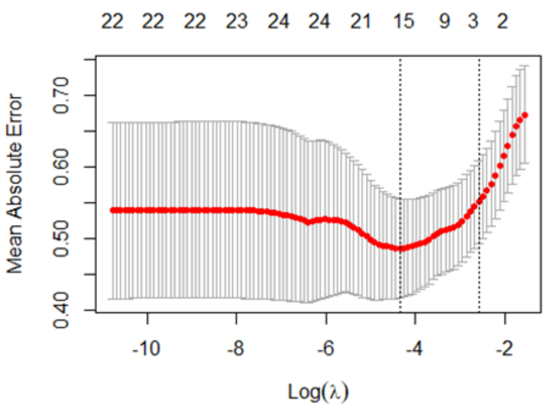

Figure 5.6 Plot of the cross-validation error according to the log of lambda n.2

The dashed lines from Figure 5.6 indicate the log of the optimal value of lambda, which minimizes the mean absolute error, is the one with a value of 0.01281736.

Assessing the classification performance of the model

After having estimated the model on the train set, it is possible to predict and assess the classification performance of the model on the test. So I calculate the probabilities of the variables being predicted in the right category, transforming these predictions into a binary result and plotting them in a confusion matrix to check at the results.

```
##         y.test
## bin.pred         0         1
##        0 0.5000000 0.1111111
##        1 0.5000000 0.8888889
```

Figure 5.7 Confusion matrix n.2

From the confusion matrix it is possible to measure the several rates of correct classification and the misclassification error.

- The sensitivity is the percentage of observations belonging to the class "1" correctly classified and in this model and it is equal to 0.889
- The specificity is the percentage of observations belonging to the class "0" correctly classified and in this model and it is equal to 0.5
- The accuracy resumes good classification cases from the confusion matrix, independently from the class. So it is a measure of the good of the general prediction of the model and in this model and it is equal to 0.784
- The misclassification error is equal to 1- accuracy, so it calculates the global percentage of misclassified observations independently from the class and in this model and it is equal to 0.216

These rates of correct classification don't support the validity of the model for prediction purposes. In fact, specificity is equal to 50%, corresponding to casualty in the model. These results are equivalent to the results of the first model.

Post Lasso OLS regression

I create a data frame in which I include just the subsets of regressors chosen by the lasso method, with coefficients different from zero.

```
Call:
glm(formula = y ~ ., family = binomial(link = "logit"), data = x_GLM)

Deviance Residuals:
     Min        1Q    Median        3Q       Max
-2.47253   0.00938   0.12430   0.30385   1.87509

Coefficients:
                                               Estimate Std. Error z value Pr(>|z|)
(Intercept)                                   -10.75465    3.26819  -3.291 0.000999 ***
Frequency                                       1.79095    0.54756   3.271 0.001073 **
ChannelInternet.websites..internet.communities -3.25285    1.19027  -2.733 0.006279 **
Ethicsecology_1                                 1.76307    0.51011   3.456 0.000548 ***
Fairprice_1                                     0.21330    0.24311   0.877 0.380283
Treasurehunter_1                                0.04535    0.29112   0.156 0.876195
Nostalgic_2                                     0.48822    0.30408   1.606 0.108367
Extraversion_1                                 -0.76135    0.34784  -2.189 0.028614 *
Opennessexperience_2                           -0.13268    0.36792  -0.361 0.718388
GenderMale                                      2.10512    1.12106   1.878 0.060409 .
GenderNon.binary...third.gender                -2.27891    1.59127  -1.432 0.152106
Generation29.42                                -0.49787    1.44128  -0.345 0.729769
Generation43.57                                -6.94794    2.38188  -2.917 0.003534 **
EducationlevelHigher.than.bachelor.degree      -1.48103    0.95640  -1.549 0.121490
MonthlyincomeLow.income                        -2.46663    1.14249  -2.159 0.030850 *
NationalityItalian                              1.48078    0.89562   1.653 0.098257 .
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for binomial family taken to be 1)

    Null deviance: 117.393  on 108  degrees of freedom
Residual deviance:  55.163  on  93  degrees of freedom
AIC: 87.163

Number of Fisher Scoring iterations: 7
```

Figure 5.8 *Output of the logistic regression model n. 2, dependent variable: BB_bin*

In this model not all the independent variables, even if their coefficients are different from zero, are significant in describing the dependent binary variable BB_bin.

In order to have insights into how well this model fits the data it is possible to look at the value of the Akaike information criterion (AIC). In this case, the AIC of the model is equal to 87.163. This value is lower compared to the one of the previous model with the dependent variables more correlated with Buyingbehaviour_1.

<u>Assessment of the goodness of fit</u>

In order to assess the goodness of fit of the model or the difference between observed values and predicted values, it is possible to calculate the likelihood ratio test.

This model can be retained because the likelihood ratio test has a p-value smaller than 5%, this very low value means that the model is accurate. In the case this value was bigger than 5%, the model would not pass the hypothesis test, meaning that the model is not able to explain the dependent variable in a better way that the null model, which contains just the intercept.

To estimate the pseudo R-square, it is needed to estimate the null model which is composed of just the intercept. The McFadden’s pseudo R-squared for this model is 0.53, it quantifies the explained variance and this value is quite high.

5.5 Comparison of the models and interpretation

According to the values of their likelihood ratio test p-value which are smaller than 5% both models can be retained. So in order to decide, which one is the best model to describe the binary dependent variable of Buying Behaviour, one must compare the value of the Akaike information criterion (AIC) calculated as AIC=2K-2ln(L) where k is the number of independent variables included in the model and L is the lo-likelihood estimate. In this case, the AIC of the first model is equal to 93,925 and the AIC of the second model is equal to 87.163. Both these values are quite low but the one of the model, built with the dependent variables more correlated with Buyingbehaviour_2, is lower. From this, it is possible to derive that the second model is better.

Taking a look at the model's pseudo-R-squares, it is possible to see that this statistic for the first model is equal to 0.524 while for the second model, it is equal to 0.53. This means that the model built with the motivations variables more correlated to BuyingBehaviour_2, has a higher explained variance and it is a better fit.

So from the summary of the second model in Figure 5.8, the relevant variables which have a significant effect on Buying Behaviour, are only Frequency, ChannelInternet.websites..internet.communities, Ethicsecology_1, Extraversion_1, Generation43.57 and MontlyincomeLow.income. In fact, their coefficients are statistically significant (with a p-value < 0.05), so our model suggests that these variables do in fact influence the probability of individuals having high values of Buying Behaviour.

In particular Frequency and Ethicsecology_1 have positive coefficients while Channel.Intenetwebsites.Internetcommunities , Extraversion_1, Monthlyincome.Lowincome, Generation.43-57 have negative coefficients. This means that an increase in these independent variables leads to a decrease in the probability for the consumers to have high values of the dependent variable Buying Behaviour.

```
##                                      (Intercept)
##                                            0.000
##                                        Frequency
##                                            5.995
## ChannelInternet.websites..internet.communities
##                                            0.039
##                                  Ethicsecology_1
##                                            5.830
##                                      Fairprice_1
##                                            1.238
##                                  Treasurehunter_1
##                                            1.046
##                                      Nostalgic_2
##                                            1.629
##                                   Extraversion_1
##                                            0.467
##                             Opennessexperience_2
##                                            0.876
##                                       GenderMale
##                                            8.208
##                   GenderNon.binary...third.gender
##                                            0.102
##                                  Generation29.42
##                                            0.608
##                                  Generation43.57
##                                            0.001
##       EducationlevelHigher.than.bachelor.degree
##                                            0.227
##                           MonthlyincomeLow.income
##                                            0.085
##                              NationalityItalian
##                                            4.396
```

Figure 5.9 Exponential transformation of the variables coefficients

Frequency is an ordinal variable; high levels correspond to a higher frequency of purchase of second-hand clothes. This categorization allows the Buying Behaviour to change from one category to the next and forces it to stay constant within it. In this case, the coefficient $\beta = 1.790$ is used to calculate $e^{\wedge\beta}$ (= $e^{\wedge 1.790}$ = 5.995) which can be interpreted as follows: Going up from 1 level of frequency to the next multiplies the odds of having high Buying Behaviour by 5.995. So increasing the level of frequency of the purchase is associated with an increase of 499,5% in the odds of having higher Buying Behaviour.

It is possible to state that, individuals who have a higher frequency in the purchase of second-hand clothes, tend to have higher values of appreciation of this way of purchasing. These values are logically correlated, the more an individual appreciates this form of shopping and prefers it to the fast fashion alternative, the more will buy his clothes in this way. It can be considered a virtuous circle, because the more someone buys second-hand, the more has the opportunity to have some positive reinforcement through experiences and pieces of clothes they might like.

The variable **Ethicsecology_1** has a coefficient of 1,763 which can be interpreted as $e^{\wedge\beta}$ ($=e^{\wedge 1,763}$=5,830) which indicated how odd the outcome (BB_bin) will change for 1 unit change in the predictor (Ethicsecology_1). An increase of 1 in the Ethicsecology_1 variable, multiplies the odds of having a high Buying Behaviour by 5,830.

So, for consumers of second-hand, the environmental aspect is very important in shaping their preference and attitude toward second-hand purchases. In particular, individuals prefer this way of shopping following their motivations connected with the values of ethics and ecology. So, the preference of someone for this way of purchasing partially derives from his agreement about the fact that, this way of purchasing help to fight against waste. In fact, this sector is growing in importance in recent times, as a response of consumers to the negative effects of fast fashion. This sector, as explained in the first chapter, is part of what is called a circular economy and consumers are becoming more sensitive to concerns about sustainability. Looking at the Figure 5.10 the leading reasons are to shop more sustainably, to give items a second life, and to perform a circular economy.

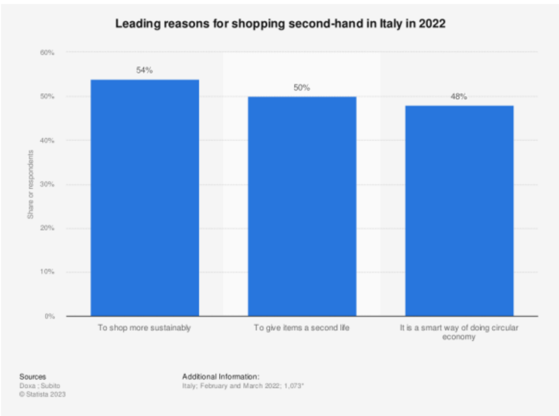

Figure 5.10 Leading reasons for shopping second-hand in Italy in 2022 (Doxa, 2022)

The variable **Extraversion_1** has a negative coefficient of -0.761 which can be interpreted as $e^{\wedge\beta}$ ($=e^{\wedge -0.761}$=0,467) that indicated how much the odds of the outcome (BB_bin) will change for each 1 unit change in the predictor (Extraversion_1). A decrease of 1 in the Extraversion_1 variable, so a decrease of 1 of the agreement on the scale, multiplies the odds of having a high Buying Behaviour by 0,467. So, a higher value of this personality trait in the consumer, has negative effects on the probability of the individual having positive perceptions of second-hand purchases. This result is in contraposition with the previous results of the regression analyses of personality traits and the different motivations. These previous analyses showed that higher values of this personality trait in individuals positively influenced their attitudes to the purchase of second-hand clothes, through their critical and hedonic motivations. To explain this different result, it is important to remember that this personality trait is a multifaceted construct, which can decline into different situational adaptations and behaviours of consumers, depending on the context in which the individual is.

The variable **Lowincome** has a negative coefficient of -2.466 which can be interpreted as $e^{\wedge\beta}$ ($=e^{\wedge-2.466}=0{,}085$) which indicates the odds ratio that associates Low income to the lower values of Buying Behaviour. Individuals who have lower monthly income have a (1 – 0.085=0,915) 91,5% reduction in the relative possibility of having higher Buying Behaviour compared to individuals with higher monthly income. So, a decrease in the probability to have a more positive consideration and preference of second-hand shopping, for individuals who have lower income availability. This result is in contraposition with the previous literature, which highlighted the importance of the economic aspect in the purchase of second-hand clothes. In fact, in the past, this channel was previously frequented by lower-income consumers and even if they are not the only type of consumers now, this aspect remains important in the purchase decision. Furthermore, according to the statistic in Figure 5.11 the share of people who are purchasing second-hand products to cut costs is around 27% globally, and it is higher for consumers in Generation Z, where this percentage is 35%, showing that consumers with lower income should have a higher preference toward this way of purchasing.

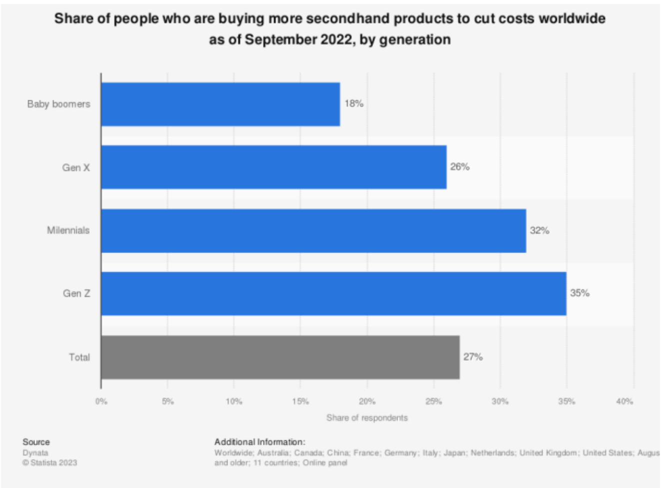

Figure 5.11 Share of people who are buying more secondhand products to cut costs worldwide as of September 2022, by generation (Dynata, 2022)

Since the independent variable, **Channel** has four categories (“Charity stores”, “Internet websites, internet communities”, “Second-hand markets, flea markets” and “Second-hand stores”) it is automatically divided into 3 binary variables and one category is left as the reference to compare as shown by Figure 5.12.

- **Internet websites, internet community** = 1 (if the consumer shops more frequently online) and 0 otherwise
- **Second-hand markets** = 1 (if the consumer shops more frequently in Second-hand markets) and 0 otherwise
- **Second-hand stores** = 1 (if the consumer shops more frequently in Second-hand stores) and 0 otherwise

The variable Charity stores is coded implicitly: if the score on all 3 variables is 0, then the person has Charity stores as their more frequent shop.

```
contrasts(factor(datasetb$Channel))

##                                         Internet websites, internet communities
## Charity stores                                                                0
## Internet websites, internet communities                                      1
## Second-hand markets, flea markets                                            0
## Second-hand stores                                                           0
##                                         Second-hand markets, flea markets
## Charity stores                                                          0
## Internet websites, internet communities                                0
## Second-hand markets, flea markets                                      1
## Second-hand stores                                                     0
##                                         Second-hand stores
## Charity stores                                           0
## Internet websites, internet communities                 0
## Second-hand markets, flea markets                       0
## Second-hand stores                                      1
```

Figure 5.12 Division of variable Channel into dummies

In this model the independent variable Channel.Internet website. Internet communities coefficient is significant (β =-3.253) with a p-value under 0.05. Individuals whose most frequented channels are Internet websites have 96,1% (1-0,039=0,961) less odds to have high Buying Behaviour compared to individuals who mostly purchase in charity stores (the reference class). It is possible to hypothesize that, this less probability to have higher values of the variable Buying Behaviour, borns from the other intrinsic reasons individuals have to purchase in charity stores. In fact, recently consumers are more concerned about the whole sustainability of their purchases, which can impact also on a social level through the exploitation of the workforce in developing countries. Choosing to purchase in charity stores, gives the consumer the possibility to contribute through their money to a benefic association.

Furthermore, purchasing second-hand clothes in a physical shop also reduces the perceived risk of consumers which negatively influences the purchase intention of products. (Lin & Chen,

2022) In the first chapter, perceived risk was defined as the consumer's beliefs about the uncertainty and the possible negative outcomes of purchasing. (Lin & Chen, 2022) In the context of second-hand clothes, consumers perceive the high risk of contamination, so that the products are damaged or ruined. The internet environment is the only channel, between those analyzed in this research, which does not provide the consumers the possibility to check the conditions of the piece of clothing before the purchase. This guarantee for the consumer can be very important, in a sector where the clothes one purchases are not new and can be ruined. This can be a possible explanation of why consumers, who purchase more online, have a lower probability to have a higher preference for second-hand shopping.

Furthermore, the internet environment is not characterized by all the hedonic opportunities of second-hand purchases. In particular, online is not possible to treasure hunting, to have the possibility to shop outside and to interact with others. While this online channel gives the possibility to compare more products and to shop in a faster and more efficient way, it can be also more stressful and impersonal for the consumers. In recent days, as explained in the first chapter, everyone has more clothes than they need and purchases most of the time are a way for individuals to spend their free time and have some hedonic benefits. Always more consumers are starting to purchase second-hand for its channel characteristics, but most of the time the internet environment doesn't provide the same satisfaction.

Since the independent variable **Generation** has four categories ("<=28", "29-42", "43-57", ">=58") it is automatically divided into 3 binary variables and one category is left as the reference group to compare them as shown by Figure 5.13.

The binary variables will be 29-42, 43-57, >=58 and their values will be 1 if the person is part of that generation and 0 if the individual is not that age. The reference group is composed of all the individuals under 28 years and it is coded implicitly if the score of all 3 variables is 0, then the person is part of Generation Z.

```
contrasts(factor(datasetb$Generation))

##       >=58 29-42 43-57
## <=28     0     0     0
## >=58     1     0     0
## 29-42    0     1     0
## 43-57    0     0     1
```

Figure 5.13 Division of the variable Generation into dummies

In this model, the independent variable Generationl.43.57 is significant with a p-value under 0.05 and a coefficient β =-6.9479. So individuals between 43 and 57 years have 99,9% (1-0.001) less odds to have high Buying Behaviour compared to individuals in Generation Z (the reference class). So, even if for a very low percentage, consumers who are part of Generation X, tend to have a low probability to have high values of the variable Buying Behaviour, which represent the preference for this way of purchasing.

The Generational Cohort Theory is a theory which divides the population into segments or generational cohorts through the years of birth, with a span between them of usually 20-25 years. (Lissitsa & Kol, 2021) This theory is extensively used because individuals tend to have similar ideas, attitudes, values and beliefs because they experience the same events at a similar age. Each generation so tends to have similar shopping behaviour and values. The generational cohorts can be divided like this: Baby Boomers, born between 1946 and 1965; Gen X, born between 1966 and 1980; Gen Y, born between 1981 and 1994; and Gen Z, born in 1995 and after. (Chaney et al., 2017 cited in Lissitsa & Kol, 2021)

Looking at Figure 5.14 the Z Generation is the one which is more interested in second-hand shopping, thanks to the benefits that this sector offers, like the possibility to have unique and original products at a lower price while respecting their sustainability beliefs and being able to satisfy some hedonic needs. (Kawulur et al., 2022) In fact, according to different research thrift shopping is more popular among young consumers, and this can explain why they are less diffident toward this way of purchasing. (Yan et al., 2015) The analysis of the model confirms this previous research, showing that consumers in the previous generations tend to be more diffident toward second-hand shopping and have a lower preference toward this shopping method.

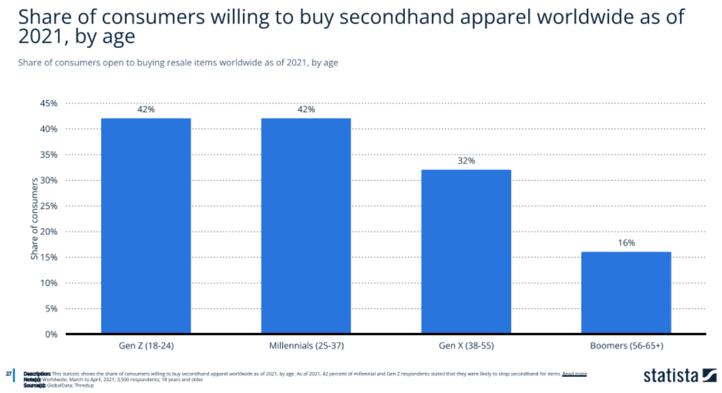

Figure 5.14 Share of consumers willing to buy secondhand apparel worldwide as of 2021, by age (Statista, 2022)

The results of the second model are in contraposition with the previous literature and with the results of the multiple regressions, so I will continue to analyse the data hoping to find better results.

5.6 First model frequency

Another variable that is interesting to study is Frequency, which indicates how frequently consumers of second-hand clothes purchase them over a year. This variable is categorical with more classes, but it is possible to transform it into a binary variable with values of 0 and 1 to study it more easily. The higher frequencies which correspond to the categories of “Quarterly”, “Monthly” and “Weekly” are codified with the value 1 while the categories “Every six months” and “Once in a year or less” are codified with the value 0. From Figure 5.15, it is possible to see that observations are almost equally split in the two groups.

```
## fre_bin
##  0  1
## 57 52
```

Figure 5.15 Distribution of frequency of the binary variable fre_bin

The first model is built including only the numeric variables of the different motivations and the personality traits, which are more correlated with the dependent variable Frequency. I will include in the model also the variable Channel and the categorical variables which are concerned with the socio-demographic characteristics of the sample.

<u>Test and Train setting</u>

To avoid incurring overfitting problems in the model, also in this case it is important to split the dataset into two different parts, train and test, with the same procedure of the models about Buying Behaviour, where the train is ⅔ of the number of observations.

The best model (out) is the one with type.measure="mae" because of the model with type.measure="class" has a lower pseudo-R-squared equal to pseudo-R-squared=0.273, while with type.measure="mae", the best model has pseudo-R-squared= 0.305

```
out

##
## Call:  cv.glmnet(x = x[train, ], y = y[train], type.measure = "mae",
nfolds = 10, alpha = 1, family = "binomial", standardize = T)
##
## Measure: Mean Absolute Error
##
##      Lambda Index Measure      SE Nonzero
## min 0.00016    73  0.9201 0.10936      25
## 1se 0.05999     9  1.0229 0.02049       9
```

Figure 5.16 Results of cross-validation n.3

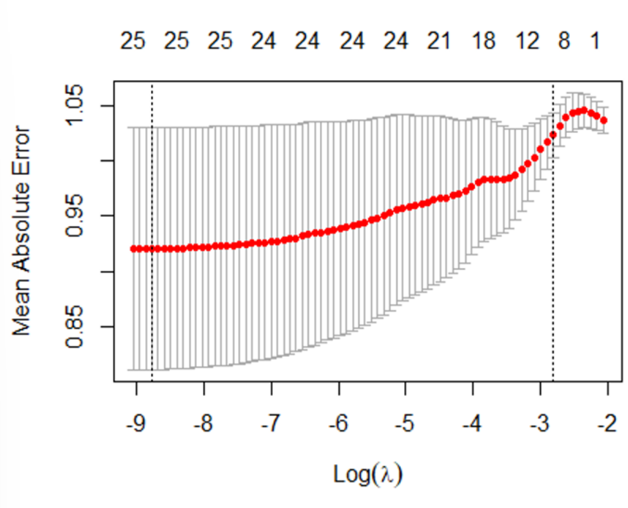

Figure 5.17 Plot of the cross-validation error according to the log of lambda n.3

The dashed lines in Figure 5.17 indicate the log of the optimal value of lambda, which minimizes the mean absolute error, is the one with a value of 0.0001556819.

Assessing the classification performance of the model

After having estimated the model on the train set, it is possible to predict and assess the classification performance of the model on the test. So I calculate the probabilities of the variables being predicted in the right category, transforming these predictions into a binary result and plotting them in a confusion matrix to check the results.

```
prop.table(table(bin.pred, y.test),2)

##          y.test
## bin.pred         0         1
##        0 0.7272727 0.4000000
##        1 0.2727273 0.6000000
```

Figure 5.18 The confusion matrix

From the confusion matrix it is possible to measure the several rates of correct classification and the misclassification error.

- The sensitivity is the percentage of observations belonging to the class “1” correctly classified and in this model, it is equal to 0.6
- The specificity is the percentage of observations belonging to the class “0” correctly classified and in this model, it is equal to 0.727
- The accuracy resumes good classification cases from the confusion matrix, independently from the class. So it is a measure of the good of the general prediction of the model and in this model it is equal to 0.676
- The misclassification error is equal to 1- accuracy, so it calculates the global percentage of misclassified observations independently from the class and in this model, it is equal to 0.324

These rates of correct classification support the validity of the model for prediction purposes. In fact, sensitivity and specificity are greater than 50%, which corresponds to casualty in the model. The value of misclassification error is not very low and can be interpreted as 32% of the observations are misclassified in the analysis.

Post Lasso OLS Regression

I create a data frame in which I include just the subsets of regressors, chosen by the lasso method, with coefficients different from zero.

```
Call:
glm(formula = y ~ ., family = binomial(link = "logit"), data = x_GLM)

Deviance Residuals:
    Min       1Q   Median       3Q      Max
-2.1062  -0.6975  -0.1814   0.8797   2.0504

Coefficients:
                                              Estimate Std. Error z value Pr(>|z|)
(Intercept)                                   -7.17386    2.93274  -2.446  0.01444 *
ChannelInternet.websites..internet.communities  0.97321    1.02980   0.945  0.34463
ChannelSecond.hand.markets..flea.markets        0.56180    0.94142   0.597  0.55067
ChannelSecond.hand.stores                       0.85864    0.97309   0.882  0.37757
Distancesystem_1                              -0.72872    0.26240  -2.777  0.00549 **
Ethicsecology_1                                0.50318    0.30629   1.643  0.10042
Roleprice_2                                    0.27449    0.23185   1.184  0.23644
Fairprice_2                                    0.26714    0.24042   1.111  0.26650
Treasurehunter_1                               0.72071    0.25876   2.785  0.00535 **
Originality_2                                  0.01782    0.23657   0.075  0.93994
Socialcontact_1                                0.07475    0.19241   0.389  0.69764
Nostalgic_2                                   -0.27321    0.24657  -1.108  0.26785
Fashionability_2                               0.22623    0.23983   0.943  0.34552
Extraversion_1                                 0.06646    0.23133   0.287  0.77388
Agreeableness_2                               -0.26197    0.29632  -0.884  0.37666
Conscientiousness_2                           -0.17439    0.27821  -0.627  0.53076
Neuroticism_2                                  0.10247    0.23778   0.431  0.66652
Opennessexperience_2                           0.36930    0.28277   1.306  0.19155
GenderMale                                    -1.10579    0.64087  -1.725  0.08445 .
GenderNon.binary...third.gender               -0.09212    1.25333  -0.074  0.94141
Generation29.42                                1.44527    1.05127   1.375  0.16920
Generation43.57                               -0.29325    1.78655  -0.164  0.86962
EducationlevelHigher.than.bachelor.degree     -0.91955    0.72744  -1.264  0.20620
EducationlevelPrimary.or.secondary.school     -0.47724    0.73369  -0.650  0.51539
MonthlyincomeLow.income                        1.09490    0.73224   1.495  0.13485
NationalityItalian                            -0.29320    0.59636  -0.492  0.62297
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

(Dispersion parameter for binomial family taken to be 1)

    Null deviance: 150.88  on 108  degrees of freedom
Residual deviance: 104.87  on  83  degrees of freedom
AIC: 156.87
```

Figure 5.19 Output of the logistic regression model n. 3, dependent variable: fre_bin

In this model not all the independent variables, even if their coefficients are different from zero, are significant in describing the dependent binary variable fre_bin.

In order to have insights into how well this model fits the data, it is possible to look at the value of the null deviance = 150.88 which is the deviance of the model built just with the null hypothesis and to compare it with the residual deviance of this model which is equal to 104.87. From this is derived that this model is able to explain part of the deviance.

Assessment of the goodness of fit

In order to assess the goodness of fit of the model or the difference between observed values and predicted values, it is possible to calculate the likelihood ratio test.

This model can be retained because the likelihood ratio test has a p-value smaller than 5%, this very low value means that the model is accurate. In the case this value was bigger than 5%, the model would not pass the hypothesis test, meaning that the model is not able to explain the dependent variable in a better way that the null model, which contains just the intercept.

To estimate the pseudo R-square, it is needed to estimate the null model which is composed of just the intercept. McFadden’s pseudo R-squared for this model is 0.305, it quantifies the explained variance and this value is quite high.

5.7 Second model frequency

I repeat the same analysis with another model in which I consider all the numeric variables of the different motivations and the personality traits, except the variables created as a sum. I will include in the model also the variable Channel and the categorical variables which are concerned with the socio-demographic characteristics of the sample.

Test and Train setting

To avoid incurring overfitting problems in the model, it is important to split the dataset into two different parts, train and test, with the same procedure of the first model

Again the best model (out) is the one with type.measure="mae" because of the model with type.measure="class" has a higher p-value and it is not interpretable.

```
out
##
## Call:  cv.glmnet(x = x[train, ], y = y[train], type.measure = "mae",
nfolds = 10, alpha = 1, family = "binomial", standardize = T)
##
## Measure: Mean Absolute Error
##
##       Lambda Index Measure       SE Nonzero
## min 0.00001   100    0.966 0.07461      36
## 1se 0.12628     1    1.038 0.01157       0
```

Figure 5.20 *Results of cross-validation n.4*

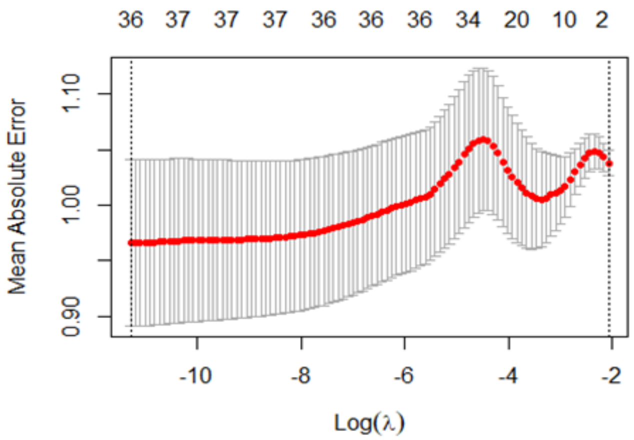

Figure 5.21 Plot of the cross-validation error according to the log of lambda n.4

The dashed lines in Figure 5.21 indicate the log of the optimal value of lambda, which minimizes the mean absolute error, is the one with a value of 1.262784e-05.

<u>Assessing the classification performance of the model</u>

After having estimated the model on the train set, it is possible to predict and assess the classification performance of the model on the test. So I calculate the probabilities of the variables being predicted in the right category, transforming these predictions into a binary result and plotting them in a confusion matrix to check at the results.

```
##          y.test
## bin.pred         0         1
##        0 0.5000000 0.4666667
##        1 0.5000000 0.5333333
```

Figure 5.22 Confusion matrix

From the confusion matrix it is possible to measure the several rates of correct classification and the misclassification error.

- The sensitivity is the percentage of observations belonging to the class “1” correctly classified and in this model, it is equal to 0.533
- The specificity is the percentage of observations belonging to the class “0” correctly classified and in this model, it is equal to 0.5
- The accuracy resumes good classification cases from the confusion matrix, independently from the class. So it is a measure of the good of the general prediction of the model and in this model, it is equal to 0.513
- The misclassification error is equal to 1- accuracy, so it calculates the global percentage of misclassified observations independently from the class and in this model, it is equal to 0.486

The rates not support the validity of the model for prediction purposes, in fact, both the sensitivity and the specificity values are really close or equal to 50%. This means that there are not significant advantages to using this model, with respect to random guessing.

<u>Post Lasso OLS Regression</u>

I create a data frame in which I include just the subsets of regressors chosen by the lasso method, with coefficients different from zero.

```
Call:
glm(formula = y ~ ., family = binomial(link = "logit"), data = x_GLM)

Deviance Residuals: 
     Min        1Q    Median        3Q       Max  
-2.21705  -0.64385  -0.06529   0.69888   1.93249  

Coefficients:
                                             Estimate Std. Error z value Pr(>|z|)   
(Intercept)                                  -6.33604    3.71687  -1.705  0.08826 . 
ChannelInternet.websites..internet.communities  1.82741    1.23280   1.482  0.13826   
ChannelSecond.hand.markets..flea.markets      0.69931    1.02517   0.682  0.49515   
ChannelSecond.hand.stores                     1.71059    1.13351   1.509  0.13127   
Distancesystem_1                             -0.60917    0.40602  -1.500  0.13353   
Distancesystem_2                             -0.29812    0.39695  -0.751  0.45264   
Ethicsecology_1                               0.61278    0.36952   1.658  0.09725 . 
Ethicsecology_2                               0.02512    0.37392   0.067  0.94645   
Roleprice_1                                   0.12488    0.32921   0.379  0.70443   
Roleprice_2                                   0.22716    0.40438   0.562  0.57428   
Fairprice_1                                  -0.29720    0.24799  -1.198  0.23076   
Fairprice_2                                   0.39309    0.30635   1.283  0.19945   
Treasurehunter_1                              0.98900    0.36169   2.734  0.00625 **
Treasurehunter_2                             -0.16055    0.29705  -0.540  0.58886   
Originality_1                                -0.26056    0.44277  -0.588  0.55621   
Originality_2                                 0.18932    0.40393   0.469  0.63928   
Socialcontact_1                               0.40091    0.38852   1.032  0.30212   
Socialcontact_2                              -0.18696    0.42796  -0.437  0.66221   
Nostalgic_1                                  -0.37739    0.30368  -1.243  0.21396   
Nostalgic_2                                  -0.01028    0.34473  -0.030  0.97621   
Fashionability_1                              0.10728    0.37402   0.287  0.77425   
Fashionability_2                              0.27743    0.32422   0.856  0.39218   
Extraversion_1                                0.46461    0.30575   1.520  0.12861   
Extraversion_2                               -0.75176    0.35970  -2.090  0.03662 * 
Agreeableness_1                               0.08137    0.36131   0.225  0.82183   
Agreeableness_2                              -0.23854    0.32970  -0.723  0.46938   
Conscientiousness_1                           0.16995    0.24300   0.699  0.48431   
Conscientiousness_2                          -0.25042    0.31216  -0.802  0.42244   
Neuroticism_1                                -0.19838    0.27500  -0.721  0.47067   
Opennessexperience_1                         -0.33704    0.26640  -1.265  0.20580   
Opennessexperience_2                          0.61150    0.37371   1.636  0.10178   
GenderMale                                   -1.78533    0.83782  -2.131  0.03309 * 
GenderNon.binary...third.gender              -0.17012    1.53233  -0.111  0.91160   
Generation29.42                               0.78397    1.25548   0.624  0.53234   
EducationlevelHigher.than.bachelor.degree    -1.23164    0.92773  -1.328  0.18431   
EducationlevelPrimary.or.secondary.school    -0.56193    0.81728  -0.688  0.49173   
MonthlyincomeLow.income                       1.40147    0.89653   1.563  0.11800   
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for binomial family taken to be 1)

    Null deviance: 150.877  on 108  degrees of freedom
Residual deviance:  95.285  on  72  degrees of freedom
AIC: 169.29

Number of Fisher Scoring iterations: 6
```

Figure 5.23 Output of the logistic regression model n. 4, dependent variable: BB_bin

In this model not all the independent variables, even if their coefficients are different from zero, are significant in describing the dependent binary variable fre_bin.

In order to have insights into how well this model fits the data, it is possible to look at the value of the null deviance = 150.88 which is the deviance of the model built just with the null hypothesis and to compare it with the residual deviance of this model which is equal to 95.285. From this is derived that this model is able to explain part of the deviance.

Assessment the goodness of fit

In order to assess the goodness of fit of the model or the difference between observed values and predicted values, it is possible to calculate the likelihood ratio test.

This model can be retained because the likelihood ratio test has a p-value smaller than 5%, equal to 0.01956. In the case this value was bigger than 5%, the model would not pass the hypothesis test, meaning that the model is not able to explain the dependent variable in a better way that the null model, which contains just the intercept. Even if this value is lower than 0,05 it is not very low, showing that the model is not very accurate.

To estimate the pseudo R-square, it is needed to estimate the null model which is composed of just the intercept. The McFadden's pseudo R-squared for this model is 0.368, it quantifies the explained variance and this value it is quite high. It is possible that this value it is influenced by other factors as overfitting problems.

5.8 Comparison of the models and interpretation

According to the values of their likelihood ratio test p-value which are smaller than 5% both models can be retained. So in order to decide, which one is the best model to describe the binary dependent variable of Frequency, one must compare the value of the Akaike information criterion (AIC) calculated as AIC=2K-2ln(L) where k is the number of independent variables included in the model and L is the log-likelihood estimate. In this case, the AIC of the first model is equal to 156.87 and the AIC of second the model is equal to 169.29. The value of this statistic is lower for the first model, build with only the variables which are more correlated with Frequency. From the value of this statistic and the results of the conclusion matrix, it is possible to conclude that the first one is the best model.

Taking a look at the model's pseudo-R-squares, it is possible to see that this statistic for the first model is equal to 0.305 while for the second model, it is equal to 0.368. These results are in contrast with the other measures, this can be explained by the fact that this statistic can be influenced by overfitting problems.

From the summary of the first model in Figure 5.19, the relevant variables which have a significant effect on Frequency are only Distancesystem_1 and Treasurehunter_1. Their coefficients are statistically significant (with a p-value < 0.05), so our model suggests that these variables do in fact influence the probability for individuals to have high values of Frequency, in a negative way for the first variable, in a positive way to the second.

```
##                                    (Intercept)
##                                          0.001
## ChannelInternet.websites..internet.communities
##                                          2.646
##       ChannelSecond.hand.markets..flea.markets
##                                          1.754
##                      ChannelSecond.hand.stores
##                                          2.360
##                               Distancesystem_1
##                                          0.483
##                                Ethicsecology_1
##                                          1.654
##                                    Roleprice_2
##                                          1.316
##                                    Fairprice_2
##                                          1.306
##                               Treasurehunter_1
##                                          2.056
##                                  Originality_2
##                                          1.018
```

Figure 5.24 Exponential transformation of the variables coefficients

The variable **Distancesystem_1** has a coefficient of -0.729 which can be interpreted as $e^{\wedge\beta}$ (=$e^{\wedge-0.729}$=0,483) that indicated how much the odds of the outcome (fre_bin) will change for each 1 unit change in the predictor (Distancesystem_1). A decrease of 1 in the Distancesystem_1 variable, so a decrease of 1 of the agreement on the scale, increases the odds of having a high value of Frequency of purchase by 0,483.

The motivation to purchase second-hand, driven by the desire of the consumers to distance themselves from the consumeristic system, reduce the probability to have higher values of Frequency of purchase for the consumers. This effect can be interpreted considering the consumers' attitudes toward sustainable purchases and the new kind of conscious consumer who wants to reduce their impact on the world. This goal is achieved through different behaviours, in fact, consumers are moving towards second-hand to reduce the negative impact of the item bought but, they are also reducing their frequency of purchases to avoid the overconsumption pushed by the fast fashion model. Individuals, with high values of the variable

Dystancesystem, don’t agree with the ideals of the fast fashion model which pushes to frequent and useless purchases. It is possible to hypothesize that individuals, with higher values of this variable, have a lower frequency of purchase toward second-hand because they buy only when they need the products and avoid worthless purchases, so they buy less in general.

The variable **Treasurehunter_1** has a coefficient of 0.721 which can be interpreted as $e^{\wedge\beta}$ ($=e^{\wedge 0.721}=2.056$) which indicated how odd the outcome (fre_bin) will change for 1 unit change in the predictor (Treasurehunter_1). An increase of 1 in the Treasurehunter_1 variable, multiplies the odds of having a high Frequency by 2.056.

The behaviour of searching around, for valuable and unique pieces of clothing, is part of the hedonic motivation to purchase something by the consumer. In fact, the research, between a multitude of objects, that leads to the finding of something unique, releases positive emotions and satisfaction in the consumer, that will crave again these emotions. Consumers, who want to satisfy their hedonic needs, will go to flea markets and other channels, where it is possible to wander around more frequently compared to others. According to Ferraro et al. (2016), there is a segment of consumers called Treasure Hunting Influencers, who are frequent shoppers with often a weekly frequency, characterized by high values of treasure hunting and the ‘thrill of a bargain’. These characteristics can lead them to be fashion influences or leaders in this second-hand context, reflecting the values of this new generation of shoppers who want to merge value seeking, fashion and distance and avoidance of the classic market systems. (Ferraro et al., 2016)

Chapter 6: Groups of consumers

In this chapter, we want to study, through the clustering methods, the possibility to divide the consumers included in the analysis into different independent groups with specific characteristics. We conducted this further analysis to search for similarities and differences of consumers' behavior toward this new sector.

6.1 Clustering methods

The term clustering refers to a variety of techniques for finding subgroups or clusters in a data set. It is a form of unsupervised learning which searches a structure within a dataset and draws inferences from unlabeled data, instead of predicting the values of a response variable as regression methods. Clustering algorithms work by grouping a set of observations into different subsets or clusters, which are coherent internally and different externally. In this way of partitioning data, it is possible to interpret the observations inside a cluster as similar between them and different from the others in separate clusters. (DataCamp Team, 2018) Clustering can be interpreted as a way to conduct market segmentation of the individuals analyzed into the dataset, and to categorize people who can be similar in their reaction to some marketing strategies.

There are many clustering methods and the more known approaches are K-means clustering and hierarchical clustering. In the first case, the observations are divided into a pre-specified number of clusters, while in the second case, it is possible to conduct the clustering analysis without knowing in advance the number of clusters desired. In fact, hierarchical clustering creates a dendrogram which shows all the cluster divisions for each possible number of clusters.

6.1.1 Hierarchical clustering

Hierarchical clustering can be of two main types, Agglomerative clustering also known as AGNES (Agglomerative Nesting) or Divisive hierarchical clustering, also known as DIANA (Divise Analysis). (*Hierarchical Cluster Analysis,* n.d.)

The first algorithm moves iteratively and follows a bottom-up approach, where each observation is considered as a single-element cluster or leaf. (*Hierarchical Cluster Analysis,* n.d.) Then the algorithm calculates the distance between every pair of observation points and stores these measures in a distance matrix. (DataCamp Team, 2018) At each step, following this distance matrix, the two more similar clusters are combined. The algorithm iterates this process, also creating a new distance matrix every time, and so decreases the number of clusters by one unit, forming a new bigger cluster or node. The algorithm repeats this procedure until all the observations are part of just one single big cluster called root. So the result of this process is a tree, plotted through a dendrogram. (*Hierarchical Cluster Analysis,* n.d.)

The Divisive Analysis follows a top-down approach, so the algorithm works in the inverse order of the one explained before. In fact, it starts by including all the objects into a single cluster and follows by splitting the most heterogeneous cluster into two. This step is repeated till all the observations are divided and belong to their own cluster. This method is better for identifying large clusters, while the first one is more indicated for small clusters. (*Hierarchical Cluster Analysis,* n.d.)

The first step in the analysis is to measure the (dis)similarity of observations using dissimilarity measures (*Hierarchical Cluster Analysis,* n.d.) and to calculate the similarity matrix. The choice of which dissimilarity measure to use depends on the type of data and it is very important because this choice has a strong effect on the dendrogram.

The so-called Linkage Methods are cluster agglomeration methods or different ways to measure the distance between groups of observations, used in order to decide the rules for clustering. The results of the analysis can change depending on which one is chosen. (DataCamp Team, 2018)

- The Minimum or single linkage clustering calculates all the pairwise dissimilarities between all the elements in two clusters and then considers the minimum of these as the linkage criterion. This method tends to produce long or “loose” clusters, where single observations are fused one at a time and so this method can be useful to detect some outliers in the dataset. (*Hierarchical Cluster Analysis,* n.d.)

- The Maximum or complete linkage clustering, after calculating all the pairwise dissimilarities between all the elements in two clusters, considers the maximum of these as the linkage criterion. It tends to produce more compact clusters.
- The Mean or average linkage clustering calculates all the pairwise dissimilarities between all the elements in two clusters and then considers the average of these as the linkage criterion.
- Ward's minimum variance method minimizes the total within-cluster variance, merging at each step the two clusters with minimum between-cluster distance.

In hierarchical clustering, it is possible to look at the differences between the Linkage Methods by looking at the dendrograms, which are similar to an upside-down tree in which each leaf corresponds to one observation and similar ones are combined into branches.

The dendrogram is built starting from the leaves and combining clusters up to the trunk. For this reason, the lower in the tree the groups are fused, the more similar they are, and so observations near the top of the tree are more different between themselves. There are $2n-1$ ways of reordering the dendrogram, where n is the number of leaves or observations. For this reason, to compare the similarity of two observations, one must look at where the branches, which contain them, combine. It is not their proximity along the horizontal axis, but the distance of split or merge, measured on the y-axis, which indicates the (dis)similarity between the two observations.

6.2 Cluster analysis

For this analysis, I will use the Agglomerative clustering method. In this case, due to the variables being continuous and numerical, I can use the Euclidean distance between two clusters. (DataCamp Team, 2018) I start the analysis by creating the distance matrix of a subset of data, containing only the variables of the different motivations (I consider the first measure), which drive the purchase of second-hand clothes according to the literature.

Comparing the different cluster dendrograms obtained through the different Linkage Methods, the best clustering solution is provided by the Ward method.

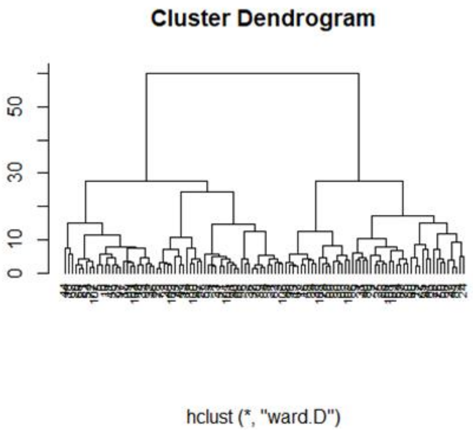

Figure 6.1 Cluster Dendrogram through Ward's minimum variance method

Looking at the Figure 6.1 it is clear that the observations are well separated in different clusters. Furthermore, the distances of agglomeration are very short at the beginning of the clustering process and then increase. This means that the observations in the clusters are very similar, while they are very different compared to observations in other clusters.

After plotting the dendrogram, it is possible to control the number of clusters in which it is possible to divide the dataset, cutting the dendrogram at a specific height. (*Hierarchical Cluster Analysis*, n.d.) This can be done through a horizontal line at a height, where this line can cross the maximum distance up and down, without intersecting the merging point. (DataCamp Team, 2018)

Looking at the distance of agglomeration and at the homogeneity of clusters, it seems that a two-cluster solution is more indicated to describe the data. To check the validity of this division, I used the NbClust package which, through varying all combinations of the number of clusters, distance measures, and clustering methods with 30 different indices, determines the best number of clusters and proposes the best clustering scheme. (DataCamp Team, 2018)

```
## *******************************************************************
## * Among all indices:
## * 8 proposed 2 as the best number of clusters
## * 3 proposed 3 as the best number of clusters
## * 1 proposed 4 as the best number of clusters
## * 3 proposed 5 as the best number of clusters
## * 2 proposed 8 as the best number of clusters
## * 1 proposed 11 as the best number of clusters
## * 1 proposed 12 as the best number of clusters
## * 3 proposed 14 as the best number of clusters
## * 1 proposed 15 as the best number of clusters
##
##                    ***** Conclusion *****
##
## * According to the majority rule, the best number of clusters is  2
##
##
## *******************************************************************
```

Figure 6.2 Best number of cluster calculated through the NbClust package

Following the majority rule of the NbClust package, the best number of clusters, in this analysis, in which it is possible to split the data is two subsets. To highlight the two clusters, we use the rect.hclust() function, which builds rectangular compartments for each cluster on the tree.

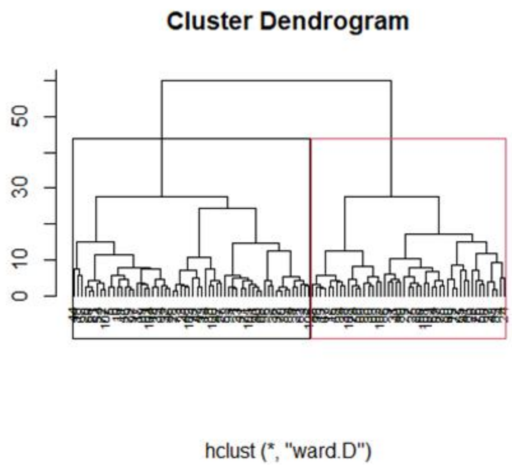

Figure 6.3 Cluster Dendrogram with the two clusters highlighted

In order to better analyze the characteristics and the differences of the observations included in each cluster, it is possible to provide some graphical representations of the two clusters. After plotting all the observations in multiple planes and in some 3D graphs, in which the axes are in turn the different clustering factors, the distribution of the observations around the different values is not clear. To avoid misinterpreting these values, it is possible to look at the Figure 6.4, in which are summed up the values of the arithmetic means of all the clustering factors, used to build the clusters.

	Distance sytem_1	Ethicsec ology_1	Roleprice_1	Fairprice _1	Treasure hunter_1	Originality_1	Socialcontact_1	Nostalgic_1	Fashionability_1	Size
clu1	5.367	6.017	5.217	4.817	6.017	5.900	3.783	5.083	5.483	60
clu2	4.408	5.327	5.082	4.122	4.327	4.122	2.673	3.612	4.245	49
Sample	4.936	5.706	5.155	4.504	5.257	5.101	3.284	4.422	4.927	109

Figure 6.4 Table of the arithmetic means of all the clustering factors

6.3 Cluster identification

The cluster identification is performed considering the different factors which are used in defining the clusters, so the various motivations.

The first difference between the two clusters is their dimension, where the first cluster is bigger with 60 individuals, while the second is composed of 49 individuals. Analyzing the values of the arithmetic means for the various motivations, displayed in Figure 6.4, it is possible to observe that individuals in the first cluster tend to be characterized by higher values for all the motivations, compared to individuals in the second cluster. It is possible to generally say that the first ones are more motivated regarding the motivations to purchase second-hand clothes defined by the previous literature.

It is possible to rename these two clusters in a way that expresses the characteristics of the individuals who are part of it.

1. It is possible to name the first cluster as the **highly motivated consumers** since it is composed of individuals who have high values for quite all the motivations to purchase second-hand clothes. These values are expressed on a 7-Likert scale, so it is possible to see that the motivation with lower values is Socialcontact_1 with medium values lower than 4, meaning that in general individuals don't consider the interaction with the seller as something which drives their purchase of second-hand clothes. Even if this value is lower than 4, it is still higher compared to the values of this motivation for individuals in the second cluster. All the other motivations have values bigger than 4, meaning that individuals consider them positively, as something that drives their behaviors. In particular, the higher motivation is Ethicsecology_1, followed by Treasurehunting_1. To conclude, it is possible to say that individuals in the first cluster are part of the interviewers who are more motivated toward second-hand clothes purchase.

2. It is possible to name the second cluster as **poorly motivated consumers**. In this cluster, individuals have lower values for all the motivations, compared to consumers in the first one. This means that they are in general less motivated to purchase second-hand clothes following these motivations. In particular, for the motivations of Socialcontact_1 and Nostalgic_1, their mean values are lower than 3, this means that they are not driven by these motivations and so they don't bring positive values toward second-hand. All the other motivations have values close to 4, which correspond to the response "Neither agree nor disagree", so these motivations

have not influenced the consumers in their purchase. The only exceptions are the variables of Ethicsecology_1 and Roleprice_1 which have a value bigger than 5. This means that also poorly motivated consumers are positively influenced by these two motivations which are connected to the critical and economic advantages of second-hand shopping.

6.4 Cluster characterization

To better define the consumers of these two different clusters, it is important to look at their similarities and differences around the other variables which were included in the analysis to describe their behavior.

First, it is possible to compare these two clusters by looking at the distribution of the binary variable Buying Behaviour, in which 0 represents the lower level of preference for the purchase of second-hand clothes and 1 is the higher level of preference. Looking at Figure 6.5 it is possible to compare the distribution of the frequency of this variable with the whole sample, where 77,06% of individuals have higher values and 22,94% have lower ones, it is possible to see that the first cluster has more individuals with positive perceptions.
In fact, for this variable, the individuals with higher values are 90% in the first cluster and 61.22% in the second, with complementary results for the frequency of individuals with lower perceptions.

```
lapply(bb_bin,round,2)

## $clusters
##
## memb     0     1
##    1 10.00 90.00
##    2 38.78 61.22
##
## $sample
##
##     0     1
## 22.94 77.06
```

Figure 6.5 Percentage distribution of the variable BB_bin in the clusters and the whole sample

To check if the difference in the proportion of this variable between the two samples is significantly different, it is possible to use the t-test for two independent samples.

This test compares the proportion of the variable in the two groups, to check if they are significantly different. In particular, in this case, I wanted to test if the proportion of the variable in the first cluster is significantly higher than the proportion in the second cluster.
Looking at the p-value, which is smaller than 0,05 it is possible to confirm that the first cluster is composed of the individuals who have higher Buying Behaviour values.

```
t.test(datasetb$BB_bin[memb==1],datasetb$BB_bin[memb==2], var.equal=TRUE,
paired=FALSE,alt="greater")

##
##  Two Sample t-test
##
## data:  datasetb$BB_bin[memb == 1] and datasetb$BB_bin[memb == 2]
## t = 3.7457, df = 107, p-value = 0.0001459
## alternative hypothesis: true difference in means is greater than 0
## 95 percent confidence interval:
##  0.1602896       Inf
## sample estimates:
## mean of x mean of y
## 0.9000000 0.6122449
```

Figure 6.6 T-test across the two independent clusters

Frequency

In Figure 6.7 are shown the values of the median of the variable Frequency, where high values of the variable correspond to higher frequencies. It is possible to see that for the first cluster, the frequency is higher compared to the whole sample and for the second it is equal. So individuals in the first sub-set tend to purchase second-hand clothes more often.

```
##                   clu1 clu2 sample
## median frequency    3    2      2
```

Figure 6.7 Median of the variable Frequency in the clusters and the whole sample

Channel

In Figure 6.8 are displayed the percentages of the type of channel, which individuals frequent more, in the different clusters. It is possible to observe that there is no difference for the variable charity stores, whose frequency is in both cases around 15%. Instead, consumers differ in their frequency of purchase in the other channels. 16,67% of individuals in the first cluster purchase more frequently second-hand clothes on internet websites and communities, 35% in second-hand markets and flea markets and 33,33% in second-hand stores. While 24,49% of individuals in the second cluster purchase more frequently in internet websites and communities, 22,45%

in second-hand markets and flea markets and 38,78% in second-hand stores. To summarize, consumers in the first cluster purchase more frequently in second-hand markets or flea markets while consumers in the second cluster, on internet channels and in second-hand stores.

```
## memb Charity stores Internet websites, internet communities
##    1         15.00                                  16.67
##    2         14.29                                  24.49
##
## memb Second-hand markets, flea markets Second-hand stores
##    1                             35.00              33.33
##    2                             22.45              38.78
##
## $sample
##
##                         Charity stores Internet websites, internet communities
##                                  14.68                                  20.18
##       Second-hand markets, flea markets                     Second-hand stores
##                                  29.36                                  35.78
```

Figure 6.8 Percentage distribution of the variable Channel in the clusters and the whole sample

6.4.1 Categorical variables

To describe the socio-demographic differences between the consumers in each cluster, it is possible to look at the percentage of respondents who are part of the different categories measured by the categorical variables Gender, Educational level, Generation, and Monthly income.

From Figure 6.9, it is possible to compare the variable Gender with the percentages for the whole sample. Individuals in the first clusters tend to be women, 75% compared to 69,72% in the whole sample, while the second cluster is composed of more males, 32,65% compared to 26,61% in the whole sample. There are no noticeable differences with the whole sample of the distribution in the clusters of the individuals who are part of a third gender or who are not binary.

```
## memb Female  Male Non-binary / third gender
##    1  75.00 21.67                      3.33
##    2  63.27 32.65                      4.08
##
## $sample
##
##                  Female                   Male Non-binary / third gender
##                   69.72                  26.61                      3.67
```

Figure 6.9 Percentage distribution of the variable Gender in the clusters and the whole sample

From Figure 6.10 the distribution of individuals across the different generations in the two samples is just slightly different compared to the whole sample. In particular, individuals in the first cluster tend to be younger than individuals in the second cluster, in fact, 86,67% of them are part of Generation Z, so are younger than 28 years. In the second cluster, the individuals part of Generation Z are 82,63% and there are higher percentages of individuals of the other generations.

```
## memb  <=28  >=58 29-42 43-57
##    1 86.67  5.00  6.67  1.67
##    2 81.63  6.12 10.20  2.04
##
## $sample
##
##  <=28  >=58 29-42 43-57
## 84.40  5.50  8.26  1.83
```

Figure 6.10 Percentage distribution of the variable Generation in the clusters and the whole sample

From Figure 6.11 the percentage of individuals for whom the higher level of education is the primary or secondary school, doesn’t differ between the two samples. In the first cluster, there are more consumers with a bachelor's degree and fewer with levels of instruction higher than a bachelor's degree. In the second cluster, there are more individuals with a level of instruction higher than the bachelor's degree, 24,49% compared to 21,10% of the whole sample and fewer individuals with a bachelor's degree as their higher level, 53,06% compared with 55,96% of the sample. Even if these differences are not deep, it is possible to say that individuals in the second cluster have reached a higher level of education, with respect to individuals in the first cluster.

```
## memb Bachelor degree Higher than bachelor degree Primary or secondary school
##    1           58.33                       18.33                       23.33
##    2           53.06                       24.49                       22.45
##
## $sample
##
##             Bachelor degree Higher than bachelor degree
##                       55.96                       21.10
## Primary or secondary school
##                       22.94
```

Figure 6.11 Percentage distribution of the variable Education in the clusters and the whole sample

From Figure 6.12 it is possible to observe that individuals in the first and second clusters highly differ between them according to their monthly income availability. In particular, individuals in the first group have a low-income level in 78,33% of the cases and a high-income level just in the 21,67%. Individuals in the second cluster, even if still have a low-income level in 59,18%

of cases, they have a high-income level in 40,82% of cases. These results are more interpretable by looking at the not equal distribution of income across the whole sample, where just 30,28% of individuals have a high-income level. So it is possible to say that individuals in the first cluster tend to be with less monetary availability compared to the whole sample and instead individuals in the second cluster are richer.

```
## memb High income Low income
##    1       21.67      78.33
##    2       40.82      59.18
##
## $sample
##
## High income  Low income
##       30.28       69.72
##
```

Figure 6.12 Percentage distribution of the variable Monthlyincome in the clusters and the whole sample

Looking at the distribution of the nationality of individuals in the two samples in Figure 6.13, there are no significant differences in the distribution of this characteristic across the whole sample. So in both samples, around 63% of the consumers are Italian.

```
## memb     0 Italian
##    1 36.67   63.33
##    2 36.73   63.27
##
## $sample
##
##       0 Italian
##    36.7    63.3
```

Figure 6.13 Percentage distribution of the variable Nationality in the clusters and the whole sample

6.4.2 The Big Five Personality Traits

Another way to characterize the individuals, who are part of the two different clusters, is to look at their values of the Big Five Personality Traits. In order to differentiate the individuals, it is possible to compare the average or the median of the quantitative variables in the two clusters with respect to the whole sample. I conducted my analysis using the arithmetic mean of the personality traits variables, which are more correlated with the variable BuyingBehaviour_2, and then I repeated the analysis using the median, which is a more robust method. Comparing the two types of analysis, I consider as more valid the comparison through the median, but I kept the mean analysis to have the possibility to compare the results.

From Figure 6.15 is possible to see that individuals in the first cluster tend to be more extroverted than individuals in the second cluster. To have values higher than 3,5 on a 5-Likert scale, corresponds to the agreement about their conformance with behaviors typical of individuals with high extraversion. For the individuals in the second cluster, it is true the opposite, they tend to be more introverted.

```
##                       clu1     clu2   sample
## mean extraversion 3.016667 2.734694 2.889908
```

Figure 6.14 Mean of the variable Extraversion_1 in the clusters and the whole sample

```
##                      clu1 clu2 sample
## median extraversion  3.5    2      3
```

Figure 6.15 Median of the variable Extraversion_1 in the clusters and the whole sample

From Figure 6.17 is possible to see that individuals in the first cluster tend to be more agreeable than individuals in the second cluster. To have values higher than 4 on a 5-Likert scale, correspond to the agreement about their conformance with behaviors typical of individuals with high agreeableness. For the individuals in the second cluster, their median is 3, meaning that they distribute between values which correspond to agreeableness and antagonism.

```
##                        clu1     clu2   sample
## mean agreeableness 3.483333 3.387755 3.440367
```

Figure 6.16 Mean of the variable Agreeableness_2 in the clusters and the whole sample

```
##                       clu1 clu2 sample
## median agreeableness    4    3      4
```

Figure 6.17 Median of the variable Agreeableness_2 in the clusters and the whole sample

From Figure 6.19 is possible to see that individuals in the first cluster tend to be less conscious than individuals in the second cluster. To have values lower than 2 on a 5-Likert scale means that they agree about their conformance with behaviors typical of individuals with low consciousness. For the individuals in the second cluster, their median is 3, meaning that they distribute between values which correspond to conscientiousness and unconscientiousness.

```
##                       clu1     clu2   sample
## mean conscientiousness  2.6 3.081633 2.816514
```

Figure 6.18 Mean of the variable Conscientiousness_1 in the clusters and the whole sample

```
##                          clu1 clu2 sample
## median conscientiousness    2    3      3
```

Figure 6.19 Median of the variable Conscientiousness_1 in the clusters and the whole sample

From Figure 6.21 is possible to see that, there is not a noticeable difference across the values of the personality trait of neuroticism between individuals in the first and second clusters. All the individuals tend to distribute around values around 3, a neutral value between neuroticism and emotional stability.

```
##                   clu1    clu2   sample
## mean neuroticism 3.05 2.77551 2.926606
```

Figure 6.20 Mean of the variable Neuroticism_1 in the clusters and the whole sample

```
##                     clu1 clu2 sample
## median neuroticism    3    3      3
```

Figure 6.21 Median of the variable Neuroticism_1 in the clusters and the whole sample

From Figure 6.23 is possible to see that there is not a noticeable difference in the personality trait of openness to experience between individuals in the first and second clusters. All the individuals across the two samples, tend to have values around 4, meaning that most of the individuals consider themselves as someone open to trying new experiences.

```
##                   clu1    clu2   sample
## mean openness 4.133333 3.77551 3.972477
```

Figure 6.22 Mean of the variable Opennessexperience_2 in the clusters and the whole sample

```
##                  clu1 clu2 sample
## median openness    4    4      4
```

Figure 6.23 Median of the variable Opennessexperience_2 in the clusters and the whole sample

To sum up, looking at the values of the personality Traits for the different clusters it is possible to observe that: the first cluster has higher values of Extraversion and lower values of Conscientiousness while the second cluster has lower values of Extraversion and lower values of Agreeableness.

6.5 Interpretation of results

The **highly motivated consumers** or individuals who have high values for quite all the motivations to purchase second-hand clothes, also differ from individuals in the second cluster, for other characteristics and behaviors. The individuals in the first cluster also have more positive perceptions towards second-hand Buying Behaviour. In fact, in the first cluster, 90% of individuals have higher values and just 10% of them have lower values for this variable. These positive inclinations toward second-hand purchases are reflected in the higher values of the frequency of purchase, compared to the whole sample and the second cluster. So individuals in the first sub-set tend to purchase second-hand clothes. Their shopping happens mainly in second-hand markets and flea markets (35% of individuals) and in second-hand stores (33,33% of individuals).

About their socio-demographic characteristics, consumers in the first cluster tend mostly to be women (75%) and young, in fact, 86,67% of them are part of Generation Z, so are younger than 28 years. Compared to the second cluster they tend to be less educated, with fewer individuals with levels of instruction higher than a bachelor's degree and also mostly with a low monthly income level, in 78,33% of the cases. So it is possible to say that individuals in the first cluster tend to be with less monetary availability compared to the whole sample. Their lower level of education achieved and lower income can derive from the fact that these consumers are mostly young.

These consumers in the two clusters also differ between themselves according to their values of the Big Five Personality Traits, where individuals in the first cluster tend to be more extroverted with values higher than 3,5 on a 5-Likert scale. These values represent their agreement about their conformance with behaviors typical of individuals with high extroversion. They also tend to be more agreeable with values higher than 4 on a 5-Likert scale.

On the opposite, they tend to be less conscious than individuals in the second cluster, as shown by their values lower than 2 on a 5-Likert scale. So they tend to be Undirectedness and not well organized.

The **poorly motivated consumers** or individuals who have lower values for all the motivations compared to the first cluster, also tend to have lower values of the variable Buying Behaviour, meaning that their perceptions about this way of shopping are less positive. These lower perceptions are reflected by their lower frequency of purchase compared to individuals in the first cluster. Their more frequented channels are second-hand stores (38,78% of individuals) and Internet websites and communities (24,49% of individuals).

Looking at the socio-demographics characteristic of this cluster, it is possible to see that despite being composed of more female consumers, the percentage of males is higher compared to the whole sample (32,65% compared to 26,61%). Individuals in this cluster tend to be slightly older, even if most of them are still part of Generation Z. There are more individuals with a level of instruction higher than a bachelor's degree and fewer individuals with a bachelor's degree as their higher level. Even if these differences are not deep, it is possible to say that individuals in the second cluster have reached a higher level of education and are more educated with respect to individuals in the first cluster. Consumers of the second cluster have more monetary availability, in fact, they have a high monthly income level in 40,82% of the cases, much higher compared to the first cluster.

These consumers of the second cluster differ according to their values of the Big Five Personality Traits, in fact, they tend to be more introverted compared to consumers of the first cluster. Their values of agreeableness and conscientiousness, are distributed around the value three, which is the median value of the 5-Likert scale. This means that they are not clearly polarized toward an aspect of this personality trait compared to individuals of the first cluster.

6.6 Comparison of results with the previous models

Comparing these clusters with the results of the multiple regression analysis and with the logistic regression analysis, it is possible to see that two groups naturally emerged from the dataset. In this way, we are able to deliver a more comprehensive view of highly and poorly motivated consumers, consistent with the previous literature and with the underlining characteristics of the sector.

Looking at the characteristics of the individuals of the first cluster, highly motivated consumers tend to have more positive perceptions towards second-hand purchases. This is coherent with the previous studies about motivations, which stated that the various motivations, such as critical, economical, hedonic and fashionability motivations, are able to explain the favorable attitude of consumers toward this sector. In the cluster analysis, this is demonstrated by the fact that in this cluster, most of the consumers, have high values for the variable Buying Behaviour. This corresponds to the consumer's agreement about preferring to buy second-hand clothing instead of comparable new ones when there is a choice and buying second-hand clothing whenever it is possible.

The only motivation, which can be considered an exception, is Social contact, this means that also consumers who consider all the other's motivations as significant, are not driven in the purchase by the interaction with the seller. A possible explanation for the low value of this motivation, in contraposition with the previous studies on the topic, can be that the medium age of the sample influences the analysis. In fact, the interviewees are mostly from Generation Z, who prefer to purchase products in a fast and efficient way, (Kawulur et al., 2022) so without losing time interacting with the seller to conclude the transaction.

This result is in contraposition with the logistic regression, which instead considers Ethics and Ecology, as the only relevant motivation to explain the purchase behaviour of individuals. This motivation is still the one with the highest value for all individuals, but it is not alone in explaining the differences between consumers.

Consumers characterized by higher values of the different motivations and positive inclinations toward second-hand purchases, also have higher values of the frequency of purchase. This group of consumers tend to purchase second-hand clothes more frequently, compared to individuals with less strong motivations. These values are logically correlated, the more an individual appreciates this form of shopping and prefers it to the fast fashion alternative, the more will buy his clothes in this way. It can be considered a virtuous circle, because the more someone buys second-hand, the more they have the opportunity to have some positive reinforcement, through experiences and pieces of clothes they might like. This result is coherent with the logistic analysis.

Analyzing the results of the cluster analysis, it is possible to see that highly motivated individuals purchase mostly in second-hand markets and flea markets (35% of individuals) and in second-hand stores (33,33% of individuals), while individuals in the second cluster prefer the internet environment. This result is coherent with the logistic analysis, in which emerged that individuals, who shop on Internet websites or communities, have less probability to have higher values of the variable Buying Behaviour.

While in the logistic analysis, the variable Gender was not significant, from this analysis it is possible to see that, there is a strong difference in the distribution of gender between the two clusters. The first cluster is mostly composed of women, who so have higher values of the various motivations and purchase more second-hand clothes, with a higher preference toward this sector. This result is coherent with some previous statistics; such as the study in Figure 6.24 which reported that in Germany women tend to purchase more second-hand clothes compared to men.

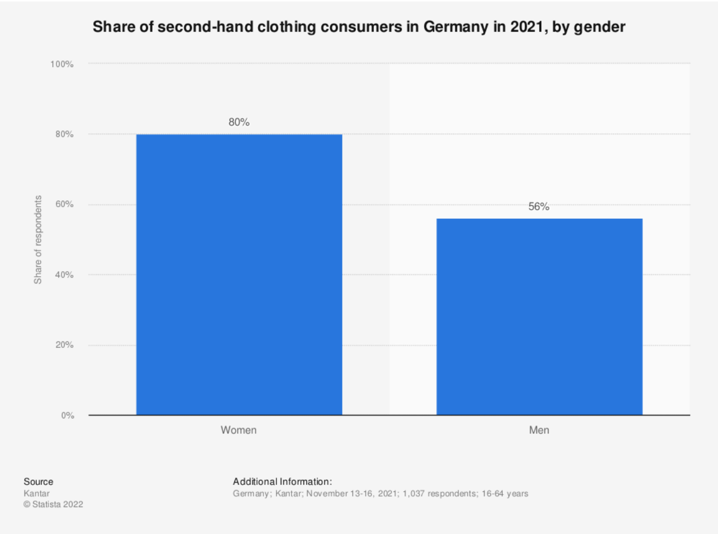

Figure 6.24 Share of second-hand clothing consumers in Germany in 2021, by gender (Momox, 2022)

The consumers in this first cluster tend to be younger and mostly of Generation Z, so it is possible to derive that consumers of younger generations tend to be more motivated to purchase second-hand clothes and this is reflected in their actual behaviors. This result is coherent with the previous literature, as shown in Figure 5.14, where the consumers of the Z Generation are the ones who are more interested in second-hand shopping. It is also coherent with the results of the logistic model, which showed that consumers of Generation X, tend to have a lower probability to have high values of the variable Buying Behaviour.

Consumers in the first cluster tend to be characterized by lower values of their monthly income availability and high motivations to purchase second-hand. This result is logical and coherent with the previous literature, which highlighted the importance of the economic aspect in the purchase of second-hand clothes. The results of the cluster analysis are coherent with Figure 5.11, which shows that the share of people who are purchasing second-hand products to cut costs is high. This percentage is even higher for consumers of Generation Z, which this cluster is mostly composed of. This shows that young individuals with low income, generally have higher values of the various motivations, higher preferences and higher frequency of purchase of second-hand clothes. This result is in contraposition with the results of the logistic analysis, which showed a decrease in the probability to have a more positive consideration and preference of second-hand shopping, for individuals who have lower income availability.

Consumers in this cluster tend to be slightly less educated, with individuals who achieved lower levels of instruction compared to the individuals in the second cluster. For the interpretation of this result, it is important to consider that the variable of Education is built considering the higher level of education, that is already achieved by the individuals. So it is possible, that most of the interviewees are still studying due to their low age, being mostly younger than 28 years old, with respect to the second cluster, where the level of instruction is higher but individuals are older.

Individuals in this first cluster tend to be more extroverted than individuals in the second cluster, so it is possible to state that more extroverted individuals have higher values of the various motivations to purchase second-hand, higher frequency and higher preference toward this way of purchasing. This is in contradiction with the results of the logistic regression, which showed that a higher value of this personality trait in the consumer, has a negative effect on the probability of the individual having positive perceptions of second-hand purchases. Instead, the cluster analysis result is coherent with the previous literature and the results of the multiple regression analysis in Chapter 4, which showed that more extroverted individuals tend to have higher values of Critical and Hedonic motivations to purchase second-hand.

<u>The personality trait of Agreeableness</u>

The cluster analysis highlights that consumers, with higher motivations to purchase second-hand clothes, also differ in having higher values of the personality trait of Agreeableness. This result is in contraposition with my previous analyses, which didn’t consider this trait as able to explain the variability in the various motivations or as able to influence the probability to have a high Buying Behaviour for the individuals. The personality trait of Agreeableness is concerned with interpersonal relationships and characterizes people who are kind, altruistic and empathetic toward others. According to the results of cluster analysis, individuals who are more agreeable have higher values of the various motivations to purchase second-hand, higher frequency and higher preference toward this way of purchasing.

As already analyzed in the second chapter, according to Wiggins (1991) individuals with high values of this trait are dominated by a sense of “communion” which can be defined as the membership of a larger community. (cited in De Raad & Perugini, 2002) If an individual is part of a community, he will search for a sense of union and solidarity with this group. This can lead these individuals to have higher values of the critical motivations to purchase second-hand clothes, because they want to improve the world not just for themselves. but also for others, which can represent an additional push toward sustainable and responsible behaviors. They also tend to participate in an activity which they consider beneficial or polite. (Tsao & Chang, 2010) This can lead agreeable individuals to become closer to second-hand purchases, because they may purchase from charity shops or people who may need that help, so this stronger empathy toward others can drive their critical motivations to purchase.

Agreeableness individuals can be considered as more soft-hearted and trusting (Costa & McCrae, 1992) and these qualities can strongly influence the purchase intention in the second-hand market, characterized by higher risks compared to the traditional sector. These consumers tend to trust more the seller of second-hand clothes and perceive less the possible risks, which can derive from this purchase. They also like the interaction, which can arise from the contact with the seller, both in the presence and on online platforms while doing online purchases. (Gohary & Hanzaee, 2014) From this it is possible to derive that, agreeableness consumers also have higher values of the hedonic motivations to purchase, such as social interaction with the seller, and purchase for their pleasure and fun.

Consumers with high values of the personality trait of agreeableness may want to conform to the social norms of the groups they are part of. This can lead them to have high fashionability motivations to purchase second-hand, because if they are surrounded by high fashion-driven individuals, they may start to purchase second-hand clothes for their unique style and the vintage image they deliver to others.

The personality trait of Consciousness

The cluster analysis highlights, that consumers with higher motivations to purchase second-hand clothes, also differ in having lower values of the personality trait of Consciousness. This result is in contraposition with my previous analysis, which didn’t consider this trait as able to explain the variability in the various motivations or as able to influence the probability to have a low Buying Behaviour for the individuals. According to the results of cluster analysis, individuals who are less conscious, have higher values of the various motivations to purchase second-hand, higher frequency and higher preference toward this way of purchasing.

This trait can be described as the aim to accomplish something by the individual. People with high values in this trait are often very systematic in their work, organized and efficient. (Goldberg, 1992) The personality traits of people are also reflected in their purchasing behavior, so it is possible to derive that, consumers characterized by high degrees of this trait, tend to be not a good fit for second-hand shopping. In fact, conscious individuals may develop methodologies also in the purchasing process, such as cognition, information processing, alternative evaluation, etc. (Karl et al., 2007, cited in Tsao & Chang, 2010) While shopping, they will use their previous experience, to make decisions about a piece of clothing, according to a set of characteristics in order to decide if buying it or if continue their research. (Tsao & Chang, 2010) From this, it is possible to derive that in their shopping experience, they are organized, cautious in their purchasing choices and determined in the achievement of their target. (Tsao & Chang, 2010) Their preference for organization and precision, can be incompatible with second-hand channels, which tend to be full of different objects and possibilities. Their layout is disorganized and for being able to purchase something, consumers need to spend time searching for valuable items, that most of the time it is not possible to find again or to compare.

Furthermore, conscientious people are able to control their impulse emotions and postpone gratification. (Joshanloo et al., 2012 cited by Gohary & Hanzaee, 2014) For this reason, they don’t tend to be impulsive shoppers, who make instant decisions, driven by their emotions and immediate thoughts. This negative relationship between impulsive behaviour and conscientiousness is highlighted by Gustavsson et al. (2003, cited by Gohary & Hanzaee, 2014) This is further evidence of the less compatibility of conscious consumers with the second-hand sector, characterized by purchases driven also by hedonic motivations.

The contradictions with the results of the previous models can derive from the relationship between the different variables in the dataset. In fact, logistic regression is useful to explain linear relationships between the variables, but it is possible that the relationship follows a different curve. The clustering method, being unsupervised, was instead able to overcome this limitation, naturally grouping individuals according to the strength of their motivation and providing a more complete analysis, which is also consistent with the previous studies.

Conclusions

This research confirmed the hypotheses of the model, demonstrating the influence of some of the Big Five personality traits on the different motivations to purchase second-hand clothes. Furthermore, it demonstrated how it is possible to categorize individuals with high values of the different motivations to purchase second-hand apparel, according to their values of the different Big Five Personality traits and some psychographic characteristics.

The results from the fourth chapter confirm the hypothesis that the Big Five Personality Traits are able to explain the variability in the different consumers' motivations to purchase second-hand clothes. In particular, more extroverted individuals tend to have higher values of Critical and Hedonic motivations to purchase second-hand, while consumers who are more neurotic are positively driven to purchase second-hand clothes following Critical motivations. Furthermore, individuals with higher values of the personality trait of Openness to experience are positively driven to purchase second-hand clothes following their Hedonic motivations.

The logistic model in Chapter 5 researched the hypothesis that it is possible to explain the variability in the attitudes toward second-hand purchase through the values of the different motivations and the values of the personality traits of the individuals.

From the results it is possible to state that individuals who have a higher frequency in the purchase of second-hand clothes tend to have higher values of appreciation of this way of purchasing. In particular, individuals prefer this way of shopping following their motivations connected with the values of ethics and ecology, so they agree that this way of purchasing help to fight against waste. The value of the personality trait of Extraversion of consumers has a negative influence on the probability of the individual having positive perceptions of second-hand purchases, a result in contraposition with the previous literature. Furthermore, individuals who have a lower monthly income have less probability of having higher Buying Behavior compared to individuals with a higher monthly income. Individuals whose most frequented channels are Internet websites have less probability to have a high attitude toward the purchase of these products compared to individuals who mostly purchase in charity stores (the reference class); the same relationship is valid for people part of Generation X compared to individuals in Generation Z.

These results are in contraposition with the previous literature and with the results of the multiple regressions, which analyzed the influence of the Big Five personality traits on the motivations. One possible reason behind this can be the fact that we only investigated the existence of linear correlations between the variables, whereas it is possible that they follow a different curve.
Moreover, the analysis in the same chapter suggests that the frequency of purchase of second-hand clothes is influenced only by the motivations of Distance from the system and Treasure hunting, where higher values of the former correspond to a lower frequency and higher values of the latter correspond to a higher frequency. A limitation of this model is that none of the Big Five Personality traits is able to explain the frequency of purchase. Also in this case we could infer the possible presence of nonlinear relationship between the studied variables.

In contrast, the unsupervised clustering method successfully overcame these limitations.
Two groups naturally emerged through the clustering analysis of the dataset, using the motivations of purchase as clustering factors. This model is able to deliver a more comprehensive view of highly and poorly motivated consumers, consistent with the previous literature and with the underlining characteristics of the sector.
From the results, highly motivated consumers tend to have more positive attitude towards second-hand purchases, with the only exception of the motivation of Social contact, which doesn’t drive positively the attitude toward the purchase. These higher values of the different motivations and the positive inclinations toward second-hand purchases are reflected in the higher values of the frequency of purchase, which happens mostly in second-hand markets and flea markets (35% of individuals) and in second-hand stores (33,33% of individuals).
These consumers tend to be young women, mostly from Generation Z, with low values of income and with a level of education corresponding to a bachelor’s degree. Regarding their personality, they tend to be more extroverted, more agreeable and less conscious.

These results might be useful for Marketing managers to define their communication strategies and customize them to target the different market segments with higher motivations and higher attitudes toward the purchase.

Limitations of the study and future research

Besides the contribution to the existing literature, this study presents some limitations which should be taken into account in future research.

The first limitation of this study regards the dimension of the sample, which is not sufficient to analyze the data in a more comprehensive way and the study can resent some biases, because responses were collected through convenience methods. This non-probabilistic sample was not representative of the real population, being composed mostly of Italian females from Generation Z. This depends partially from the fact that the categories with higher representation, are the ones which showed higher motivations toward the second-hand sector, higher attitudes and frequency to purchase. This bias in the distribution may partially derive by the fact that individuals who do not fall into these categories were reached by the questionnaire but were unable to respond because the survey required participants to have made at least one purchase of a second-hand clothing item in order to proceed. It will be interesting in future studies to repeat the research with a more balanced sample, to check if the results will be the equivalent or if new observations can be drawn from the research.

Secondly, most of the respondents are Italian and so the study cannot demonstrate universal conclusions. In future studies, it can be interesting to compare two samples of second-hand consumers across two countries, which highly differ in their purchasing culture, to assess the universality of the influence of the Big Five Personality traits on second-hand purchases.

Another limitation of this study is that it analyzes together all the categories of second-hand clothes, not distinguishing the vintage pieces from the rest. This can negatively influence the results, because it is possible to hypothesize that consumers who purchase just vintage pieces have higher values of fashionability and nostalgic motivations, compared to consumers who never had the opportunity or interest to purchase vintage. It is possible that these consumers are more relevant to other motivations and so the difference in the type of clothes purchased through second-hand can moderate some strong relations, which otherwise will be interesting to observe. Some future research can assess if significant differences exist across the motivations and the personality traits, for those consumers who are vintage enthusiast. It is possible to do so, by comparing the vintage enthusiast with second-hand consumers, who never purchased specific vintage clothes, and checking for differences across the two groups of consumers. It is also possible to repeat the same research comparing vintage enthusiast and individuals who never purchase second-hand clothes.

Some future research can analyze the same model through interviews conducted on-site, interviewing actual consumers of a variety of second-hand channels and consumers of fast fashion shops and checking for different values of the personality traits and the motivations to purchase in the two groups.

Some future studies can use the results of this research to suggest and create marketing campaigns, which target the more motivated consumers who have already higher attitudes toward this sector. Through specific campaigns, it is possible to increase their frequency and make them become affectionate clients for those brands and channels which search for new clients.

Appendix

Survey about second-hand purchase behavior

Start of Block: Introduction

Dear participant,
I'm a student from Ca' Foscari University of Venice and I'm conducting a research project for my Master’s degree.
The purpose of this study is to discover which motivations are able to drive the purchasing behaviour of the consumers of second-hand clothing and how these motivations of purchase are determined by the combination of the personality traits of the individuals.
I invite you to participate in this research project by completing a short online questionnaire, that will take about 6-8 minutes to be completed.
If you consent to participate, your responses will be kept confidential. The information provided will be used solely for the purpose of this research project and only aggregated results will be reported.
Your participation is voluntary, and you are free to withdraw consent at any time and to withdraw any unprocessed data you have previously supplied. Upon completion of the research, all questionnaires will be securely stored.
If you have any questions regarding this project, feel free to contact me:
ilacannazza@gmail.com
Thank you for your time and consideration.
Ilaria Cannazza

Do you want to give your consent to partecipate?

- ○ I consent to partecipate in this research survey (1)
- ○ I do not consent to partecipate in this research survey (2)

Skip To: End of Survey If Do you want to give your consent to partecipate? = I do not consent to partecipate in this research survey

---

Have you ever bought second-hand clothes?

- ◯ Yes (1)
- ◯ No (2)

Skip To: End of Survey If Have you ever bought second-hand clothes? = No

End of Block: Introduction

Start of Block: SHC buying behaviour

When there is a choice, I prefer to buy second-hand clothing instead of comparable new ones

- ◯ Strongly disagree (1)
- ◯ Somewhat disagree (2)
- ◯ Neither agree nor disagree (3)
- ◯ Somewhat agree (4)
- ◯ Strongly agree (5)

I buy second-hand clothing whenever it is possible

- ◯ Strongly disagree (1)
- ◯ Somewhat disagree (2)
- ◯ Neither agree nor disagree (3)
- ◯ Somewhat agree (4)
- ◯ Strongly agree (5)

End of Block: SHC buying behaviour

Start of Block: Second-hand purchase behaviour

I buy second-hand clothes

- Weekly (1)
- Monthly (2)
- Quaterly (3)
- Every six months (4)
- Once a year or less (5)

Which type of second-hand purchase channel you frequent most frequently?

- Second-hand markets, flea markets (1)
- Second-hand stores (2)
- Charity stores (3)
- Internet websites, internet communities (4)
- Other (5)

End of Block: Second-hand purchase behaviour

Start of Block: Critical motivations- Distance from the system

By buying second-hand clothes, I feel like I'm escaping the (consumption) system

- Strongly disagree (1)
- Disagree (2)
- Somewhat disagree (3)
- Neither agree nor disagree (4)
- Somewhat agree (5)
- Agree (6)
- Strongly agree (7)

Buying second-hand clothes enables me to distance myself from the consumerist society

- Strongly disagree (1)
- Disagree (2)
- Somewhat disagree (3)
- Neither agree nor disagree (4)
- Somewhat agree (5)
- Agree (6)
- Strongly agree (7)

End of Block: Critical motivations- Distance from the system

Start of Block: Critical motivations- Ethics and ecology

I enjoy buying second-hand clothes because I don’t like clothes being thrown away that can still be of use

- ◯ Strongly disagree (1)
- ◯ Disagree (2)
- ◯ Somewhat disagree (3)
- ◯ Neither agree nor disagree (4)
- ◯ Somewhat agree (5)
- ◯ Agree (6)
- ◯ Strongly agree (7)

By buying second-hand clothes, I feel I'm helping to fight against wastefulness

- ◯ Strongly disagree (1)
- ◯ Disagree (2)
- ◯ Somewhat disagree (3)
- ◯ Neither agree nor disagree (4)
- ◯ Somewhat agree (5)
- ◯ Agree (6)
- ◯ Strongly agree (7)

End of Block: Critical motivations- Ethics and ecology

Start of Block: Economic motivations- Gratificative role of price

I can afford more clothes because I pay less second-hand

- Strongly disagree (1)
- Disagree (2)
- Somewhat disagree (3)
- Neither agree nor disagree (4)
- Somewhat agree (5)
- Agree (6)
- Strongly agree (7)

I feel like I can have more clothes for less money by buying them second hand

- Strongly disagree (1)
- Disagree (2)
- Somewhat disagree (3)
- Neither agree nor disagree (4)
- Somewhat agree (5)
- Agree (6)
- Strongly agree (7)

End of Block: Economic motivations- Gratificative role of price

---

Start of Block: Economic motivations- Search for fair price

I don't want to pay more for a piece of clothing just because it's new

- Strongly disagree (1)
- Disagree (2)
- Somewhat disagree (3)
- Neither agree nor disagree (4)
- Somewhat agree (5)
- Agree (6)
- Strongly agree (7)

By buying second-hand, I feel I'm paying a fair price for clothes

- Strongly disagree (1)
- Disagree (2)
- Somewhat disagree (3)
- Neither agree nor disagree (4)
- Somewhat agree (5)
- Agree (6)
- Strongly agree (7)

End of Block: Economic motivations- Search for fair price

Start of Block: Hedonic/recreational motivations- Treasure hunting

I like wandering around second-hand outlets or roam second-hand websites because I always hope I'll come across a real find

- Strongly disagree (1)
- Disagree (2)
- Somewhat disagree (3)
- Neither agree nor disagree (4)
- Somewhat agree (5)
- Agree (6)
- Strongly agree (7)

In certain second-hand outlets or websites, I feel rather like a treasure hunter

- Strongly disagree (1)
- Disagree (2)
- Somewhat disagree (3)
- Neither agree nor disagree (4)
- Somewhat agree (5)
- Agree (6)
- Strongly agree (7)

End of Block: Hedonic/recreational motivations- Treasure hunting

---

Start of Block: Hedonic/recreational motivations- Originality

I hope to come across clothes that nobody else has

- Strongly disagree (1)
- Disagree (2)
- Somewhat disagree (3)
- Neither agree nor disagree (4)
- Somewhat agree (5)
- Agree (6)
- Strongly agree (7)

I hope to come across original clothes that are not found in mainstream stores or websites

- Strongly disagree (1)
- Disagree (2)
- Somewhat disagree (3)
- Neither agree nor disagree (4)
- Somewhat agree (5)
- Agree (6)
- Strongly agree (7)

End of Block: Hedonic/recreational motivations- Originality

Start of Block: Hedonic/recreational motivations- Social contact

What I like about certain second-hand outlets or websites is the pleasure of meeting and talking/chatting to people

- Strongly disagree (1)
- Disagree (2)
- Somewhat disagree (3)
- Neither agree nor disagree (4)
- Somewhat agree (5)
- Agree (6)
- Strongly agree (7)

I enjoy the social interaction you find in certain second-hand outlets or websites

- Strongly disagree (1)
- Disagree (2)
- Somewhat disagree (3)
- Neither agree nor disagree (4)
- Somewhat agree (5)
- Agree (6)
- Strongly agree (7)

End of Block: Hedonic/recreational motivations- Social contact

---

Start of Block: Hedonic/recreational motivations- Nostalgic pleasure

I like buying second-hand clothes because they evoke the past

- ○ Strongly disagree (1)
- ○ Disagree (2)
- ○ Somewhat disagree (3)
- ○ Neither agree nor disagree (4)
- ○ Somewhat agree (5)
- ○ Agree (6)
- ○ Strongly agree (7)

I like buying second-hand clothes because I find them authentic

- ○ Strongly disagree (1)
- ○ Disagree (2)
- ○ Somewhat disagree (3)
- ○ Neither agree nor disagree (4)
- ○ Somewhat agree (5)
- ○ Agree (6)
- ○ Strongly agree (7)

End of Block: Hedonic/recreational motivations- Nostalgic pleasure

---

Start of Block: Fashionability motivations

I enjoy buying second-hand clothes because they are fashionable

- Strongly disagree (1)
- Disagree (2)
- Somewhat disagree (3)
- Neither agree nor disagree (4)
- Somewhat agree (5)
- Agree (6)
- Strongly agree (7)

I enjoy buying second-hand clothes because their attractive styling is very important to me

- Strongly disagree (1)
- Disagree (2)
- Somewhat disagree (3)
- Neither agree nor disagree (4)
- Somewhat agree (5)
- Agree (6)
- Strongly agree (7)

End of Block: Fashionability motivations

Start of Block: Big Five Inventory

How well do the following statements describe your personality?
I see myself as someone who...

	Disagree strongly (1)	Disagree a little (2)	Neither agree nor disagree (3)	Agree a little (4)	Agree strongly (5)
...is reserved (1)	○	○	○	○	○
...is generally trusting (2)	○	○	○	○	○
...tends to be lazy (3)	○	○	○	○	○
...is relaxed, handles stress well (4)	○	○	○	○	○
...has few artistic interests (5)	○	○	○	○	○
...is outgoing, sociable (6)	○	○	○	○	○
...tends to find fault with others (7)	○	○	○	○	○
...does a thorough job (8)	○	○	○	○	○
...gets nervous easily (9)	○	○	○	○	○
...has an active imagination (10)	○	○	○	○	○

End of Block: Big Five Inventory

Start of Block: Demographics

What gender do you identify as?

- ○ Male (1)
- ○ Female (2)
- ○ Non-binary / third gender (3)

What is your age?

- ○ <=28 (1)
- ○ 29-42 (2)
- ○ 43-57 (3)
- ○ >=58 (4)

What is the highest degree or level of education you have completed?

- ○ Primary or secondary school (1)
- ○ Bachelor degree (2)
- ○ Higher than bachelor degree (3)

What is your monthly income?

- ◯ Less than €500 (1)
- ◯ €500 - €1000 (2)
- ◯ €1000 - €1499 (3)
- ◯ €1500 - €2499 (4)
- ◯ More than €2500 (5)

What is your nationality?

____

End of Block: Demographics

Figures

References

1. Ajzen, I. (2005). *EBOOK: Attitudes, Personality and Behaviour*. McGraw-hill education (UK).
2. Allport, G. W. (1937). Personality: A psychological interpretation.
3. Aycock, M., Cho, E., & Woods, L. (2022, September). Understanding Online Second-Hand Luxury Fashion Shopping Motivations and Perceived Value Among Young Adult Consumers. In *International Textile and Apparel Association Annual Conference Proceedings* (Vol. 78, No. 1). Iowa State University Digital Press.
4. Bardhi, F., & Arnould, E. J. (2005). Thrift shopping: Combining utilitarian thrift and hedonic treat benefits. *Journal of Consumer Behaviour: An International Research Review*, *4*(4), 223-233.
5. Barnes, L., & Lea-Greenwood, G. (2010). Fast fashion in the retail store environment. *International Journal of Retail & Distribution Management*, *38*(10), 760-772.
6. Beatty, S. E., & Ferrell, M. E. (1998). Impulse buying: Modeling its precursors. *Journal of retailing*, *74*(2), 169-191.
7. Cervellon, M. C., Carey, L., & Harms, T. (2012). Something old, something used: Determinants of women's purchase of vintage fashion vs second-hand fashion. *International Journal of Retail & Distribution Management*, *40*(12), 956-974.
8. Costa Jr, P. T., & McCrae, R. R. (1992). The five-factor model of personality and its relevance to personality disorders. *Journal of personality disorders*, *6*(4), 343-359.
9. DataCamp Team (July, 2018). Hierarchical Clustering in R. https://www.datacamp.com/tutorial/hierarchical-clustering-R
10. De Raad, B. (2000). *The big five personality factors: the psycholexical approach to personality*. Hogrefe & Huber Publishers.
11. De Raad, B. E., & Perugini, M. E. (2002). Big five assessment. Hogrefe & Huber Publishers.
12. DeLong, M., Heinemann, B., & Reiley, K. (2005). Hooked on vintage!. *Fashion Theory*, *9*(1), 23-42.
13. Donnellan MB, Lucas RE. (2008) Age differences in the Big Five across the life span: evidence from two national samples. Psychol Aging. 2008 Sep;23(3):558-66. doi: 10.1037/a0012897. PMID: 18808245; PMCID: PMC2562318.

14. Doxa . (May 20, 2022). Leading reasons for shopping second-hand in Italy in 2022 [Graph]. In *Statista*. Retrieved May 27, 2023, from https://www-statista-com.scd-rproxy.u-strasbg.fr/statistics/976687/reasons-for-shopping-second-hand-in-italy/
15. Duong, C. D. (2022). Big Five personality traits and green consumption: bridging the attitude-intention-behavior gap. *Asia Pacific Journal of Marketing and Logistics*, *34*(6), 1123-1144.
16. Dynata. (September 15, 2022). Share of people who are buying more secondhand products to cut costs worldwide as of September 2022, by generation [Graph]. In *Statista*. Retrieved May 27, 2023, from https://www-statista-com.scd-rproxy.u-strasbg.fr/statistics/1353230/cost-saving-buying-secondhand-worldwide-generation/
17. Evans, F., Grimmer, L., & Grimmer, M. (2022). Consumer orientations of secondhand fashion shoppers: The role of shopping frequency and store type. *Journal of Retailing and Consumer Services*, *67*, 102991.
18. Eysenck, H. J. (1990). Genetic and environmental contributions to individual differences: The three major dimensions of personality. *Journal of personality*, *58*(1), 245-261.
19. Ferraro, C., Sands, S., & Brace-Govan, J. (2016). The role of fashionability in second-hand shopping motivations. *Journal of Retailing and Consumer Services*, *32*, 262-268.
20. Fujiwara, K., & Nagasawa, S. Y. (2015). Relationships among purchase intentions for luxury brands and personality traits based on the Big Five. *American Journal of Industrial and Business Management*, *5*(11), 631.
21. Gohary, A., & Hanzaee, K. H. (2014). Personality traits as predictors of shopping motivations and behaviors: a canonical correlation analysis. *Arab Economic and Business Journal*, *9*(2), 166-174.
22. Goldberg, L. R. (1981). Language and individual differences: The search for universals in personality lexicons. *Review of personality and social psychology*, *2*(1), 141-165.
23. Goldberg, L. R. (1990). An alternative" description of personality": the big-five factor structure. *Journal of personality and social psychology*, *59*(6), 1216.
24. Grasso, M. M., McEnally, M., Widdows, R., & Herr, D. G. (2000). Consumer behavior toward recycled textile products. *Journal of the Textile Institute*, *91*(2), 94-106.
25. Guiot, D., & Roux, D. (2010). A second-hand shoppers' motivation scale: Antecedents, consequences, and implications for retailers. *Journal of retailing*, *86*(4), 355-371.

26. Herjanto, H., Scheller-Sampson, J., & Erickson, E. (2016). The increasing phenomenon of second-hand clothes purchase: Insights from the literature. *Jurnal manajemen dan kewirausahaan*, *18*(1), 1-15.
27. *Hierarchical Cluster Analysis* (n.d.), *UC Business Analytics R Programming Guide*. (n.d.). https://uc-r.github.io/hc_clustering
28. Hirsh, J. B. (2010). Personality and environmental concern. *Journal of environmental psychology*, *30*(2), 245-248.
29. James, G. M., Witten, D., Hastie, T., & Tibshirani, R. (2021). An Introduction to Statistical Learning. In Springer texts in statistics. Springer International Publishing. https://doi.org/10.1007/978-1-0716-1418-1
30. Jang KL, Livesley WJ, Vernon PA (1996) Heritability of the big five personality dimensions and their facets: a twin study. J Pers. 1996 Sep;64(3):577-91. doi: 10.1111/j.1467-6494.1996.tb00522.x. PMID: 8776880.
31. Kawulur, A. F., Sumakul, G., & Pandowo, A. (2022) Purchase Intention of Second-Hand: A Case Study of Generation Z. In *SHS Web of Conferences* (Vol. 149, p. 02026). EDP Sciences.
32. Ki, C., Lee, K., & Kim, Y. K. (2017). Pleasure and guilt: how do they interplay in luxury consumption?. *European Journal of Marketing*, *51*(4), 722-747.
33. Koay, K. Y., Cheah, C. W., & Lom, H. S. (2023). Does perceived risk influence the intention to purchase second-hand clothing? A multigroup analysis of SHC consumers versus non-SHC consumers. *Journal of Product & Brand Management*, *32*(4), 530-543.
34. Laitala, K., & Klepp, I. G. (2018). Motivations for and against second-hand clothing acquisition. *Clothing cultures*, *5*(2), 247-262.
35. Lin, P. H., & Chen, W. H. (2022). Factors That Influence Consumers' Sustainable Apparel Purchase Intention: The Moderating Effect of Generational Cohorts. *Sustainability*, *14*(14), 8950.
36. Lissitsa, S., & Kol, O. (2021). Four generational cohorts and hedonic m-shopping: association between personality traits and purchase intention. *Electronic Commerce Research*, *21*, 545-570.
37. Mano, H., & Elliott, M. T. (1997). Smart shopping: the origins and consequences of price savings. *ACR North American Advances*.
38. Matzler, K., Bidmon, S., & Grabner-Kräuter, S. (2006). Individual determinants of brand affect: the role of the personality traits of extraversion and openness to experience. *Journal of product & brand management*.

39. McCrae, R. R., & Costa Jr, P. T. (2008). The five-factor theory of personality.
40. Milfont, T. L., & Sibley, C. G. (2012). The big five personality traits and environmental engagement: Associations at the individual and societal level. *Journal of Environmental Psychology*, *32*(2), 187-195.
41. Mitchell, M., & Montgomery, R. (2010). An examination of thrift store shoppers. *The Marketing Management Journal*, *20*(2), 94-107.
42. Momox. (January 20, 2022). Share of second-hand clothing consumers in Germany in 2021, by gender [Graph]. In *Statista*. Retrieved May 29, 2023, from https://www-statista-com.scd-rproxy.u-strasbg.fr/statistics/1341141/second-hand-clothing-consumers-gender-germany/
43. Niinimäki, K., Peters, G., Dahlbo, H., Perry, P., Rissanen, T., & Gwilt, A. (2020). The environmental price of fast fashion. *Nature Reviews Earth & Environment*, *1*(4), 189-200.
44. Norman, W. T. (1963). Toward an adequate taxonomy of personality attributes: Replicated factor structure in peer nomination personality ratings. *The journal of abnormal and social psychology*, *66*(6), 574.
45. Ögel, İ. Y. (2022). Is it sustainability or fashion? Young educated consumers' motivations for buying second-hand clothing. *Business & Management Studies: An International Journal*, *10*(3), 817-834.
46. Padmavathy, C., Swapana, M., & Paul, J. (2019). Online second-hand shopping motivation–Conceptualization, scale development, and validation. *Journal of Retailing and Consumer Services*, *51*, 19-32.
47. Petrescu, M., & Bhatli, D. (2013). Consumer behavior in flea markets and marketing to the Bottom of the Pyramid. *Journal of Management Research*, *13*(1), 55-63.
48. Piedmont, R. L. (1998). The revised NEO Personality Inventory: Clinical and research applications.
49. Pookulangara, S., & Shephard, A. (2013). Slow fashion movement: Understanding consumer perceptions—An exploratory study. *Journal of retailing and consumer services*, *20*(2), 200-206.
50. Pretner, G., Darnall, N., Testa, F., & Iraldo, F. (2021). Are consumers willing to pay for circular products? The role of recycled and second-hand attributes, messaging, and third-party certification. *Resources, Conservation and Recycling*, *175*, 105888.

51. Rammstedt, B., & John, O. P. (2007). Measuring personality in one minute or less: A 10-item short version of the Big Five Inventory in English and German. *Journal of research in Personality*, *41*(1), 203-212.
52. Roux, D., & Guiot, D. (2008). Measuring second-hand shopping motives, antecedents and consequences. *Recherche et Applications En Marketing (English Edition)*, *23*(4), 63-91.
53. Şener, T., Bişkin, F., & Kılınç, N. (2019). Sustainable dressing: consumers' value perceptions towards slow fashion. *Business Strategy and the Environment*, *28*(8), 1548-1557.
54. Sherry Jr, J. F. (1990). A sociocultural analysis of a Midwestern American flea market. *Journal of consumer research*, *17*(1), 13-30.
55. Sprotles, G. B., & Kendall, E. L. (1986). A methodology for profiling consumers' decision-making styles. *Journal of Consumer Affairs*, *20*(2), 267-279.
56. Statista (2022) Second-hand e-commerce. Digital & Trends
57. Stone, J., Horne, S., & Hibbert, S. (1996). Car boot sales: a study of shopping motives in an alternative retail format. *International Journal of Retail & Distribution Management*, *24*(11), 4-15.
58. Sun, Y., Wang, S., Gao, L., & Li, J. (2018). Unearthing the effects of personality traits on consumer’s attitude and intention to buy green products. *Natural Hazards*, *93*, 299-314.
59. Tsao, W. C., & Chang, H. R. (2010). Exploring the impact of personality traits on online shopping behavior. *African journal of business management*, *4*(9), 1800.
60. Turkyilmaz, C. A., Erdem, S., & Uslu, A. (2015). The effects of personality traits and website quality on online impulse buying. *Procedia-Social and Behavioral Sciences*, *175*, 98-105.
61. Westbrook, R. A., & Black, W. C. (1985). A motivation-based shopper typology. *Journal of retailing*.
62. Wiggins, J. S. (Ed.). (1996). *The five-factor model of personality: Theoretical perspectives*. Guilford Press.
63. Xu, Y., Chen, Y., Burman, R., & Zhao, H. (2014). Second-hand clothing consumption: a cross-cultural comparison between A merican and C hinese young consumers. *International Journal of Consumer Studies*, *38*(6), 670-677.
64. Yan, R. N., Bae, S. Y., & Xu, H. (2015). Second-hand clothing shopping among college students: the role of psychographic characteristics. *Young Consumers*.

65. Zaman, M., Park, H., Kim, Y. K., & Park, S. H. (2019). Consumer orientations of second-hand clothing shoppers. *Journal of Global Fashion Marketing*, *10*(2), 163-176.
66. Zarley Watson, M., & Yan, R. N. (2013). An exploratory study of the decision processes of fast versus slow fashion consumers. *Journal of Fashion Marketing and Management: An International Journal*, *17*(2), 141-159.