ERASMUS UNIVERSITY ROTTERDAM

Erasmus School of Economics

Master Thesis Data Science & Marketing Analytics

"Examining drivers of consumer behavior among Western Europeans related to masstige products using a machine learning approach"

Name student: Desirée Gründler

Student ID number: 615219

Supervisor: Marina Khismatullina

Second assessor: Michiel Van Crombrugge

Date final version: 30.06.2022

Table of Contents

Abbreviations

DT Decision Tree

LR Logistic Regression

ML Machine Learning

OOB Out-Of-Bag

RF Random Forest

SL Super Learner

1. Introduction

"While value brands satisfy, luxury brands delight." – Hagtvedt & Patrick (2009)

Different motivations drive different consumers around the world to buy different products. These motivations may vary depending on the consumer's attitude, culture, or environment, as well as the need the product is intended to fulfill (Solomon, 2013). Due to its promising profit margins that are far from justifiable product prices the luxury segment has always been of great interest to marketers and researchers. It is a well-researched area in the field of marketing, which brought a lot of helpful insights for marketing managers to light.

Close to the luxury segment, there exists a relatively new segment including so-called "masstige" products, where masstige is a neologism consisting of the words mass and prestige (Silverstein & Fiske, 2003). Silverstein and Fiske (2003) coined the term and defined it as "prestigious but attainable" brands which are marketed as luxurious but are relatively inexpensive and mass-produced, hence, this segment is positioned between the luxury and the mass segment (Paul, 2018). The strategy is to distribute premium products to the maximum number of (middle-class) customers while prices are much lower than traditional luxury products (Truong et al., 2009). Since this new-luxury segment is promising to be highly profitable, the main purpose of this study is to investigate drivers of masstige consumption within consumers living in Western European countries, to estimate their purchase behavior related to a masstige product, and finally to derive managerial implications for marketers who distribute masstige products or want to distribute them within this area since the right ingredients can lead to highly successful strategies.

Many drivers of luxury consumption have been identified and researched, as well as differences in motivations, attitudes, or cultures (e.g., Drèze & Nunes, 2009; Rucker & Galinsky, 2008; Sun et al., 2014). There is consensus in research that luxury consumption is used as an indicator of exclusivity, uniqueness and status signaling in society (Belk, 1985), which can be assured by high prices and limited sales volumes at limited point of sales. This strategy proved to be successful in many examples. One of the best-known cases that perfectly combines the beforementioned characteristics might be the Birkin Bag from Hermès: a bag with a waiting list of up to two years, starting from a purchase price of around €7000, only on personal request and not in the assortment of Hermès stores or on their website available, tailored to personal requirements of the buyer, and worldwide used as a status symbol (Wüpper, 2016). Obviously, such luxury consumption is not affordable for the average person, but there is increasing interest in luxury brands, which can be deduced from increasing sales numbers and volumes within the industry. While the volume of sales in the overall luxury consumer goods market has nearly doubled over the past 15 years, sales at LVMH (Moët Hennessy Louis Vuitton) have quadrupled, and sales at Hermès have even increased six-fold over the same period (Statista, 2022a, 2022b, 2022c).

Improved economic factors like lower production costs, lower unemployment rates, and an increasing work rate among women may be reasons that have led to an increasing interest in luxury consumption (Silverstein & Fiske, 2003). Companies can take advantage of the increasing interest in luxury consumption to promote and distribute their products to a broader mass at a lower price so that the lower classes of society who couldn't buy them in the past can now afford them (Solomon, 2013). Kumar et al. (2020) describe masstige marketing as "the next big paradigm shift in brand management" and Solomon (2013) noticed mass marketing moving upscale. This growing attention of customers and luxury brands towards masstige opens new possibilities for marketers and needs to be investigated separately (Paul, 2015) since the target group differs to the average luxury consumer (Truong et al., 2009). The approach of masstige marketing clearly diminishes the traditional concept of exclusiveness and uniqueness of luxury products (Purohit & Radia, 2022). Therefore, it should be treated as a pre-level to traditional luxury brands, but as a level above mid-price segment (Paul, 2018). Brands like Michael Kors, Calvin Klein, or Victoria's Secret belong to the masstige segment since they create a trade-off between mass market and prestige (Purohit & Radia, 2022). Whereas the market is already making use of the strategy to target this new category near the luxury segment, the research is not exhaustive. Many researchers have acknowledged the masstige segment, as well as the existence of masstige marketing strategies (e.g. Das et al., 2021; Paul, 2015; Purohit & Radia, 2022), but the research is still in its infancy phase. Existing research focusses mainly on the terminology (Silverstein & Fiske, 2003), positioning (Truong et al., 2009) and differentiation to the traditional luxury segment (Paul, 2015, 2018), whereas less research is done on examining underlying motivations and drivers of the purchase intention towards masstige brands and products. Nevertheless, these are crucial factors to tailor marketing efforts and communications to a specific target group to successfully distribute a product and position a brand in the market (Homburg, 2017).

As a first step in this direction, Purohit & Radia (2022) explored factors that affect behavioral intentions towards masstige brands in India. This is a good basis for further research and brings initial insights for marketers. Nevertheless, cultural differences have been detected among the consumers of luxury products (Gentina et al., 2016; Zhang & Zhao, 2019) hence, it is necessary to prove external validity across cultures also within the masstige segment. To the best of the authors' knowledge, no research has been found that relates to Western European consumers in the context of masstige consumer behavior. Additionally, besides the factors which have been already considered in the study by Purohit & Radia (2022), many other factors are known to drive the consumption of luxury products that have not previously been studied in a masstige context. Accordingly, the scope of this study is both building a better theory of masstige consumption, as well as providing an empirically proven and understandable basis for marketing managers to tailor marketing efforts in the masstige segment.

Existing research on this topic agrees with the lack of theory and underlines the need of investigating the masstige segment (e.g., Kumar et al., 2020; Paul, 2015, 2018; Purohit & Radia, 2022; Silverstein & Fiske, 2003). First, little research has been done on investigating specific attitudes or motivations towards masstige products. Second, existing research on masstige consumption does not represent Western European countries and the results are not necessarily valid across borders. Third, existing research mostly uses complicated statistical models that are based on many assumptions and are difficult for marketing managers to understand and adapt.

Thus, in addition to the beforementioned purpose of this study, this work also seeks to combine marketing research with the field of machine learning to provide marketers and researchers a relatively easy-to-use and understandable but also reliable model with which they can investigate drivers of (masstige) consumption to derive useful directions for marketing efforts. Although machine learning methods have received increasing attention in various research areas (e.g., in biology, see van der Laan et al., 2007; in physics, see Varmuza et al., 2003; for fraud detection see Viaene et al., 2002) they are rarely used for marketing research purposes, even though different researchers have verified their usability in the field of marketing (e.g., De Caigny et al., 2018; Lemmens & Croux, 2006; Lessmann et al., 2015). Therefore, this study utilizes different machine learning methods, as well as a traditional method to investigate a data set gathered separately for this purpose to derive several insights and compare the different approaches. The data contain information about different consumer attitudes, motivations, and demographics that could theoretically influence their purchase behavior.

By addressing these research gaps, several contributions are done to the literature. First, this thesis focusses on examining additional factors which could theoretically drive masstige consumption to create an expanded empirically proven basis for masstige marketing strategies. Second, it compares prior research across cultures to check external validity. Third, it contributes to the approach of combining marketing research with machine learning methods. Finally, it builds a framework for marketers who can make use of this approach to investigate their own brand or product, even if the data set is much bigger than the one used for this study. This is especially important in the era of big data and enormous online consumer engagement.

This thesis is structured as follows. The next section maintains an overview of existing literature on consumer behavior, luxury consumption, and masstige brands and strategies. This is followed by the methodology section including a brief introduction to machine learning in marketing, the analytical approach of this thesis, and detailed explanations of the methods of random forests, logistic regression, stacking, and hierarchical clustering, as well as their measurement and implementation. Afterwards, the used data set is introduced, and descriptive analyses are performed on it. Then, the results section follows, leading to a general discussion, methodological discussion, theoretical contribution, managerial implications, limitations and directions for further research, and a brief conclusion.

2. Literature Review

The following section aims to provide an overview of the existing literature in the field of consumer behavior, luxury and masstige and to present its findings.

2.1. Consumer Behavior

In order to predict and influence consumer behavior, it is necessary to understand internal psychological processes, as well as external determinants that affect an individual and therefore influence their behavior and the decisions they make.

Psychological processes related to consumer behavior can be divided into activating and cognitive processes. Activating processes are those that activate an action or behavior due to a stimulus, whereas cognitive processes are those that classify and decode the stimulus (Kroeber-Riel & Gröppel-Klein, 2019). These stimuli can be internal, such as thoughts, or external, i.e., perceived externally via sense organs. With every cognitive process, new information is linked to a specific stimulus, which is included in future assessment processes (Kroeber-Riel & Gröppel-Klein, 2019). Motivation arises from subjective human needs, which can be utilitarian (describes functional needs) or hedonistic (describes needs for pleasure or enjoyment) (e.g., Solomon, 2013). If an individuum detects a discrepancy between a given state and a desired state that wants to be achieved, they form the motivation to pursue a specific goal which reduces or eliminates the discrepancy (Solomon, 2013). Attitudes are an interplay of motivation and cognitive processes to evaluate an object or thing based on how much it contributes to the satisfaction of a specific motivation (Kroeber-Riel & Gröppel-Klein, 2019).

Maslow's (1943) hierarchy of needs divides human needs into physiological, safety, social, esteem, and self-actualization ones (see Figure 2.1). The model states that each demand is satisfied by humans in ascending order. Since physiological and safety needs are to a large extent covered in the western world due to state and political structures, society here focuses more on love and belongingness, esteem, and self-actualization. The motivation to achieve a desired state increases or decreases according to how achievable the desired goal appears to the individual (Schouten, 1991).

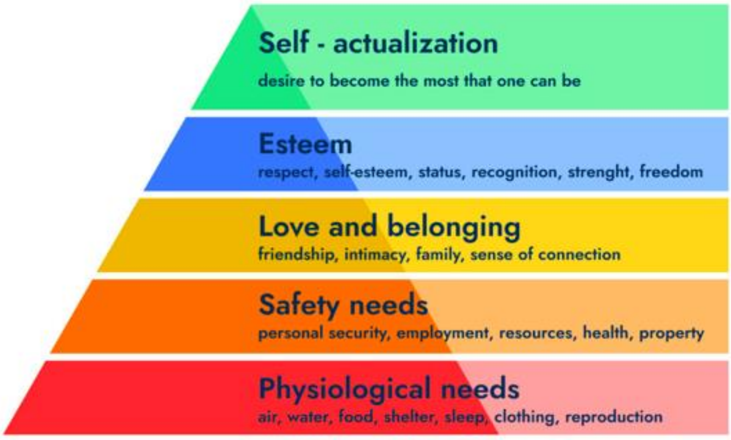


Figure 2.1: Maslow's Hierarchy of Needs

To fulfill the psychological needs of love, belongingness, and esteem, we early start building a self-concept about the perception of ourselves and how we want to be seen by others to place ourselves into the construct of society (Mummendey, 1997). A general distinction is made between the "real" self-image and the "ideal" self-image. The real self-image describes how the person sees themselves, regardless of the objective realism, whereas the ideal self-image refers to a construct of attitudes and characteristics that appear to the individual as a desirable self (Solomon, 2013). Based on this self-concept individuals make decisions and behave in a certain way to achieve the desired version of themselves or simply act coherently to be perceived as authentic from others (Mummendey, 1997). The mere theory seems very simple and grants all a place somewhere in society, nevertheless the self-concept is not static, but driven by a continuous need for prestige, accomplishment, and self-fulfillment (Maslow, 1943). The social psychologist Gordon Allport (1937) even sees self-enhancement as a central goal of a human's existence. Therefore, humans strive to get better and gain a higher status. The motivation to reduce a self-discrepancy in order to achieve the goal of an ideal self is often reflected in the consumer behavior of the individual, since consumer goods can have a psychological value that transcends their utilitarian value (Mandel et al., 2017).

External determinants and the understanding of one's own self in society or in a group also plays a major role. There are several so-called reference groups through which an individual can identify. Those can be divided into associative, aspired, and demarcated reference groups (Hoyer et al., 2013). While the associative reference group represents the one to which an individual currently belongs, the aspired reference group is that which the individual admires and wants to belong to, and the demarcated reference group is the one individuals want to distance themselves from. This grouping goes hand in hand with the real and ideal self-concept described above. In addition to personal influence, reference groups also have a normative influence. This ensures sanctions, punishment, or ridicule as soon as an individual does not comply with socially applicable norms (Hoyer et al., 2013). The human urge for social belonging and validation leads individuals to a tendency to behave in a group-conforming manner rather than individually, for fear of rejection or social exclusion from the group (Hoyer et al., 2013).

Different decisions require different levels of emotion and cognition, with extensive decisions (e.g., buying a car) requiring a high proportion of cognition and emotion, while habitual decisions (e.g., buying cotton pads) tend to be less cognitive and emotional (Kroeber-Riel & Gröppel-Klein, 2019). The first step that leads to a choice is the problem identification, i.e., the recognition of a discrepancy between an ideal and an actual state. The larger this discrepancy, the more motivated is an individual to reduce it. After information processing has taken place for the individual in a decision-oriented manner, the consumer proceeds to purchase the item or service. Emotions or expected emotions play a role here, especially in the case of extensive decisions, but also availability, personal resources such as income, or the quality of the product itself (Homburg, 2017). The final step in the decision-making process is evaluating the purchased product. Here it is important to what extent the expectations of the consumer have been fulfilled and how satisfied the person is with it. The focus is therefore on whether and to what extent a discrepancy could be resolved or reduced with the help of the product. This in turn affects the attitude towards the purchased product, which influences future decision-making behavior.

2.2. Luxury Consumption

Since the masstige segment is positioned close to the luxury segment (Truong et al., 2009), it is crucial to take a closer look at the luxury segment to understand its consumers and drivers.

Luxury products can be defined as goods that give the consumer status and prestige beyond any functional utility (Goor et al., 2020). They are associated with buzz words like superior quality, craftmanship, indulgence, and aspirational lifestyle (Kapferer, 1997). Luxury in the field of marketing is often defined as products and brands that rank at the top of the product and brand hierarchy (e.g., Morhart et al., 2020). It is interchangeable used with the term “status goods” since consumers from all social statuses perceive luxury as a status symbol (Husic & Cicic, 2009). Anderson et al. (2015) argue that possessing status is a fundamental human motive because people with higher status enjoy higher levels of positive affect, life satisfaction, and self-esteem. In contrast, people with lower status experience more negative affect, depression and anxiety, and negative physical health outcomes, such as higher blood pressure (Anderson et al., 2015). Whereas status used to be a person's place in the ancient social hierarchy in which they were born or which they achieved through ordainment, it changed during the last centuries, and status is nowadays linked to a person's achievements, which we assume will lead to success and wealth (De Botton, 2004), as well as power and influence (Han et al., 2010). A common way to show off achievements to a broader society is the consumption of luxury products to signal status and express one's identity (Belk, 1985). A lot of research has been done on investigating luxury consumption. Besides signaling status, luxury products can also be used to restore and maintain status (Gao et al., 2009), symbolizing privilege, and being superior or better than others (Drèze & Nunes, 2009). Luxury consumption is closely related with the motivation to impress others (Wiedmann et al., 2009), get socially recognized, respected, and better treated (McFerran & Argo, 2014), as well as feeling powerful (Rucker & Galinsky, 2008) and proud (Bellezza & Keinan, 2014). In addition to the function of communicating with society, luxury products are also used for the purpose of quality (utilitarian need), reward, or indulgence (hedonic needs), as well as compensatory for a lack of interpersonal relationships (Kroeber-Riel & Gröppel-Klein, 2019). Therefore, a distinction between luxury consumption in public and in private should be made since luxury can have a different meaning to the environment compared to the self (Kroeber-Riel & Gröppel-Klein, 2019). American teens, for example, want to buy luxury products because they hope to be perceived as unique, whereas French teens see it as an opportunity to be accepted by society (Gentina et al., 2016). Heine & Phan (2011) further distinguish luxury products by their concrete product characteristics (price, quality, and rarity) and abstract product characteristics (aesthetics, extraordinariness, and symbolic meaning).

Obviously, different motivations can drive consumers to buy luxury products, but besides the motivation to buy a luxury item it is also necessary to have the resources (e.g., money) to buy it. Therefore, luxury consumption is highly dependent on the level of income or wealth (Han et al., 2010), but also shaped by cultural values and the ideology in a society, which affect materialism and attraction towards luxury consumption (Sun et al., 2014). Besides the above stated positive outcomes and connotations, luxury consumption can also have negative effects, both, related to the self and society. Indulging in luxury consumption can entail feelings of guilt and regret if the consumer considers the consumption as wasteful (Kivetz & Simonson, 2002), as well as convey a feeling of unreasonable status or privilege which leads the consumer to feel inauthentic (Goor et al., 2020). It is noteworthy that this effect emerges across people of all income levels. Furthermore, luxury consumption can also lead to negative social aftereffects. Consumers of luxury products are often seen as less social and arrogant because they are perceived as focusing on making impressions (McFerran et al., 2014). Likewise, it can be depreciated when the consumption is not based on one's own merit (e.g., using parents' money) since the idea of signaling success and achievements then no longer applies (Lee et al., 2018).

2.3. Masstige Brands & Strategy

With their article „Luxury for the Masses", Silverstein and Fiske (2003) introduced a new way of luxury consumption. This "new-luxury" goods seek to satisfy the consumer's aspiration for a better life through products that are well-designed, well-engineered, and well-crafted to convey the appearance of a traditional luxury good but are priced way below traditional luxury goods. The researchers argue that the emergence of the need for new-luxury brands can be traced back to two main changes within society. First, a change on the demand-side of consumption, and second, a change on the supply-side of consumption. The necessity on the demand side appeared trough higher real income rates, rising home equity, and the changing role of women and family structures in society, but also due to the nationwide availability of discount retailers for daily goods leading to minimized expenses for daily needs (Silverstein & Fiske, 2003). On the other hand, the change on the supply-side is driven by better educated entrepreneurs, more sophisticated and accessible knowledge about consumer behavior, and the access to flexible supply-chain networks and cheaper, as well as global resources (Silverstein & Fiske, 2003). As a result, consumers engage in trading-up products that are important to them and trading down those that aren't. Accordingly, this new-luxury consumption is not necessarily linked to a higher income level (Silverstein & Fiske, 2003).

Contrary to traditional luxury goods, these new luxury goods are available to the mass middle market and are therefore sold in high sales volumes regardless that they reach prices far higher than the average middle-class products. Silverstein & Fiske (2003) divided this category into (1) downward extensions of traditional luxury brands like Armani Exchange or Versace Jeans Couture which are, compared to the mother brand, relatively cheap, (2) accessible super-premium products like Starbucks coffee, and (3) mass prestige or "masstige" brands, that are placed between mid-priced and super-premium products like Coach or Victoria's Secret. It is important to stress out that all three categories can follow a masstige strategy but only the latter ones are called masstige brands or products.

Nevertheless, targeting the mid-price segment is always challenging since the consumers within this category are in general more price sensitive and not very loyal (Homburg, 2017). The economist Michael Porter (1980) describes firms suffering from a poorly defined marketing strategy as "stuck in the middle". Indeed, new-luxury products have high potential to fail, if companies or brands are unable to meet the expectations of a better life, since they are unqualified to compete with the prices of low-cost products. Therefore, it is important to define a clear position within the segment, as well as separate the brand or goods from other mid-range products. Goor et al. (2020) argue that new-luxury brands need to define a reasonable level of perceived prestige to differentiate them from middle-range brands and justify a higher price. Therefore, the critical success factor lies in the perception of new-luxury products closer to traditional luxury brands to ensure that the consumer's emotional engagement outweighs higher costs (see Figure 2.2; Truong et al., 2009). The perceptual map derived from the study of Truong et al. (2009) which assumed Calvin Klein and Polo Ralph Lauren as masstige brands can be seen in Figure 2.3. In their study, they furthermore confirmed that new-luxury brands can successfully be marketed using the masstige strategy but arose doubt of brand dilution, if the brand is a downward extension of a traditional luxury brand.

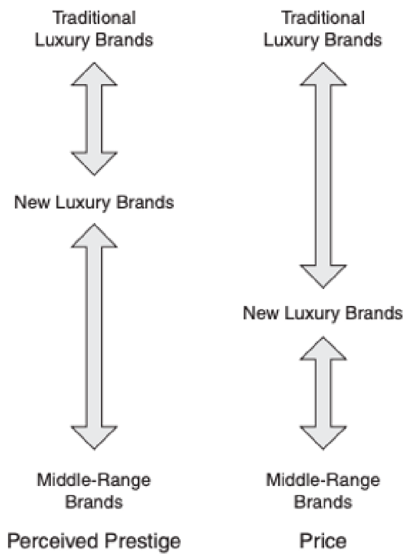


Figure 2.2: The perceived prestige of new-luxury products is closer to traditional luxury brands, while the price is closer to mid-range brands (Truong et al., 2009)

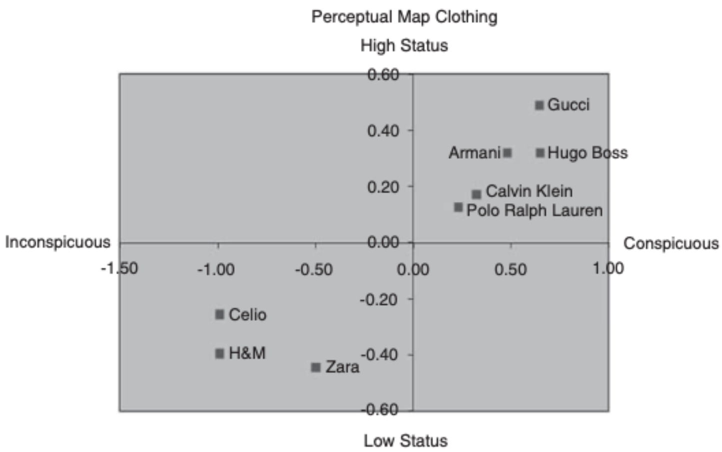


Figure 2.2: Perceptual map indicating Calvin Klein and Polo Ralph Lauren as masstige brands (Truong et al., 2009)

To circumvent concerns about assuming brands fruitfully market as masstige, Paul (2015) conducted a Masstige Mean Score which indicates the likelihood of a brand to succeed based on a masstige strategy, and indicates e.g., Apple as masstige brand. A study of Kumar et al. (2021) also revealed the Apple iPhone as a masstige brand, as well as the consumption of masstige brands leading to brand happiness.

Purohit & Radia (2022) were the first researchers to conceptualize buying behavior towards masstige brands. They analyzed the influence of different value perceptions that are rooted in the luxury value framework (Wiedmann et al., 2007) and brand aspirations rooted in conspicuous consumption theory (Truong et al., 2010) on masstige purchase intention. They found a significant positive relationship between functional values (e.g., quality), vanity values (e.g., enhancing appearance), achievement signaling (e.g., emulating role models) and the purchase intention of masstige products, and insignificant relationships between experiential values (e.g., uniqueness), social recognition (e.g., appreciation from society) and the purchase intention of masstige products. Furthermore, they found that masstige purchase intention generates brand happiness among customers and increases attitudinal loyalty. Nevertheless, the study was conducted in India, hence, these findings may not be valid across boarders as several researchers noted cultural differences in luxury context (e.g., Hung et al., 2011; Sun et al., 2014).

3. Methodology

For this research, several common machine learning methods are used, with machine learning defined as "computing the capacity of computers to learn and adapt without following explicit instructions, by using algorithms and statistical models to analyse and infer from patterns in data" (OED, n. d.). In other words, data is used to explore relationships within it and predict future observations (James et al., 2021). This can happen in a supervised or unsupervised manner, meaning that a specific outcome is either present or absent prior to the analysis (James et al., 2021). In addition, the choice of methods is also based on the scaling of the outcome variables (Homburg, 2017). Since the dependent variable in this case is binary, the methodological focus is on solving a classification problem, hence only methods that can handle a classification task are considered. The methodology section starts with a brief overview of marketing research in machine learning, leading to a discussion of different methods. Then, the analytical approach of this work is explained, followed by a more detailed explanation of the chosen models. This section ends with the implementation of the methods.

3.1. Machine Learning in Marketing

Although machine learning methods are used in various fields of society like healthcare (van der Laan et al., 2007) or legal administration (Barocas & Selbst, 2016), as well as in the marketing business e.g., for churn prediction (De Caigny et al., 2018) or product recommendations (Amatriain & Basilico, 2012), marketing research does not strongly make use of it yet. Despite the wide range of possible applications, marketing researchers often stick to traditional statistical methods, although there are already initial analyses based on machine learning methods that outperform traditional methods (e.g., Lemmens & Croux, 2006; Lessmann et al., 2015). Miguéis et al. (2017) compared different approaches to predict direct marketing response in the banking sector and point out that the random forest technique outperforms traditional methods like logistic regression but also other machine learning models (e.g., Support Vector Machines). Similar results can be derived from Lessmann et al. (2015) who compared 41 different classifiers in the field of marketing credit scoring. They found several methods significantly outperforming the industry standard method of logistic regression and recommend using RF as benchmark model for comparing different methods. Furthermore, they found heterogeneous ensemble classifier like the method of stacking as especially well performing. De Caigny et al. (2018), as well as Lemmens & Croux (2006) discover different tree-based methods as the strongest classifiers in customer churn prediction, and Kumar et al. (2019) supports this finding in the prediction of consumer repurchase intention. Finally, Lilhore et al. (2021) describes the random forests method as "an efficient filter in high-dimensional data to reliably classify consumer behavior factors". Therefore, machine learning methods have proven themselves in solving marketing problems, characterized by very accurate predictions, and resolving old problems in new ways. Nevertheless, regression analysis is still one of the most widely used analysis methods in market research (Homburg, 2017). In addition, cluster analysis is a common method for reducing complexity within the data (Homburg, 2017). It is used for market segmentation, which seeks to divide a heterogeneous market into homogeneous clusters which can then be investigated separately (Homburg, 2017).

3.2. Method Selection

Based on the previous review on machine learning methods in the field of marketing research, the methods of random forests (RF), logistic regression (LR), stacking, and cluster analysis are utilized in the scope of this research. Several reasons led to the decision to use a variety of model. In addition to the performance of a model, each method has additional advantages and limitations. Since drivers can change over time (D. Dubois et al., 2021), it is necessary for marketers to update and adjust findings in an easy and fast way, ideally based on a small set of data. Since logistic regression is based on explicit assumptions about the distribution of the variables, the input data needs to be investigated and accordingly transformed to ensure reliable results (James et al., 2021). In contrast, the RF method is a non-parametric method which means that no strong assumptions are made (James et al., 2021). Therefore, the danger to use data that does not meet the requirements is completely avoided. However, non-parametric methods typically require more observations to obtain accurate results (James et al., 2021). Furthermore, the method of random forest can model non-linear relationships between input features and the target class (James et al., 2021), so there is no pre-specification of the relationship needed. Moreover, random forests perform well even in the presence of multicollinearity among predictors, which can lead to problems in logistic regression because variables that are actually significant are indicated as not significant (Sarstedt & Mooi, 2019). Besides technical data pre-processing and assumptions, consumers often process different attributes of a product in a certain way to evaluate and eventually buy them (Sarstedt & Mooi, 2019). Therefore, it is crucial to investigate the relative importance of the different features. Tree-based methods can capture these relative priorities among customers easily and hence identify the key drivers of the independent variable (Breiman, 2001). Nevertheless, in contrast to logistic regression, RF models cannot compute the significance level of the input variables, hence the error probability is unknown. Lastly, stacking outperforms logistic regression and tree-based models in predictions (e.g., Džeroski & Ženko, 2004; Lessmann et al., 2015) but is considered as a "black box" algorithm which veils the exact contribution of each base learner to the prediction (Naimi & Balzer, 2018).

3.3. Analytical Approach

The analysis is divided into two parts: a global analysis and a narrowed local analysis. The global analysis aims to gain general insights from the entire data set (including all available observations and variables). To do so, multiple methods are used: random forest, logistic regression and stacking to combine random forest and logistic regression. Using multiple methods has the advantage that the individual insights can be combined to reach more robust conclusions and to avoid the disadvantages of each method in general. Therefore, the RF method is used to get the relative variable importance among the input features, while the estimated coefficients derived from the LR indicate the direction of the variables and their influence on the outcome, as well as their significance level. Lastly, stacking seeks to advance the accuracy of prediction among the models and can also serve as a robustness check.

The local analysis focuses on subsets of the complete data set. First, cluster analysis is applied to segregate the data which may provide further insight or increase the prediction accuracy since the objects in the subgroups are more similar to each other. Afterward, the different clusters will be explored separately by a machine learning method. Although stacking promises better results, a RF approach will be used to explore the different clusters. This is due to the necessity to provide a test set to measure the performance of stacked models which implies splitting the data further. This should be avoided since the number of observations per cluster will be relatively small. A RF on the other hand can indicate the model performance based on the OOB error and does not necessarily require a test set (Breiman, 2001). Additionally, the relative variable importance within the cluster will be computed so that the different clusters can be compared easily. Afterward, a subset of variables is investigated further. This subset contains all variables that belong to the top 5 in the global analysis or in at least one of the derived clusters. Since all observations can be used for this part of the local analysis, it is possible to split the data for prediction, hence, the most promising method of stacking is used in this case. This part of the local analysis can indicate the loss of information if only a few features are considered, and whether less important variables are still adding value to the analysis. Lastly, an additional hierarchical cluster analysis is performed on only the subgroup of interest to gain insights into differences within the sample. See Figure 3.1 for an overview of the analytical approach of this thesis.

Analytical approach of the thesis

Global Analysis: *using whole data set*

- Random Forest
  - → Prediction
  - → Variable importance
- Logistic Regression
  - → Prediction
  - → Influence of variables
- Stacking
  - → Prediction
  - → Variable importance
  - → Influence of variables
  - → Robustness check

Local Analysis: *using subsets*

- Hierarchical clustering
  - → Clusters including subsets of observations, but all variables
- Random Forest on each cluster
  - → Variable Importance
- Stacking (SL) on subset of most important variables, but all observations
  - → Prediction
  - → Variable importance
  - → Influence of variables
- Additional cluster analysis
  - → Insights on subgroup of interest

Figure 3.1: Overview analytical approach of the thesis

3.4. Random Forests

The method of random forests was first introduced by Breiman (2001) and is a combination of multiple decision trees, leading to improved accuracy, lower error rates, and more robust results compared to a single decision tree. In order to understand the procedure of random forests, it is necessary to understand the methodology of decision trees and bagging as well, as it is fundamental for the random forest approach. Since the dependent variable is binary, the algorithm is based on decision trees for classification, also called classification trees.

3.4.1. Classification Trees

Classification trees are a subcategory of decision trees (DT), which is a supervised learning method to classify observations into pre-existing classes. It was introduced by Breiman et al. (1984) as a non-parametric machine learning algorithm. It is a graphical model that splits the data step by step into smaller subgroups using a top-down approach which means the starting point is at the top of the tree (so-called root node). At each split, an observation is assigned to another subgroup (internal node) and eventually ends up in one of the pre-defined classes. These classes are represented by so-called leaf nodes which are based on the majority vote of the actual outcome of the observations that fall into the specific leaf node. It is a greedy algorithm, which means that each split is done based on the best possible result on exactly that split, without taking the following splits into consideration (James et al., 2021). To decide which variables and which thresholds to choose for each split within a classification tree, a measurements called Gini impurity can be used (James et al., 2021). It measures the impurity of a node and thus searches for a split that leads to the lowest impurity of the resulting nodes. Purity means that all observations within a subgroup belong to the same class. Its goal is to gain the purest nodes possible, and it is defined as 1 minus the sum of all squared probabilities belonging to the class i, whereas c represents the number of target classes as seen in (3.1).

$$Gini\ impurity = 1 - \sum_{i=1}^{c}(p_i)^2 \tag{3.1}$$

The Gini impurity is based on the true outcomes of the data set, ranges between 0 and 0.5 (whereas zero indicates purity) and is calculated at each node for all possible splits of a variable (Aggarwal, 2014). Afterward, the split leading to the lowest Gini impurity is chosen.

3.4.2. Bagging

Unfortunately, decision trees are very unstable and non-robust (James et al., 2021). Therefore, Breiman (1996) introduced the idea of ensemble methods, that combine multiple classifiers based on bootstrapped data sets. Bootstrapping is a tool which helps to improve the robustness of machine learning models by building different subsets of the original training data (James et al., 2021). It is used with ensemble methods to reduce variance and increase accuracy (James et al., 2021). The subsets are built from replicated but randomly chosen observations of the original training data set based on sampling with replacement. This can lead to duplicated data points, as well as missing observations within the bootstrapped data set, since the new training data has the same number of observations as the original one. These new training sets (B) are used as the basis for training separate prediction models on the b_tth bootstrapped training set, which can lead to different predictions. A majority vote among the prediction outcomes $\hat{f}^{*t}(x)$ builds a new classifier $\hat{f}_{bag}(x)$ that is called Bagging (James et al., 2021) and is defined in (3.2).

$$H(x) = \hat{f}_{bag}(x) = arg\,max \sum_{t=1}^{B} \hat{f}^{*t}(x) \tag{3.2}$$

See Algorithm 3.1 for a description of the Bagging algorithm (Aggarwal, 2014). The number of bootstrapped subsets (B) and hence the number of independent classifiers is determined depending on the OOB error, which is explained in the following part.

Algorithm 3.1 Bagging

Input: Training data $D = \{x_i, y_i\}_{i=1}^{n}$ $(x_i \in \mathbb{R}^m,\ y_i \in \Upsilon)$
Output: Ensemble classifier H

  for t ← 1 to B do

    Build a bootstrapped subset b_t by randomly sampling with replacement in D

    Learn base classifier h_t based on b_t

  end for

  return $H(x) = arg\,max_{y \in \Upsilon} \sum_{t=1}^{B} 1(h_t(x) = y)$

As described above, bagging uses bootstrapped data sets that are based on sampling with replacement. As a result, some observations are not used for prediction in this specific model. On average, each bagged classifier uses around two-thirds of the observations of the original training set (James et al., 2021). The remaining observations that are left out in the specific model are assigned to the Out-Of-Bag (OOB) data set. Since the specific model is not based on these OOB observations, they can be used as test set to measure the accuracy of the independent model (Breiman, 2001). After the bagged model is built, the outcome of each OOB observation is predicted by all independent models that are not built on this specific observation, and eventually classified based on the average or majority vote of these separate models. The proportion of incorrectly classified OOB observations is called the OOB error, which gives reliable indications of how well the model performs on new data (Breiman, 2001). Furthermore, it is used to determine the number of bootstrapped data sets (B) that are needed to achieve stable OOB error rates (Breiman, 2001). Although a higher number of independent classifiers does not lead to overfitting (James et al., 2021), the computational time increases with an increase in B, since it specifies the number of models within the ensemble.

3.4.3. Random Forests

The method of random forests (RF) was initiated by Breiman in 2001 as an advanced bagging algorithm using an ensemble of decision trees. Due to the fact that the independent trees within a bagged model based on decision trees are highly correlated to each other (James et al., 2021), researchers used different approaches of randomness within their models to minimize correlation (e.g., Dietterich, 2000; Tin Kam Ho, 1998). Breimans method of RF makes use of randomly selected input variables at each split when growing the separate trees. Therefore, the method of RF can be described as a combined method of bagging and random feature selection (Breiman, 2001). The number of trees ($ntrees$) in a RF is equal to the number of bootstrapped subsets (B) and can be derived by the OOB error rate as described in the bagging algorithm. The number of random features ($mtry$) that are used for each split within a tree is approximately calculated as $mtry \approx \sqrt{p}$ that is the square root of the total number of independent variables (James et al., 2021). The independently grown trees are not pruned (Breiman, 2001) which means the size of the leaf node is equal to 1. After building B independent trees, each separate tree has a unit vote for the RF classifier, which is determined by the majority vote (Breiman, 2001). The algorithm is summarized in Algorithm 3.2 (Aggarwal, 2014). Since the method of RF is based on tree-structured classifiers, it is a non-parametric and supervised learning method.

Algorithm 3.2 Random Forest

Input: Training data $D = \{x_i, y_i\}_{i=1}^{n}$ $(x_i \in \mathbb{R}^m,\ y_i \in Y)$
Output: Ensemble classifier H
for t $\leftarrow$ 1 to B do
Build a bootstrapped subset b_t by randomly sampling with replacement in D
Use $mtry_t \approx \sqrt{p}$ at each split by randomly sampling features in $\mathbb{R}^m$
Find best feature $q_t^* \in mtry_t$ based on Gini impurity
Learn decision tree h_t
end for
return $H(x) = arg\,max_{y \in Y} \sum_{t=1}^{B} 1(h_t(x) = y)$

In addition to prediction, an estimate of variable importance can be derived from a RF (Breiman, 2001). It can be described as an overall summary of the importance of variables that provide the predictive accuracy (James et al., 2021). To rank the features accordingly, two measures can be used: the Gini impurity (MeanDecreaseGini) and the permutation measure (MeanDecreaseAccuracy) (Chapman & Feit, 2019). The Gini measurement calculates the average reduction in Gini impurity (3.1) across all trees, achieved by the splits using the specific feature (James et al., 2021). The permutation measurement assesses the variable impact on accuracy as follows: After a tree is constructed, the values of a specific variable are randomly permuted within the OOB data set. Then, the outcomes of the OOB data set including the permuted variable is predicted by the constructed tree. Afterwards, the model accuracy of the permuted OOB data set is compared to the accuracy of the original OOB data set. The bigger the decrease in accuracy (in percent), the more important is the variable for accurate predictions. This is repeated for all variables in the data set so that a ranking of variable importance averaged over all trees can then be derived. In other words, both measurements rank the variables according to their usefulness in correctly classifying the OOB data. (Breiman, 2001)

3.5. Logistic Regression

The underlying function of logistic regression was firstly introduced in the early 19th century to describe the growth of populations (Cramer, 2022). Nowadays, after several adjustments, it is mainly used in cases with qualitative response values due to its ability to model the probability that an observation falls into a pre-defined class (James et al., 2021). Therefore, it can serve as a supervised machine learning method for classification. The method is based on the logistic function, which calculates the probability of the class outcome X as seen in (3.3), where β_i are unknown constants called the model coefficients or parameters, and x_i represent the independent input variables (James et al., 2021).

$$p(X) = \frac{e^{\beta_0 + \beta_1 X_1 + \cdots + \beta_p X_p}}{1 + e^{\beta_0 + \beta_1 X_1 + \cdots + \beta_p X_p}} \tag{3.3}$$

While β_0 represents the constant called intercept, all other β_i are the slopes in the model. By manipulating the formula and taking the logarithm of both sides, the so-called log odds are derived which are represented by the left side of the equation (3.4) and are linear in X. (James et al., 2021)

$$log\left(\frac{p(X)}{1-p(X)}\right) = \beta_0 + \beta_1 X_1 + \cdots + \beta_p X_p \tag{3.4}$$

Therefore, a one-unit increase in any independent variable x_i leads to a change in log odds by e^{β_i}. Hence, the additional unit increases or decreases the probability that the observation falls into the positive class by $\left(e^{\beta_i} - 1\right) * 100$ ceteris paribus. (James et al., 2021)

The estimation of the coefficients is done by fitting the logistic model using the maximum likelihood method. Likelihood is the probability of an observed outcome y_i conditioned on its corresponding variable values x_i and a set of parameters β_i. The goal is to estimate those parameter values such that the predicted probabilities (3.3) for an outcome are as closely as possible to the observed outcome, where the probability is p if $y_i = 1$, and $1 - p$ if $y_i = 0$. This can be formalized by the likelihood function in (3.5) that seeks to make the occurrence of the observations most likely. Consequently, the likelihood function is maximized. (James et al., 2021).

$$L(\beta_0, \beta) = \prod_{i=1}^{n} p(x_i)^{y_i} \, (1 - p(x_i)^{1-y_i} \tag{3.5}$$

After the coefficients are estimated, the probability of a new observation can be computed using the logistic function as described above (3.3). After calculating the probability, the observation is classified to one of the pre-existing classes, depending on a set threshold (James et al., 2021). This threshold indicates a certain probability assigning the observations in a binary case either to class “zero” when below the threshold value, or class “one” when above. In general, the default value is set to 0.5 which is equal to a 50% probability.

Additionally, the method of logistic regression computes so-called p-values, indicating the probability that the relationship between an independent variable and a dependent variable is observed by chance (James et al., 2021). In other words, a smaller p-value indicates a higher probability that there is a real association between independent and dependent variable (James et al., 2021). Therefore, a threshold indicating a certain probability is used to decide which variables are taken into consideration when interpreting the model. Typically, this threshold is set at a probability of 5% (James et al., 2021), which corresponds to a p-value of 0.05 and is referred to as significance level. A potential problem regarding the significance of variables is multicollinearity which means that two or more independent variables are correlated to each other (James et al., 2021). It can lead to the problem that actually significant variables are indicated as not significant as their influence is captured by the correlated variable instead (Sarstedt & Mooi, 2019). Multicollinearity can be detected by computing the variance inflation factor VIF for each variable which is formulized in in (3.6), where $R^2_{X_i|X_{-i}}$ is the R^2derived by a regression of X_i on all other independent variables (James et al., 2021).

$$VIF\left(\hat{\beta}_i\right) = \frac{1}{1-R^2_{X_i|X_{-i}}} \tag{3.6}$$

The VIF ratio shouldn't exceed a value of 10 to ensure an unproblematic amount of collinearity (James et al., 2021).

3.6. Stacking

Stacked generalization or stacking is a machine learning method which combines several base classifiers to a so-called meta classifier. It was first introduced by Wolpert (1992) to minimize the bias of the base models by adding an additional meta-level. Nowadays, it is used as a supervised ensemble algorithm searching for the optimal combination of prediction models (Aggarwal, 2014). The main idea is to learn a certain number of base or first-level classifiers T on the original training set and use the predictions of these models as input to learn a meta or second-level classifier, whereas the true outcome labels remain as dependent variable Y (Wolpert, 1992). Therefore, a new data set is constructed containing $T + Y$ columns and the same number of observations as in original training set. The exact algorithm can be derived from Algorithm 3.3 (Aggarwal, 2014).

Algorithm 3.3 Stacking

Input: Training data $D = \{x_i, y_i\}_{i=1}^{n}$ $(x_i \in \mathbb{R}^m,\ y_i \in \Upsilon)$
Output: Ensemble classifier H
Step 1: Learn first-level classifiers
for t $\leftarrow$ 1 to T do
Learn a classifier h_t based on D
end for
Step 2: Construct new data sets from D
for $i \leftarrow 1$ to n do
Construct a new data set $\{x'_i, y_i\}$, where $x'_i = \{h_1(x_i), h_2(x_i), \dots, h_T(x_i)\}$
end for
Step 3: Learn second-level classifier
Learn a new classifier h' from the newly constructed data set
return $H(x) = h'(h_1(x), h_2(x), \dots, h_T(x))$

The methods that are used to train the meta classifier can be homogeneous or heterogeneous, and even other ensemble methods can serve as first-level classifiers (Aggarwal, 2014). In other words, any supervised machine learning method can be used as base classifier either the same or different ones. Additionally, any method can be used to train the meta-model (Aggarwal, 2014). Different to other ensemble methods, stacking does not only accumulate and average the predictions derived by the first-level methods, but learns how to best combine the models. Nevertheless, the nature of stacking is prone to overfit the data, therefore, the technique of cross-validation is used to train the base classifiers (Aggarwal, 2014). The combination of stacking with cross-validation is also called super learner (SL) (van der Laan et al., 2007).

3.6.1. Super Learner

To prevent models from overfitting the data cross-validation is a useful tool (James et al., 2021). The most commonly used technique is K-fold cross validation. It splits the training data randomly in a certain number K of disjoint subsets of the same size, whereas one of the K subsets is left out and used as test set, the other $K - 1$ sets are used to train the model on (James et al., 2021). This procedure is repeated over K iterations so that each of the subsets serves once as test set. Hence, K number of models are built within the cross-validation process. Typically, $K = 5$ is used as it is empirically shown that it yields good results (James et al., 2021). In the case of stacking, the predictions made for the sample that was hold-out and used as test set, are the predictions used for learning the meta learner. After fitting the second-level classifier, the first-level classifiers are re-trained on the full training set. This returns the final ensemble classifier, consisting of the T first-level models and the second-level model, which can then be used for predicting new data. This procedure avoids overfitting since the base learners are trained on a different data set than the second-level classifier. The super learner algorithm is summarized in Algorithm 3.4. (Aggarwal, 2014)

Algorithm 3.4 Super Learner

Input: Training data $D = \{x_i, y_i\}_{i=1}^{n}$ $(x_i \in \mathbb{R}^m,\ y_i \in Y)$
Output: Ensemble classifier H
  Step 1: Cross-validation approach
  Randomly split D into K equal sized subsets: $D = \{D_1, D_2, \dots, D_K\}$
  for k $\leftarrow$ 1 to K do
    Step 1.1: Learn first-level classifiers
    for t $\leftarrow$ 1 to T do
      Learn a classifier h_{kt} from D/D_k
    end for
    Step 1.2: Construct a training set for second-level classifier
    for $x_i \in D_k$ do
      Get a record $\{x'_i, y_i\}$, where $x'_i = \{h_{k1}(x_i), h_{k2}(x_i), \dots, h_{kT}(x_i)\}$
    end for
  end for
  Step 2: Learn second-level classifier
  Learn a new classifier h' from the collection of $\{x'_i, y_i\}$
  Step 3: Re-learn first level classifiers
  for t $\leftarrow$ 1 to T do
    Learn classifier h_t based on D
  end for
  return $H(x) = h'(h_1(x), h_2(x), \dots, h_T(x))$

If the methods LR and RF are used as first-level models within the SL approach, the model computes coefficients from the LR part and variable importance derived from the RF part. The coefficient estimation proceeds as described above using the likelihood-function (see 3.4). However, the variable importance derived by the RF part is calculated differently. To decide which variables and which thresholds to choose, the method makes use of the classification error rate which is defined in (3.7) where p_i is the probability that an observation belongs to the class i (James et al., 2021).

$$E = 1 - \max_i\ (p_i) \tag{3.7}$$

Hence, the variable importance is achieved by calculating the average reduction in the classification error across all trees (MeanDecreaseError) achieved by the splits using the specific feature.

3.7. Hierarchical Cluster Analysis

Hierarchical cluster analysis is an unsupervised machine learning method, searching for clusters within the data. It is used to determine subgroups with specific characteristics that are possibly related to the dependent variable (James et al., 2021). The main idea of cluster analysis is to group similar observations together. Goal of the algorithm is that observations within a cluster are as similar as possible, whereas the different clusters are as dissimilar as possible (James et al., 2021). To measure similarity, different measurements can be used. In this research, similarity is gained from the proximities derived from a random forest (Breiman & Cutler, n.d.). Since hierarchical cluster analysis is based on a distance matrix, the values in the proximity matrix are subtracted from 1, resulting in a distance matrix (Breiman & Cutler, n.d.). The exact procedure for obtaining proximities is described in the next paragraph.

The algorithm treats all observations n as their own cluster. The two clusters that are most similar according to the distance matrix are merged into one cluster. Hence, $n-1$ clusters are left. Afterward, the remaining clusters are merged according to their similarity, until only one big cluster is left. The procedure is illustrated in Algorithm 3.5. (James et al., 2021)

Algorithm 3.5 Hierarchical Clustering

Input: Data $D = \{x_i, y_i\}_{i=1}^{n}$ $(x_i \in \mathbb{R}^m,\ y_i \in \Upsilon)$
Output: N clusters
    for i $\leftarrow 1$ to N do
        Measure similarity based on Ward's distance
        Merge clusters based on $min(ESS)$
    end for
    return $N(x) = \arg\min\ (ESS)$

Since the clusters contain several observations at a certain point, similarity cannot be measured between single observations but a group of them. Therefore, a linkage technique is needed to define similarity between clusters containing more than one observation (James et al., 2021). In this case, Ward's minimum variance method is used. It clusters the observations according to the minimum error sum of squares (ESS) within a cluster, which is formulized in (3.8), where x_i is the score of the i^{th} individual observation, and n is the total number of observations (Ward, 1963).

$$ESS = \sum_{i=1}^{n} {x_i}^2 - \frac{1}{n} \left(\sum_{i=1}^{n} x_i\right)^2 \tag{3.8}$$

After the clusters are formed, their different characteristics can be visualized by a heatmap, which uses colors to represent the mean value of each variable within the cluster (James et al., 2021). An example heatmap can be seen in Figure 3.2, where lighter colors indicate a low mean value of the variable within the cluster, whereas darker colors represent a relatively high variable mean. This visualization tool is very intuitive for getting a fast overview of the different characteristics between clusters.

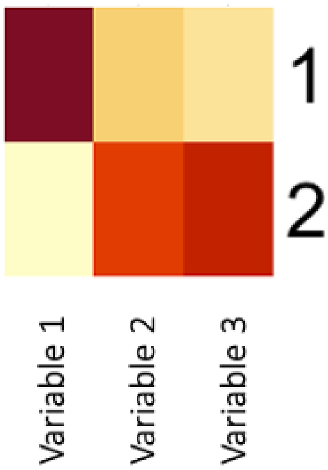


Figure 3.2: Example heatmap showing two clusters indicating very high values in cluster 1 for variable 1, but low values in cluster 2 for variable 1. Variables 2 and 3 have high values in cluster 2, and low values in cluster 1.

In the RF method, proximities can be calculated based on the similarity of observations (Breiman & Cutler, n.d.). Similarity in this case means that observations end as often as possible in the same leaf node among all independent trees in a RF. The proximities are originally based on a $n\;x\;n$ matrix, meaning that the matrix consists of a row for each sample, as well as a column for each sample. After building an independent tree, all data is run down the tree. If two samples end up in the same leaf node, their proximity value increases by one. After applying all samples to all trees, the matrix values are normalized. This means the proximities are divided by the number of independent trees within a forest ($ntrees$). All values of the normalized proximity matrix range between 0 and 1, whereas 0 indicates that the samples never ended up in the same leaf node, hence are highly dissimilar, a value of 1 indicates that the samples ended up in the same leaf node among all independent trees. The proximity matrix can also be derived in the case of an unsupervised RF. Then, the original data is considered as class 1, whereas a synthetically created second data set is considered as class 2. The class 2 observations consist of randomly copied values for each variable from the original data set. Therefore, the synthetic data set destroys the dependency structure in the original data but provides an artificial classification problem which can be handled by the random forest approach as described above. (Breiman & Cutler, n.d.)

3.8. Measurements

To evaluate the performance of a prediction model, different measurements can be used. This thesis focuses on the accuracy, as well as the sensitivity of the model. Both measurements are based on the confusion matrix of previously unseen observations of the data set, which summarizes the totality of right and wrong predictions (see Table 3.1; James et al., 2021).

	TRUE CLASS = 1	TRUE CLASS = 0
PREDICTED CLASS = 1	True Positive	False Positive
PREDICTED CLASS = 0	False Negative	True Negative

Table 3.1: Confusion Matrix

Whereas the accuracy captures the proportion of correctly classified observations, the sensitivity represents the false negative rate (James et al., 2021). This means the lower the false negative predictions, the higher the sensitivity. This measurement is used when predicting observations false negative could lead to negative consequences. The definitions can be derived from (3.9) and (3.10).

$$Accuracy = \frac{True\ Negative + True\ Positive}{True\ Negative + False\ Negative + True\ Positive + False\ Positive} \tag{3.9}$$

$$Sensitivity = \frac{True\ Positive}{True\ Positive + False\ Negative} \tag{3.10}$$

3.9. Implementation

For the data analysis and resulting graphics, the software environment R for statistical computing and graphics was used. To implement the methods of RF, LR, and SL the packages “randomForest” (Liaw & Wiener, 2002), “stats” (R Core Team, 2021), and “h2o” (LeDell et al., 2022) were used respectively. The plots were generated using the “ggplot2” package (Wickham, 2016). For cluster analysis, a combination of the “randomForest” package and the “stats” package is used.

4. Data

The following chapter covers data set related topics such as data collection, description, cleaning, and pre-processing, as well as descriptive analysis.

4.1. Collection & Description

The data on which the quantitative part of the thesis is based was collected through an online survey between May 7th and May 29th, 2022. The study was conducted in English and German on qualitrics.com and distributed via WhatsApp, Facebook, Instagram, and LinkedIn to people living in Western Europe, which includes the countries Austria, Belgium, France, Germany, Liechtenstein, Luxembourg, Monaco, Netherlands, and Switzerland (United Nations Statistics Division, n. d.). This resulted in a total number of 280 participants. Since research on masstige brands is still in its infancy phase and no clear clusters or target groups have yet been identified, a convenience sample was used, which means that the sample contains participants from that part of the population that is easy to reach (Sarstedt & Mooi, 2019). This implies that no specific group was targeted a priori and that there is only limited control over the participants type, which is highly influenced by situational factors (Sarstedt & Mooi, 2019). Nevertheless, to ensure a greater demographic diversity among the respondents, the snowball sampling method was used additionally. It means that existing participants shared the questionnaire with people within their reach (Sarstedt & Mooi, 2019).

The questionnaire utilizes existing scales (Kappes et al., 2021 for measuring attitudes towards wealth; Richins & Dawson, 1992 for materialism and compensatory; Barbopoulos & Johansson, 2017 for utilitarian factors and acceptance; Purohit & Radia, 2022 for attraction, appreciation, and appearance; Dubois et al., 2005 for uniqueness; Lay & Furnham, 2019 for impress, achievement, and success), as well as additional questions capturing factors derived from the literature review. Besides demographics, it consists of 33 items with response options on a 5-point Likert scale ranging from strongly disagree = 1 to totally agree = 5 (Sarstedt & Mooi, 2019). The dependent variable captures past purchasing behavior for a masstige product. For this study, the Apple iPhone was chosen as a representative masstige product since Kumar et al. (2021) and Paul (2015) have already demonstrated the Apple iPhone as a masstige product in previous research. Therefore, the outcome variable captures whether a person owns an Apple iPhone or not ("Do you have an iPhone?"). Related features were divided into six categories, including demographics, utilitarian, self-related, and society-related variables, various attitudes, as well as additional factors for further information. The entire questionnaire and its subcategories can be seen in Table 4.1. The variables were selected based on prior research on consumer behavior, which affirmed them having a causal impact on buying behavior in general (e.g. Solomon, 2013), and on more specific factors that appeared in research related to luxury consumption (e.g. Heine & Phan 2011; Belk, 1985), as masstige products are considered a subcategory of the luxury segment (Solomon, 2013; Paul, 2015). After a pretest in which seven individuals participated, several items were modified based on feedback on clarity, required time, design, and willingness to answer (Sarstedt & Mooi, 2019).

Category	Question	Measurement
Demographic	What is your gender?	Gender
Demographic	What is your age?	Age
Demographic	What is your annual income (before taxes)?	Income
Demographic	In which country have you mainly lived in the last 5 years?	Country
Demographic	In which country do you currently live?	Country
Additional Factors	Are you happy with the brand of your mobile phone?	Brand happiness
Additional Factors	Would you buy the brand of your current mobile phone again?	Satisfaction
Additional Factors	Did you buy your mobile phone using your own money?	Own merit
Additional Factors	Is the brand of your mobile phone important to you?	Trading up / down
Additional Factors	I can imagine buying a mobile phone of a different brand.	Brand loyalty
Additional Factors	I notice when somebody has an iPhone.	Awareness (iPhone)
Additional Factors	I notice when somebody does not have an iPhone.	Awareness (no iPhone)
Additional Factors	I'm likely to purchase luxury brands.	Attraction
Additional Factors	I feel guilty spending money on luxury products.	Guilt
Utilitarian	My mobile phone is of high quality.	Quality
Utilitarian	My mobile phone is resonably priced.	Price
Utilitarian	Deciding on the brand of my mobile phone was a matter of convenience.	Convenience
Self-related	I feel comfortable with the brand of my mobile phone.	Authenticity
Self-related	I can identify with the brand of my mobile phone.	Identification
Self-related	I am proud to own my mobile phone.	Proudness
Self-related	I bought my mobile phone to reward myself.	Reward
Self-related	The brand of my mobile phone makes me feel superior to others.	Superior
Self-related	The brand of my mobile phone makes me feel respected by others.	Respect
Self-related	The brand of my mobile phone makes me feel more powerful.	Power
Self-related	The brand of my mobile phone makes me feel exclusive.	Exclusiveness
Society-related	The brand of my mobile phone reveals who I am.	Uniqueness
Society-related	The brand of my mobile phone is approved by my friends.	Acceptance (friends)
Society-related	The brand of my mobile phone gets me appreciation from others.	Appreciation
Society-related	The brand of my mobile phone is perceived as prestigious.	Prestige
Society-related	My mobile phone is a good way to impress others.	Impress
Society-related	My mobile phone helps me enhance my appearance.	Appearance
Society-related	I use my mobile phone to signal status to society.	Status Signaling
Attitudes	I think the iPhone is a luxury product.	Luxury product
Attitudes	Money is a really good indicator of a person's life achievements.	Achievement
Attitudes	Money is a really good indicator of a person's success.	Success
Attitudes	Spending a lot indicates that someone is wealthy.	Wealth
Attitudes	I admire people who own expensive homes, cars, and clothes.	Materialism
Attitudes	I have all the things I really need to enjoy life.	Compensatory

Table 4.2: Questionnaire to capture possible predictors of iPhone ownership

4.2. Data Cleaning & Pre-Processing

The total of 280 responses collected was cleaned based on several basic principles. These include examining aborted, unrealistic, unlikely, and exceptional observations, as well as missing values and suspicious response patterns (Sarstedt & Mooi, 2019). Among all responses, 41 drop-outs were recorded, most of them without answering a single question. Observations with very unlikely age/income combinations (e.g., 16 years old with annual income above € 75,000) were removed, as were responses from participants from the outside of Western Europe. Further, individual missing values were examined and replaced by the median conditional on income, as they only occur within the age variable. The time spent completing the survey served as a measure to verify whether participants actually paid attention to the questions. Therefore, all responses with a duration of less than 120 seconds were removed since the author personally took 114 seconds in a test run after creating the questionnaire. Additionally, all observations between 120 and 150 seconds were investigated more closely, and another observation was removed after detecting suspicious response patterns within the second half of the questionnaire. After cleaning the data as described, 229 observations remained for the analysis. These respondents were split into a training set containing 187 observations (representing around 82% of the total data) which was used to train the different models on, and a test set including 42 unseen observations (around 18% of the total data) to verify the predictive accuracy.

4.3. Descriptive Analysis

First, the distribution of the dependent variable was checked. It contains 74.24 % participants with an iPhone and 25.76 % who own a mobile phone of another brand. In the test set, 73.81% are iPhone owners, which represents the so-called no-information rate. This is equal to the accuracy of the best guess, i.e., when all observations are simply assigned to the majority group.

Further, all independent variables were examined more detailed. Based on the demographics, the modus within the gender variable is female, within income it is less than € 20,000, within age it is generation Z (born after 1995), and within the country variable it is Germany. Whereas the gender variable is almost equally distributed between males and females, and the variables income and generation show at least some variance in their responses, it is particularly striking that 80% of the participants have lived mainly in Germany within the last 5 years. The exact details of the demographic variables can be found in Table 4.2. It contains the count and proportion of all respondents after cleaning the data. It is important to note that the country variable had more answer options than indicated in the table. This is due to the fact that no data was gathered for the categories Liechtenstein, Luxembourg, and Monaco. The range of the age variable is 17-74 years, with a mean value of 29.41 years.

Variables and Categories	Respondents (n = 229)	Proportion (%)
Gender		
Male	108	47,16%
Female	119	51,97%
Other	2	0,87%
Income		
< 20k €	108	47,16%
20k - 40k €	35	15,28%
41k - 55k €	24	10,48%
56k - 75k €	30	13,10%
> 75k €	27	11,79%
Unknown	5	2,18%
Generation		
X	13	5,68%
Y	105	45,85%
Z	108	47,16%
Unknown	3	1,31%
Country		
Austria	3	1,31%
Belgium	4	1,75%
France	2	0,87%
Germany	184	80,35%
Netherlands	17	7,42%
Switzerland	1	0,44%
Moved to Western Europe	18	7,86%

Table 4.3: Count and proportion of each demographic variable

Besides the demographics, the descriptive analysis indicates higher agreement with the questionnaire statements for the additional and utilitarian factors, while participants mostly disagree with the self-related and society-related statements. The modus as well as the median for each feature can be seen in Table 4.3. The exact distribution of each variable can be found in the Appendix 1. Since these numbers give a general sense of the data, Table 4.3 also compares the percentage of the total number of participants who totally agreed with a statement (regardless of which mobile phone they own) to participants who totally agreed with a statement but own an iPhone. The largest differences can be seen in the additional and utilitarian factors, whereas the agreement of self-related, society-related, and attitudes differs less.

Of all variables, product satisfaction (approx. a 9 percentage points difference), brand happiness (approx. a 10 percentage points difference), price (approx. a 10 percentage points difference), and brand loyalty (approx. a 14 percentage points difference) are the ones that differ the most. This means that iPhone users are on average more satisfied with their mobile phone and its quality, happier with their mobile phone brand, more loyal to this brand, but less satisfied with its price. Other noticeable variables that iPhone users agree with more strongly regarding their mobile phones are convenience, authenticity, acceptance by their friends, and prestige, as well as general enjoyment of life. It is also noteworthy that iPhone owners themselves consider the iPhone less of a luxury product than the average participant.

			Totally agree (5)	
Measurement	Modus	Mean	no iPhone %	iPhone %
Additional Factors				
Brand happiness	5	4.5	40.68	79.41
Satisfaction	5	4.6	47.46	81.76
Own merit	5	4.0	83.05	61.76
Trading up / down	4	3.7	11.86	37.06
Brand loyalty	1	2.8	59.32	7.65
Awareness (iPhone)	4	3.0	15.25	14.12
Awareness (no iPhone)	4	3.0	5.08	21.18
Attraction	1	2.6	11.86	7.65
Guilt	1	2.5	1.69	5.29
Utilitarian				
Quality	5	4.4	35.59	59.41
Price	4	3.4	47.47	10.00
Convenience	4	3.4	15.25	24.12
Self-related				
Authenticity	5	4.4	54.24	65.29
Identification	3	3.5	16.95	22.94
Proudness	3	3.0	8.47	12.94
Reward	1	2.2	5.08	5.29
Superior	1	1.6	0	0.59
Respect	1	1.7	0	2.35
Power	1	1.5	0	1.76
Exclusiveness	1	1.7	0	2.35
Society-related				
Uniqueness	1	1.7	0	2.35
Acceptance (friends)	3	2.9	0	10.59
Appreciation	1	1.9	0	1.18
Prestige	1	2.7	0	10.00
Impress	1	1.6	0	1.18
Appearance	1	1.7	0	2.35
Status Signaling	1	1.5	0	1.76
Attitudes				
Luxury product	4	3.5	33.9	22.94
Achievement	1	2.4	3.39	2.94
Success	4	2.7	5.08	4.71
Wealth	1	2.1	1.69	1.18
Materialism	1	2.4	6.78	7.06
Compensatory	5	4.1	37.29	44.12

Table 4.4: Descriptive Analysis stating the modus and mean of all variables, as well as the proportion of participants that totally agreed (=5 on the Likert scale) with the statement belonging to the variable split by non-iPhone owners and iPhone owners.

5. Results

The next section deals with the results of the analysis. It begins with the global results, followed by the local results, both including brief discussions. This section ends with an overall summary discussion.

5.1. Global Analysis

The global analysis, as described in the methodology section, is based on the totality of variables and observations available in the data set. This means that 37 independent variables operate as input to predict the dependent variable of iPhone ownership. The methods used for this part of the analysis are RF, LR, and SL. Table 5.1 provides an overview of the accuracy and sensitivity per model in the scope of the global analysis.

Model	Accuracy	Sensitivity
RF global	88.1%	93.55%
GLM global	90.48%	93.55%
SL global	95.24%	96.77%

Table 5.5: Model performance for the global analysis, indicating the stacked model performs best, both in accuracy and sensitivity

The RF model is built on the training set, using the parameters mtry = 6 and ntree = 2000, resulting in an OOB-error of 13.37%. A graphical representation of all OOB error rates derived from this entire study can be found in Appendix 2. After applying the model to the unseen test set, an accuracy of 88.1% is achieved, with a sensitivity of 93.55%. Since the difference between the OOB error and the actual test error is relatively small, with a difference of 1.47 percentage points, the OOB error appears quite reliable in this case. Figure 5.1 shows the variable importance plot including all predictors. It indicates the variables capturing brand loyalty and price as the most important features in classifying the data.

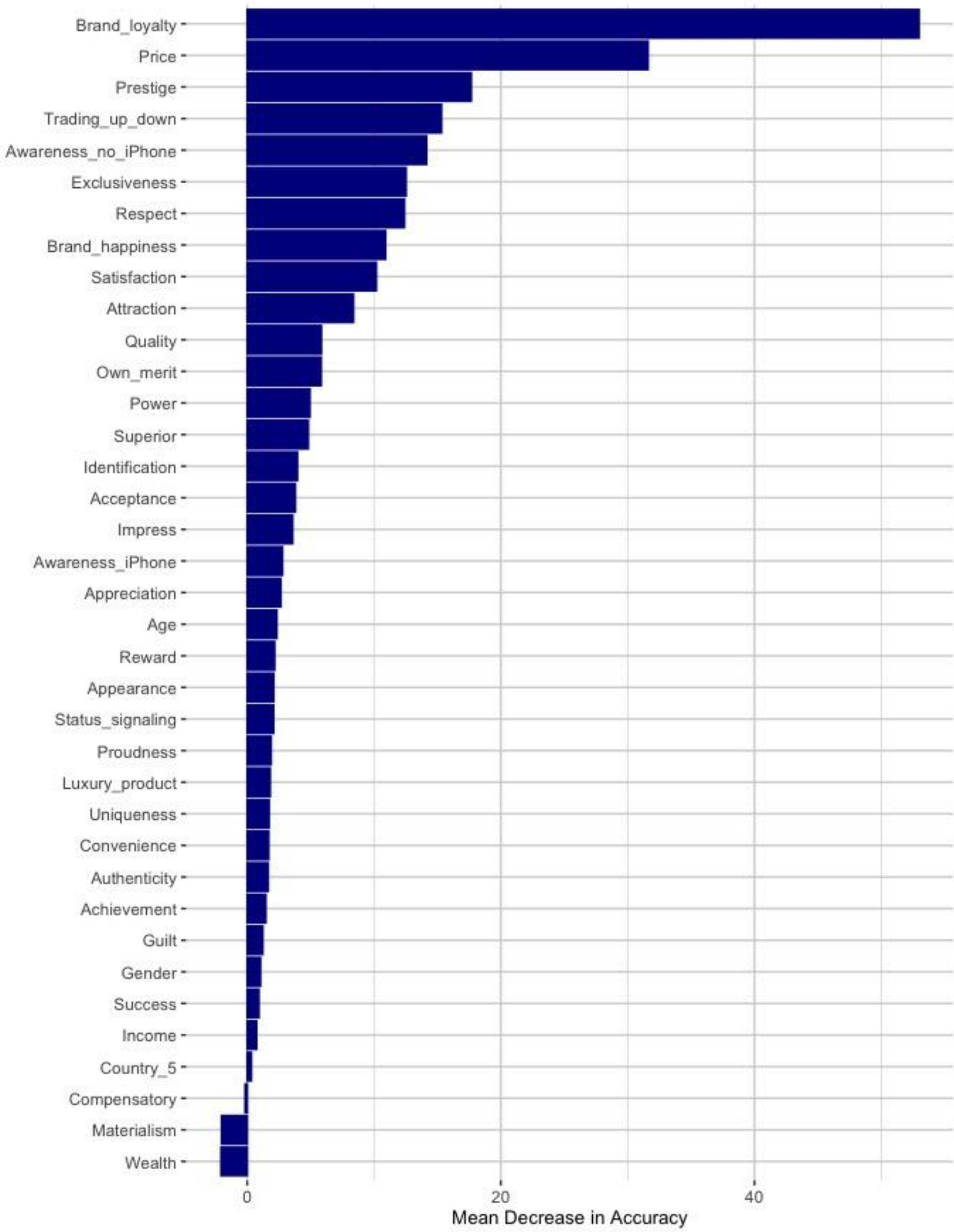


Figure 5.1: Variable importance plot based on the full data set (229 observations, 37 independent variables), indicating the variables capturing brand loyalty and price as most important in distinguishing iPhone owners from non-iPhone owners.

Next, a LR model was built on the training data set, resulting in an accuracy of 90.48% and a sensitivity of 93.55% after applying to the test set. To assign the observations to one of the classes, the default threshold of 50% probability was used. The outcoming significant variables (based on a significance level of 5%, equal to a p-value < 0.05) and their coefficients can be derived from Table 5.2. Since multiple testing could be a potential problem due to the large number of predictors, the threshold is adjusted using the Bonferroni correction. Taking this correction into account, only the variables brand loyalty and price remain significant at a 5% level. Additionally, it is important to note, that the input data do not face the problem of strong multicollinearity (VIF < 7) but violates the assumption of linearity between the independent variables and the log odds of the dependent variable (see equation 3.4 of the methodology section). Therefore, the results should not be used to take exact inferences from the model coefficients, but their signs can still indicate the direction of the corresponding variable. Hence, higher values in the variables capturing authenticity, brand loyalty, price, and whether the participant bought their mobile phone on their own merit lead to a lower probability of owning an iPhone, whereas higher values in age, prestige, and the awareness of other people not owning an iPhone lead to a higher probability of owning an iPhone. Note that the wording "higher values" related to brand loyalty and price does not mean "more", as the questionnaire asked in reverse. In other words: higher values in brand loyalty refer to less loyal customers, and higher values in price refer to customers who think their phone is not reasonably priced.

Variable	Estimated Coefficient	p-Values
Intercept	1.06	>0.01
Brand_loyalty	- 0.15	>0.01
Price	- 0.08	>0.01
Authenticity	- 0.11	0.01
Prestige	0.05	0.03
Age	0.01	0.03
Own_merit	- 0.04	0.02

Table 5.6: Coefficient estimations derived from the LR model, indicating the variables brand loyalty and authenticity as most influential to the probability of iPhone ownership. All variables presented are significant at a 5% significance level, but only brand loyalty and price remain significant on a 5% significance level after adjusting based on the Bonferroni correction.

As last part of the global analysis, a SL model based on the methods of RF and LR is built on the training set. Using 5-fold cross-validation leads to an accuracy of 95.24% and a sensitivity of 96.77% after applying the SL model to the test set. It indicates the variables brand loyalty and price as most important based on the RF part of the model. The estimated coefficients based on the LR part can be derived from Table 5.3, where some of the variables are missing since the SL model only estimates the coefficients of the variables that are used for prediction. Unfortunately, the p-values cannot be derived from the LR part of the model.

Variable	Estimated Coefficient
Intercept	3.17
Age	0.01
Brand_happiness	0.10
Satisfaction	0.02
Own_merit	- 0.09
Trading_up_down	0.07
Quality	0.17
Price	- 0.57
Authenticity	- 0.03
Brand_loyalty	- 0.67
Exclusiveness	0.03
Uniqueness	- 0.01
Prestige	0.17
Awareness_no_iPhone	0.10
Achievement	- 0.05
Compensatory	0.07

Table 5.7: Coefficient estimations derived from the LR part of the stacked model, indicating the variables brand loyalty and price as most influential to the probability of iPhone ownership.

Overall, the SL model outperforms the RF and LR models. Therefore, it should be used for prediction. Additionally, the SL model indicates the variables brand loyalty and price as most important which is in line with the RF model. Further, the list of influential features based on the SL model includes all significant coefficients derived from the LR model plus additional ones, to distinguish iPhone owners from non-iPhone owners. Although the indicated signs of the estimated coefficients match between the LR and the SL, their values differ partly very strongly, especially for the most influential variables. Therefore, the exact influence is not discussed further in this thesis, as the estimates do not appear robust, which could be due to the violation of assumptions mentioned above. Furthermore, the sensitivity is higher than the accuracy for all models. This means that the models are better at predicting iPhone ownership than ownership of alternative mobile phone brands. It appears that iPhone owners are more similar in terms of the questionnaire than people who use alternative mobile phone brands.

Next, the robustness of the variable importance is checked. Therefore, three methods of calculating the variable importance are used. First, based on the mean decrease in accuracy (derived from the RF model), second, based on the mean decrease in Gini impurity (derived from the RF model), and third based on the mean decrease in classification error (derived from the SL model). See Figure 5.2 for the different variable importance plots. All methods indicate brand loyalty as the most important variable in distinguishing iPhone owners from non-iPhone owners, followed by the variable capturing reasonable pricing. Further, the top 5 variables are the same across methods, although the ranking of the last ones differs. Nevertheless, it can be said that the ranking of variable importance among the top variables is somewhat robust.

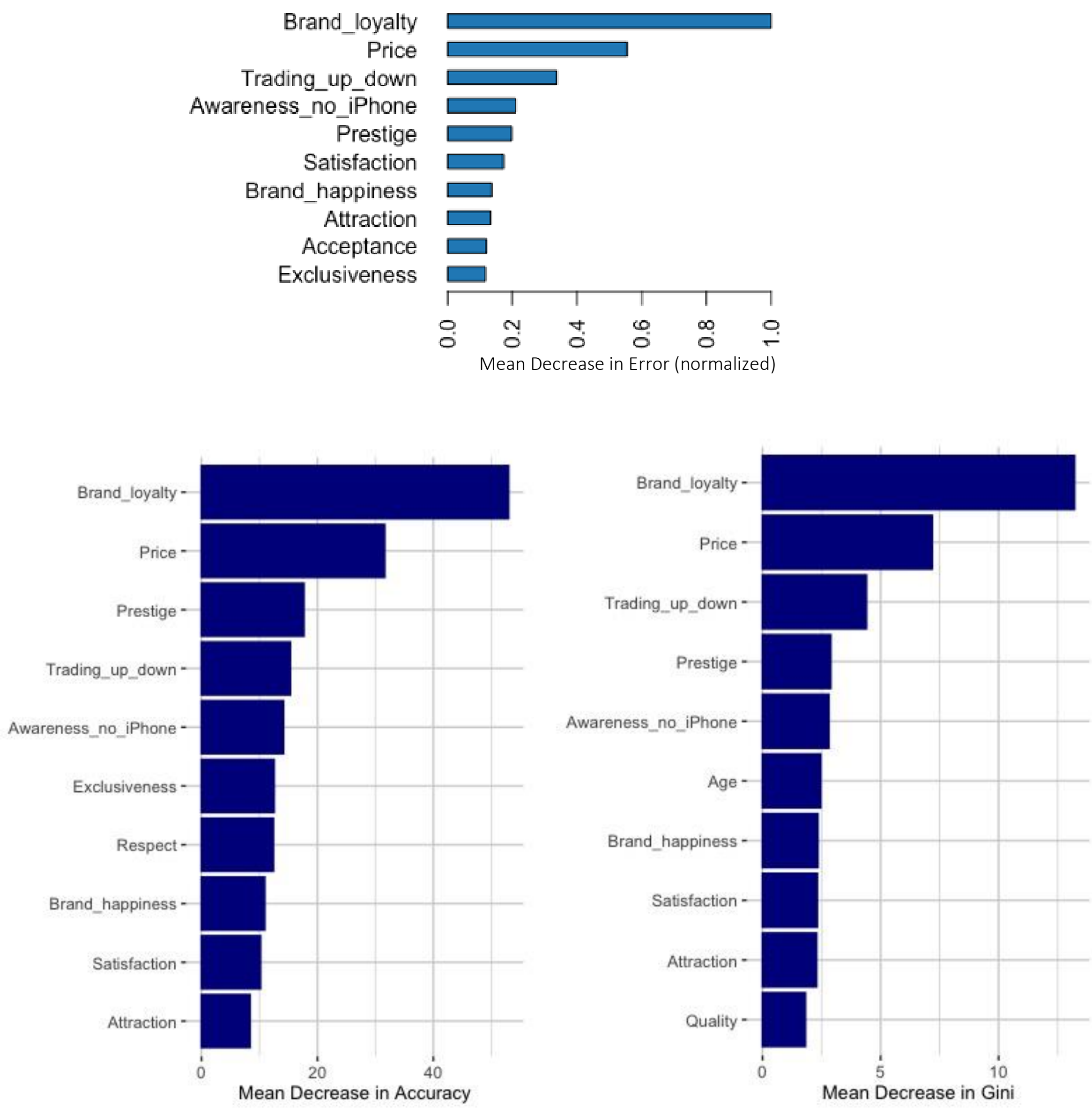


Figure 5.2: Variable importance plots based on the mean decrease in classification error, mean decrease in accuracy, and mean decrease in Gini impurity indicating consistently brand loyalty and price as most important variables, followed by prestige, trading the importance of one's mobile phone up / down, and capturing the awareness if somebody has no iPhone, in varying order across measurements

5.2. Local Analysis

The local analysis is performed on subsamples of observations, as well as on a subsample of variables. This leads to five additional models in total: two RF models based on a two-cluster solution derived by hierarchical clustering on all variables, two RF models based on a two-cluster solution derived by hierarchical clustering on only a subset of demographic variables, as well as a SL model based on all top 5 variables of the models up to this point. Additionally, hierarchical cluster analysis is performed on only observations belonging to iPhone owners to get an impression of possibly existing subgroups.

Clustering the full data into smaller subgroups is the first step of the local analysis. Based on an unsupervised RF, the data is split into two clusters. The clusters differ mainly in the variables capturing brand loyalty, income, age, price, and whether the mobile phone was bought on their own merit, whereas cluster 2 has on average higher values in the beforementioned features, compared to cluster 1. The heatmap can be derived from Figure 5.3 accordingly. Afterward, a supervised RF is performed on each cluster leading to an OOB-error of 12.24% for cluster 1 and 18.67% for Cluster 2. Note that the clusters are due to their size not further split into training and test sets and the OOB-error is used as performance measure. Cluster size, parameters of the RF, and OOB-errors can be derived from Table 5.4. Although predictions based on cluster 1 slightly improve the model performance compared to the main model in the global analysis, they perform worse when averaged with their affiliate cluster. This is due to the fact that the proportion of iPhone owners is higher in cluster 1, that can therefore be better predicted by the model. The variable importance of predicting both clusters is inspected, indicating that cluster 2 is in line with the global analysis indicating the variables price and brand loyalty as the most crucial ones, whereas cluster 1 ranks brand loyalty and the awareness of people not owning an iPhone on top. The variable importance plots can be seen in Figure 5.4. It indicates that iPhone users in cluster 1 are more concerned about factors detached from their mobile phone like exclusiveness, perceived prestige, or to impress which probably drive their brand loyalty. This claim is supported by the highly important variable of brand awareness if other people do not own an iPhone. Contrary, cluster 2 indicates utilitarian variables like price and quality as more important than any self- or society-related factor. Their brand loyalty seems to be rooted in variables like brand happiness, quality, and satisfaction with the product, hence, factors that are directly related to the product. The inferences about the directions are based on the coefficients estimated in the global analysis.

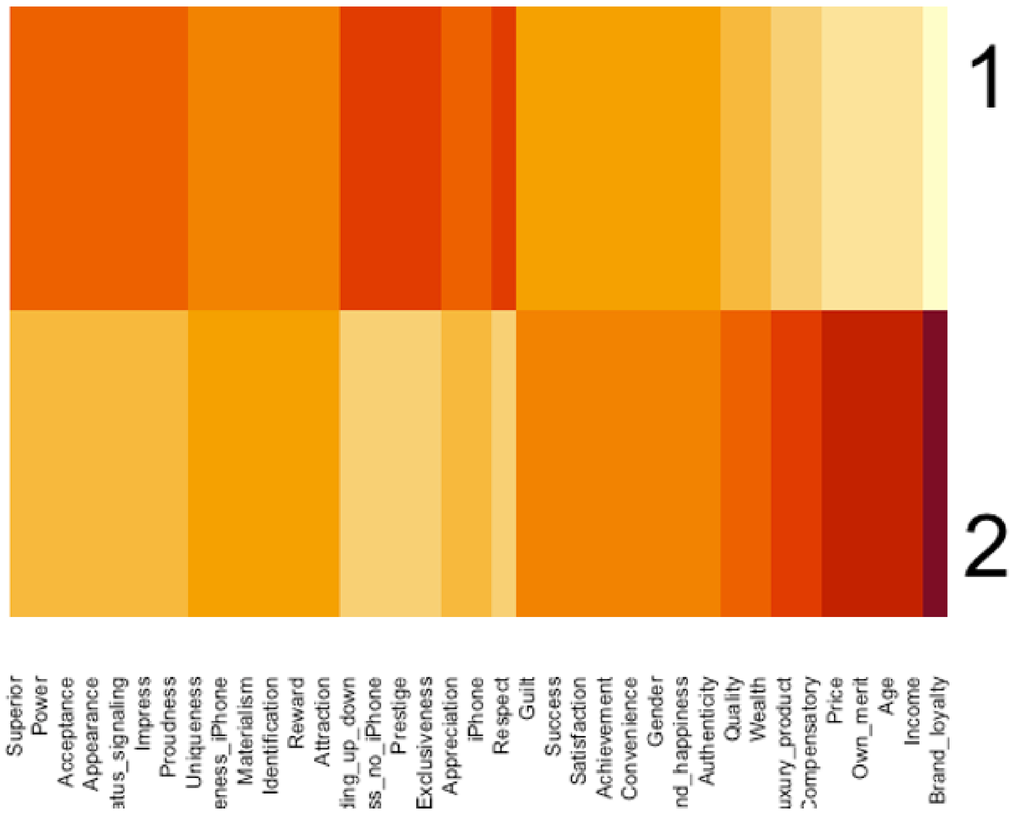


Figure 5.3: Heatmap of the clusters derived from an unsupervised RF. The clusters differ mainly in the variables capturing brand loyalty, income, age, price and whether the mobile phone was bought from own merit, whereas observations in cluster 2 have on average higher values in the mentioned features compared to cluster 1.

Model	Cluster size	OOB-error	mtry	ntree
Cluster 1 all	147	12.24%	6	1500
Cluster 2 all	75	20.00%	6	2000
Cluster female	115	13.91%	6	2000
Cluster male	107	12.15%	6	1500

Table 5.8: Cluster size, OOB-error, and parameters of all RF models built on subsets derived from cluster analysis. Some clusters show slight improvements compared to the main model in the global analysis but perform worse averaged with their affiliate cluster. The OOB-error rates leading to the number of trees can be derived from Appendix 2.

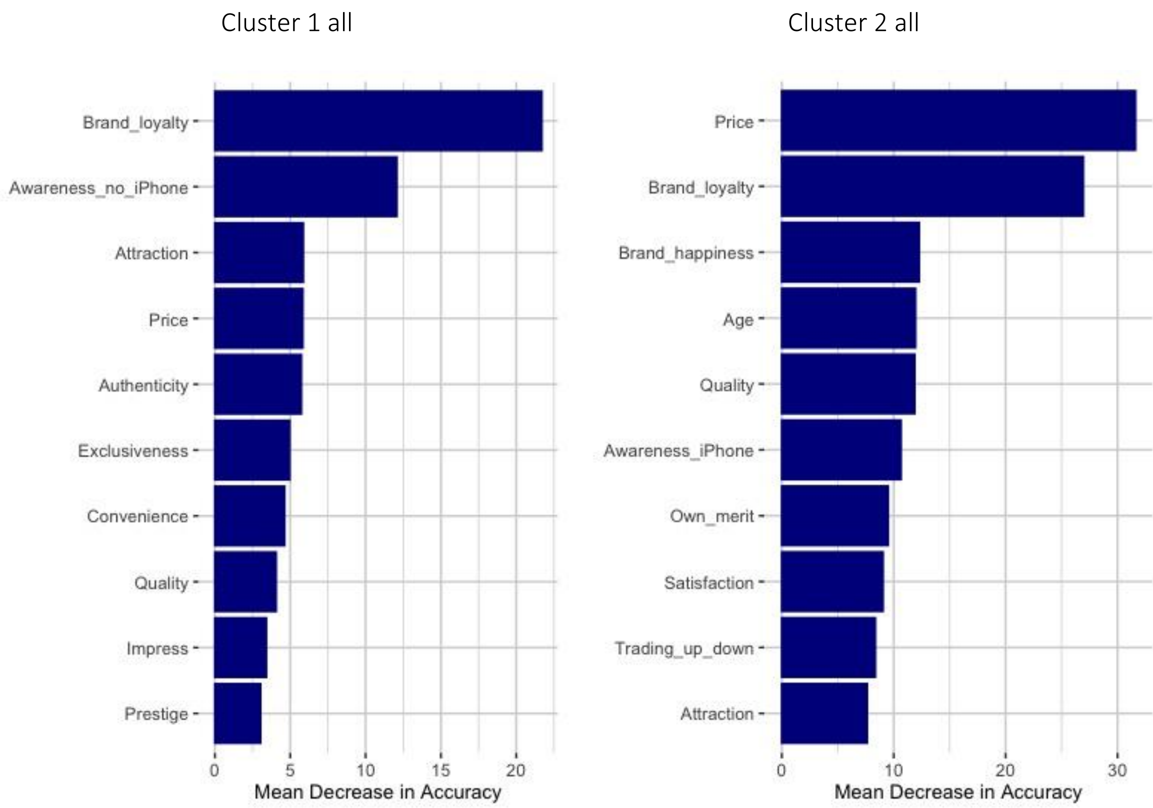


Figure 5.4: Variable importance plots derived from RF models built on the two clusters. Whereas the left plot (based on cluster 1) indicates brand loyalty and no-iPhone-awareness as most important in predicting iPhone ownership, the right plot (based on cluster 2) indicates price and brand loyalty as most important ones. These insights differ from the global analysis.

Further, the data set is clustered based on only the demographic variables gender (male/female), age, and income with the help of an unsupervised RF. In a two-cluster solution, the data is perfectly separated by gender. The according heatmap can be derived from Figure 5.5.

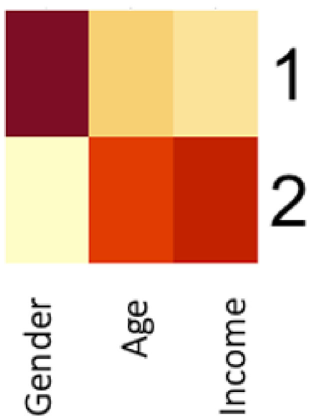


Figure 5.5: Heatmap derived from two-cluster solution based on demographic inputs, indicating a separation by gender with smaller differences in age and income

After separating the subgroups from each other, a RF model is built on each. While the size of the clusters as well as the exact OOB-errors can be derived from Table 5.4, it can be summarized that the models did not lead to a noteworthy improvement in prediction compared to the global analysis. Nevertheless, the variable importance of the models is inspected and can be derived from Figure 5.6. Both RF models indicate brand loyalty and price as the most important variables in distinguishing iPhone owners from other mobile phone owners, what is in line with the global analysis. Nevertheless, it is striking that the variables acceptance, power, and superiority appear as more important within the female cluster compared to the male cluster and compared to the global analysis. Whereas brand happiness plays a more important role in the male cluster, it does not appear important within the female cluster which means that it is not crucial for distinguishing female iPhone owners from others. Overall, slight differences are existent between the female and male clusters, but the top two variables remain the same.

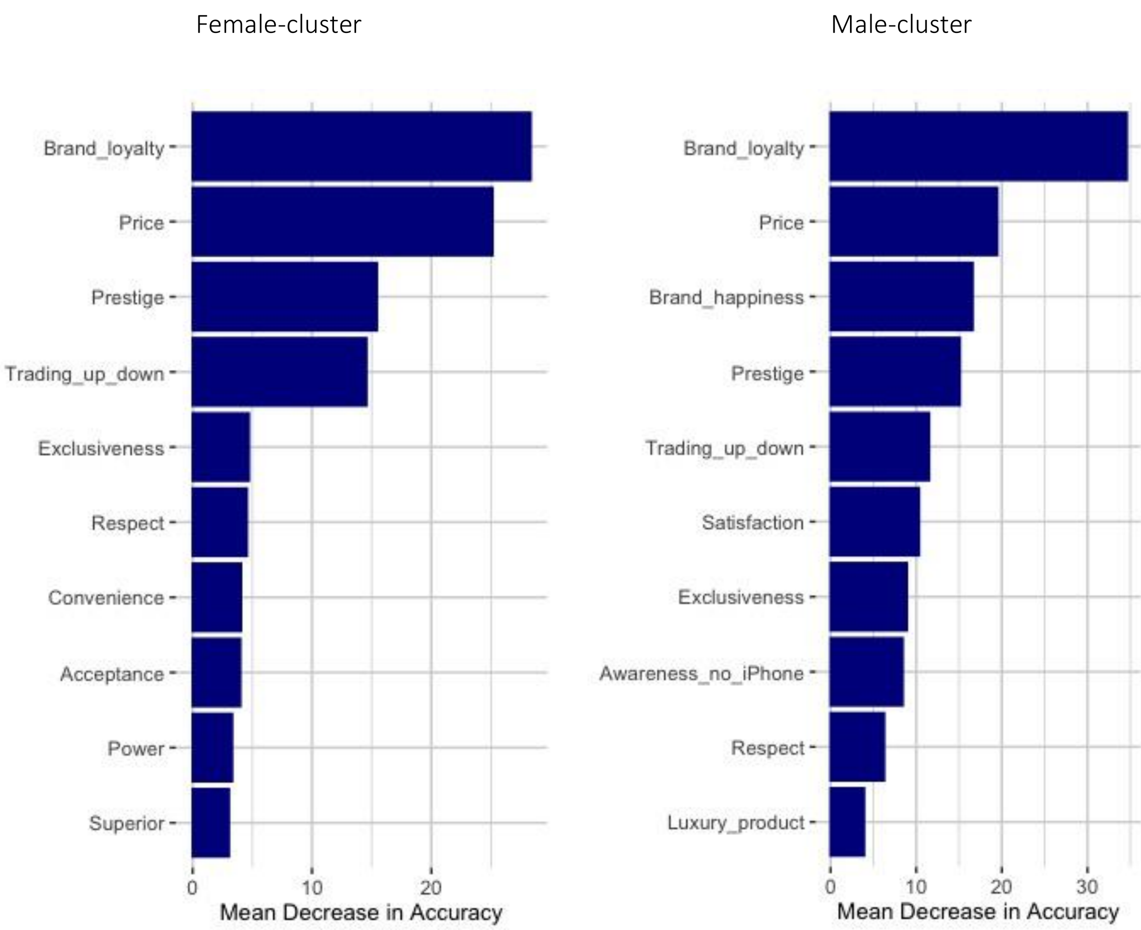


Figure 5.6: Variable importance plots derived from the clusters based on demographic variables. Both indicating brand loyalty and price as most important features for prediction. The left plot belongs to the cluster including females, the right one includes males.

Besides splitting the data into smaller subgroups of observations, it is also possible to investigate results based on only a subset of input variables. Therefore, a model was built based on those variables, occurring in the top five of all models above. That is a totality of eleven independent variables including brand loyalty, price, prestige, trading up/down, brand happiness, awareness of non-iPhone owners, exclusiveness, attraction, quality, authenticity, and age. Since all observations can be used for this approach, the additional split into training and test set can be made, hence a SL can be trained for prediction and analysis. This leads to an accuracy of 95.24% in predicting the test set. Therefore, the performance is equal to the SL model in the global analysis. The derived variable importance plot can be seen in Figure 5.7, and the derived coefficients in Table 5.5. It indicates the features capturing brand loyalty and price as most important, whereas higher values lead to a lower probability of owning an iPhone respectively. The variables prestige, trading up/down, and brand happiness follow, and higher values rise the probability of iPhone ownership, as well as higher values in age and quality. Since the accuracy is as high as based on all variables, this part of the analysis suggest that the majority of independent variables add noise to the model instead of providing accurate information for prediction. Therefore, the described subset is sufficient to derive a high accuracy. Nevertheless, it does not mean that all other variables are unimportant for prediction. A different allocation of variables could lead to similar results, as the variable importance of the different clusters indicate. Nevertheless, the totality of all independent variables measured are unnecessary for a prediction purpose.

Figure 5.7: Variable importance plot derived from SL model based on the top 5 variables among all models, indicating the features capturing brand loyalty and price as most important

Variable	Estimated Coefficient
Intercept	3.07
Brand_loyalty	- 0.67
Price	- 0.58
Quality	0.19
Prestige	0.17
Awareness_no_iPhone	0.09
Brand happiness	0.08
Trading_up_down	0.06
Authenticity	- 0.01
Age	0.01
Exclusiveness	0.02

Table 9.5: Coefficient estimations from SL model based on the top 5 variables among all models, indicating brand loyalty and price as most influential, where higher values lead to a lower probability of owning an iPhone

Lastly, hierarchical cluster analysis is performed on the subsample of observations that belong to iPhone owners. Although insights do not directly serve in distinguishing iPhone owners from people with alternative mobile phones, it can help deriving a broader background knowledge for interpretation and further research. Indeed, the resulting heatmap in Figure 5.8 indicates clear differences between clusters. Whereas cluster 1 (containing 80 observations) has high mean values in most of the self- and society-related variables and is very brand loyal (low cluster mean), cluster 2 (85 observations) has high mean values in the utilitarian variables quality and price, as well as in authenticity, age, and income. The two clusters are similar to the clusters derived by separating observations based on all variables. It supports the claim that there are subgroups within the data that consider different factors as important.

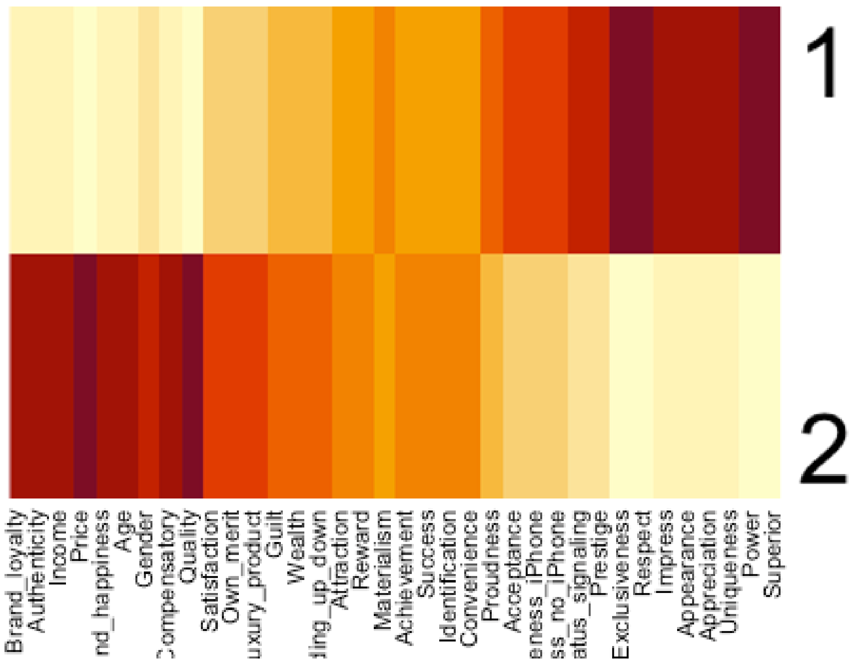


Figure 3.8: Heatmap based on observations belonging to iPhone owners, with clear differences between clusters, suggesting that there are differences within the subgroup

5.3. Summary Results

Although some of the results were expected in connection with the Apple iPhone, like very brand loyal customers, an unreasonable price, and the perception of prestige towards the product, other variables that were indicated as important by this research were less expected. These include that younger age, as well as the feeling of authenticity is closer related to alternative mobile phone brands. Remarkable is also, that the variable capturing income seems not influential to iPhone ownership at all. Although the difference between genders is relatively small, the cluster analysis based on all variables, together with the cluster analysis based on only iPhone users, showed that there are subgroups within the data with different characteristics, hence, divergent drivers.

6. Discussion

The following chapter discusses the results derived from the analysis, as well as the performance of the different methods. Additionally, the theoretical contribution is highlighted, followed by managerial implications, limitations of the study, and directions for further research. The chapter ends with a brief conclusion.

6.1. General Discussion

The previous study reveals brand loyalty, price, and perceived prestige as the most important features in distinguishing iPhone owners from non-iPhone owners. From the results, it can be deducted that iPhone users are, on average, more brand loyal, perceive the product price as unreasonably high, but the product as more prestigious. Since the Apple iPhone is used as a representative product of the masstige segment, it is assumed that the results of this study are transferable to masstige consumers in general. Therefore, this study empirically shows that consumers of masstige products are on average more brand loyal, less price sensitive, and seek for prestige in their products. These results are consistent with previous research explaining that the perceived prestige level distinguishes masstige products from other mid-range products and thus justifies a higher price (Truong et al., 2009), or that luxury values positively affect brand loyalty (Chung & Kim, 2020). Moreover, the data literally reflect the definition of mass-prestige, since iPhone owners are the majority group within the data, indicating the iPhone as prestigious.

According to prior research, brand loyalty is closely related to brand happiness, whereas the latter one positively influences the former one (e.g. Gelbrich, 2011; Purohit & Radia, 2022). Kumar et al. (2021)

found that masstige purchase intention leads to brand happiness, supported by Purohit & Radia (2022) who additionally showed a significant influence of brand happiness on customer loyalty. In the present study, brand happiness is as well among the top variables indicating masstige consumption. Further, Schnebelen & Bruhn (2018) found brand happiness positively influencing re-purchase intentions, which can also be derived from the data indicating iPhone users are more likely to repurchase the product.

Regarding utilitarian features, prior research agrees on the importance of high quality of luxury products (e.g., Kapferer, 1997; Kroeber-Riel & Gröppel-Klein, 2019). Kapferer et al. (2014) argue that it also drives masstige consumption, what was additionally observed by Kumar et al. (2020), and empirically proven in the study of Purohit & Radia (2022). Nevertheless, although quality is not entirely unimportant in this study, it is not indicated as a particularly important characteristic in determining masstige consumption. Hence, the willingness to pay an unreasonably high price in this case is not rooted in perceptions of quality. Mandel et al. (2017) argue that consumer goods usually have an emotional and psychological value that is detached from their utilitarian value. This insight becomes more important the more extensive the decision for a product is (Kroeber-Riel & Gröppel-Klein, 2019). Since buying a mobile phone is clearly not a habitual decision, it requires a higher level of cognition that includes a higher emotional level (Kroeber-Riel & Gröppel-Klein, 2019). This can be clearly seen from the data, where masstige customers place more focus on society-related factors, which are detached from utilitarian value. It is particularly notable that none of the participants owning an alternative mobile phone indicated any society-related factor as very important. However, the global analysis does not indicate the society-related factors as particularly crucial in accurately distinguish consumers.

According to previous research, exclusiveness and uniqueness play an important role in the luxury consumption segment (e.g., Belk, 1985; Gentina et al., 2016). However, the study by Purohit & Radia (2022) suggests that these factors diminish in the masstige segment. Although the global analysis supports this finding, the part of the local analysis based only on the subsample of iPhone owners shows a clear contrast. It indicates a group of participants, mainly younger men, precisely ranking such attributes as exclusiveness, uniqueness, power, superiority, and respect as more important than the average masstige consumer. Thus, it cannot be confirmed that these factors completely diminish in the masstige segment. Purohit & Radia (2022) further found a positive relationship between vanity values (e.g., enhanced appearance, and signaling status or success) and masstige consumption, which is also prevalent in the luxury sector (Belk, 1985), but was not observed in the global analysis in this work. These mixed results mirror previous research on luxury consumption by Soh et al. (2017), who found that vanity values and uniqueness influence purchase intention but do not explain purchase behavior.

Furthermore, this research empirically supports Silverstein & Fiske's suggestion (2003) that masstige consumption is not dependent on the income level compared to luxury consumption (Han et al., 2010). Silverstein & Fiske's theory (2003) that masstige consumers trade some product categories up and trade down other product categories is supported by the present study, which reveals that individuals who report the representative product category as important are more likely to own a masstige product. This feature is among the most important consumer differentiators in this study. Since consumption is therefore rather a matter of trade-off than a matter of income, the additional finding that masstige consume is not perceived as a reward is consistent with this. Thus, while reward plays an important role in the luxury segment (Kroeber-Riel & Gröppel-Klein, 2019), it disappears completely in the masstige segment.

In the present work, it was found that acceptance and appreciation are not crucial in the masstige segment, so consumers do not seek social approval. These results match previous research findings where social recognition is neither found to be significant in the field of luxury (Shukla, 2012), nor in the field of masstige (Purohit & Radia, 2022). Nevertheless, the fact that consumers notice when others do not own an iPhone, plays a major role. It could be inferred that people do not consider masstige products as social approval, but rather as a commodity. This is supported by the data, according to which the majority of participants do not regard the masstige product as a luxury.

Purohit & Radia's study (2022) suggests that women place more value on social recognition and men place more value on non-functional attributes but indicate all other variables as not gender specific. The present study indicates the factors acceptance, power, and superiority as more important to women, whereas men put more emphasis on brand happiness in contrast. Further, the analysis indicates that there are no differences in the crucial variables. Therefore, it is assumed that gender does not have a strong influence on masstige consumption, which is consistent with the study of Purohit & Radia (2022).

According to the results of this study, general attitudes toward wealth, success, or achievement do not play a role in distinguishing masstige consumers, and the concept of compensatory consumption is also not observed among masstige consumption, although it does play a role in the luxury segment (Kroeber-Riel & Gröppel-Klein, 2019).

In summary, the results also confirm that cultural differences in value perception within a segment are also prevalent in the masstige sector, as previous research has already found in the luxury context (Gentina et al., 2016; Zhang & Zhao, 2019).

6.2. Methodological Discussion

The performance of the different models was quite surprising. The predictions of the LR model were unexpectedly accurate, while the RF model performed worse. These results are not in line with previous research which clearly indicate tree-based methods and RF as particularly better performing (e.g., De Caigny et al., 2018; Lemmens & Croux, 2006). Nevertheless, the data set does not formally match the assumptions on which the LR method is based. Therefore, the method can be used for prediction, but conclusions from the results should be drawn with caution as it can lead to restricted or even wrong interpretations (Sarstedt & Mooi, 2019). Additionally, the trade-off between taking effort on investigating and adjusting the data according to the assumptions and deriving additional insights must be considered. Since LR is often used due to its good interpretability, it is a major drawback when inferences are not reliable due to unsuitable data. Hence, the method of RF is more convenient especially for inexperienced marketers. Nevertheless, its reliability is reduced for small data sets. The prediction results based on cluster analysis brought further thought-provoking insights. Therefore, the method is recommended if there is an interest in analyzing possible subgroups.

Since the combination method of SL yields the best predictive performance for the dataset and both main insights from LR and RF can be derived, it is highly recommended for marketing research. This confirms previous findings showing that SL performs best compared to other ML methods (Džeroski & Ženko, 2004; Lessmann et al., 2015). In addition, some of the variables were already classified as non-influential in the first SL model (by estimating their coefficients to zero). All eleven variables independently derived from the top 5 of all models were already estimated to be non-zero in the global SL analysis, thus were considered influential already at the beginning. Since only 15 predictors were estimated to be non-zero, the model already gave a very good indication of the following results. It suggests that the majority of independent variables add noise to the model instead of providing accurate information for prediction, which is supported by the results of the last model where a subset of variables leads to the same prediction performance.

In summary, this work shows that each method has its own advantages and disadvantages regarding, complexity, insights, and performance, but proves that SL is a powerful method for marketing research that largely combines the insights and advantages of LR and RF. Although it also combines their disadvantages, it remedies those that can be remedied by complementary application.

6.3. Theoretical Contribution

Although prior work exists on masstige consumer behavior (e.g., Paul, 2015, 2018; Kumar et al., 2020; Purohit & Radia, 2022), research is still in its infancy phase and researchers called for a deeper understanding of the segment (e.g., Paul, 2018; Kumar et al., 2020), as well as an external validity check (Purohit & Radia, 2022). This thesis contributes to the theory of masstige consumption by empirically examining additional factors (e.g., general attitudes) in relation to masstige products and better differentiating the masstige segment from the luxury segment. Additionally, the study focusses only on the Western European area to verify previous findings across borders (e.g., Purohit & Radia, 2022; Truong et al., 2009).

By using a machine learning approach for analysis, this research also contributes to the small body of research that combines machine learning methods with the field of marketing research (e.g., Lemmens & Croux, 2006; Lessmann et al., 2015). It identifies the super learner as the most powerful and useful method for predictions as well as for deriving consumption drivers, although there are some drawbacks to consider. Moreover, it is shown that even non-parametric machine learning approaches can lead to important and meaningful insights, while their application is much simpler than traditional approaches. Therefore, this thesis contributes to the field of machine learning by showing that it is an appropriate approach for marketing research in general, but also that the super learner in particular excels in this area.

6.4. Managerial Implications

This study provides several contributions that are useful for marketing managers to design marketing activities within the masstige domain efficiently. The results clearly show that utilitarian values are not of great interest in the masstige sector. Therefore, brand managers should not emphasize features such as quality and price, as they are not crucial for the purchase of masstige products. In addition, the promotional strategy should be built on an emotional level that goes beyond functional value (Mandel et al., 2017). In the case of masstige consumption in the Western European market, brand loyalty and perceived prestige are the main drivers. Therefore, these two aspects will be considered in more detail in the following part.

Based on brand loyalty, five main concepts have been derived from prior research (Khamitov et al., 2019): brand attachment (Park et al., 2010), brand love (Batra et al., 2012), self-brand connection (Escalas & Bettman, 2003), brand identification (Stokburger-Sauer et al., 2012), and brand trust (Chaudhuri & Holbrook, 2001). Based on several studies (e.g., Homburg et al., 2009; Mazodier &

Merunka, 2012), these aspects are positive predictors of customer brand loyalty, and their positive promotion should therefore be the focus of a masstige marketing strategy. In addition, consumption of a masstige product should be publicized to transform a utilitarian product into a lifestyle product that can be used for self-expression (Jensen Schau & Gilly, 2003). This, in turn, promotes the self-brand connection as it increases with publicly consumed brands (Khamitov et al., 2019). Complaint management also plays an important role in increasing brand loyalty (Morgeson et al., 2020), as loyalty is built on any interaction with the brand when it meets the expectations of its customers (Court et al., 2009). In addition, consumers are more likely to develop brand loyalty when the brand is perceived as prestigious (Choi et al., 2017).

To improve perceptions of prestige, brand endorsement can be effective (Chen & Wyer, 2020). However, the choice of the brand endorser is critical to the success or failure of prestige endorsement. This statement is based on the research of Chen & Wyer (2020), who found that the endorser's social prestige level is transferred to the product when the actual prestige value of the product is unknown. Therefore, an endorser who is considered to have a high status should be selected, as they are expected to use prestigious products (Chen & Wyer, 2020). Nevertheless, the endorsement must be authentic, so that consumers assume that the endorser actually uses the product themselves (Kroeber-Riel & Gröppel-Klein, 2019). While Chen & Wyer (2020) additionally found that smiling male endorsers and non-smiling female endorsers convey prestige value better than their counterpart, van Kleef et al. (2021) found in the context of leadership endorsement that risk-taking protagonists are perceived as more prestigious than those who avoid risk.

6.5. Limitations & Further Research

This study was conducted to develop a better theory of the masstige segment and to derive the drivers of masstige consumption. Also, an attempt was made to combine machine learning with the field of marketing research. Therefore, all possible efforts were made to obtain robust results. Nevertheless, this study has its limitations. First, as in previous research, the Apple iPhone was examined in this study. However, in order to make broader generalizations, it is necessary to examine different products and brands from different industries. In addition, many participants indicated the iPhone as a luxury product, which could be due to the fact that there is no established luxury brand in the mobile phone market. This assumption should be further investigated to ensure that the current and future results are indeed related to the masstige sector. In addition, this study attempts to find differences between masstige consumers and non-masstige consumers. However, some results suggest that there are differences between consumers even within the masstige segment. Therefore, future studies could explore differences among masstige consumers to provide a better understanding. As it was suggested that masstige products may not be considered higher value products but have become the new average products. Future research could include studies in this direction to gain a deeper understanding of the masstige segment. In terms of methodology, RF's poor performance was very surprising. Therefore, it would be interesting to see where such a difference to previous research comes from, whereby a cross-industry comparison could be a first step.

6.6. Conclusion

This research was conducted to gain deeper insights into the masstige segment and additionally contributing to the sparse body of literature towards machine learning in marketing research. To this end, a study was conducted among Western Europeans in which various factors were measured in relation to a masstige product. Particular attention was paid to the drivers of luxury consumption, as the masstige sector is considered as new luxury, as well as to previous findings from a study conducted in India. Crucial differences to alternative products were found in brand loyalty, the judgement of price, and perceived prestige value, whereas functional values such as quality were completely eclipsed. In addition, cultural differences in the masstige segment and a clear distinction from the luxury segment were found. The results are helpful for marketing managers working in the masstige domain to develop or improve their marketing strategy. In addition, the analysis finds that machine learning methods, and super learner in particular, are applicable to marketing-related questions. This insight expands the possibilities of analytical approaches for marketing researchers.

References

Aggarwal, C. C. (Ed.). (2014). *Data classification: Algorithms and applications*. CRC Press, Taylor & Francis Group.

Allport, G. W. (1937). *Personality: A psychological interpretation*. Henry Holt & Co.

Amatriain, X., & Basilico, J. (2012, April 6). Netflix Recommendations: Beyond the 5 stars (Part 1). *Netflix TechBlog*. https://netflixtechblog.com/netflix-recommendations-beyond-the-5-stars-part-1-55838468f429

Anderson, C., Hildreth, J. A. D., & Howland, L. (2015). Is the desire for status a fundamental human motive? A review of the empirical literature. *Psychological Bulletin*, *141*(3), 574–601. https://doi.org/10.1037/a0038781

Barbopoulos, I., & Johansson, L.-O. (2017). The Consumer Motivation Scale: Development of a multi-dimensional and context-sensitive measure of consumption goals. *Journal of Business Research*, *76*, 118–126. https://doi.org/10.1016/j.jbusres.2017.03.012

Barocas, S., & Selbst, A. D. (2016). Big Data's Disparate Impact. *California Law Review*, *104*, 671–732. https://doi.org/10.15779/Z38BG31

Batra, R., Ahuvia, A., & Bagozzi, R. P. (2012). Brand Love. *Journal of Marketing*, *76*(2), 1–16. https://doi.org/10.1509/jm.09.0339

Belk, R. W. (1985). Materialism: Trait Aspects of Living in the Material World. *Journal of Consumer Research*, *12*(3), 265. https://doi.org/10.1086/208515

Bellezza, S., & Keinan, A. (2014). Brand Tourists: How Non–Core Users Enhance the Brand Image by Eliciting Pride. *Journal of Consumer Research*, *41*(2), 397–417. https://doi.org/10.1086/676679

Breiman, L. (1996). Bagging predictors. *Machine Learning*, *24*(2), 123–140. https://doi.org/10.1007/BF00058655

Breiman, L. (2001). Random Forests. *Machine Learning*, *45*(1), 5–32. https://doi.org/10.1023/A:1010933404324

Breiman, L., & Cutler, A. (n.d.). *Random Forests*. https://www.stat.berkeley.edu/~breiman/RandomForests/

Breiman, L., Friedman, J. H., Olshen, R. A., & Stone, C. J. (1984). *Classification And Regression Trees* (1st ed.). Routledge. https://doi.org/10.1201/9781315139470

Chapman, C., & Feit, E. M. (2019). *R For Marketing Research and Analytics*. Springer International Publishing. https://doi.org/10.1007/978-3-030-14316-9

Chaudhuri, A., & Holbrook, M. B. (2001). The Chain of Effects from Brand Trust and Brand Affect to Brand Performance: The Role of Brand Loyalty. *Journal of Marketing*, *65*(2), 81–93. https://doi.org/10.1509/jmkg.65.2.81.18255

Chen, Y., & Wyer, R. S. (2020). The effects of endorsers' facial expressions on status perceptions and purchase intentions. *International Journal of Research in Marketing*, *37*(2), 371–385. https://doi.org/10.1016/j.ijresmar.2019.10.002

Choi, Y. G., Ok, C. "Michael," & Hyun, S. S. (2017). Relationships between brand experiences, personality traits, prestige, relationship quality, and loyalty: An empirical analysis of coffeehouse brands. *International Journal of Contemporary Hospitality Management*, *29*(4), 1185–1202. https://doi.org/10.1108/IJCHM-11-2014-0601

Chung, Y., & Kim, A. J. (2020). Effects of mergers and acquisitions on brand loyalty in luxury Brands: The moderating roles of luxury tier difference and social media. *Journal of Business Research*, *120*, 434–442. https://doi.org/10.1016/j.jbusres.2019.11.030

Court, D., Elzinga, D., Mulder, S., & Vetvik, O. J. (2009). The consumer decision journey. *McKinsey Quarterly*, *3*, 96–107. Business Source Premier.

Currim, I. S., Meyer, R. J., & Le, N. T. (1988). Disaggregate Tree-Structured Modeling of Consumer Choice Data. *Journal of Marketing Research*, *25*(3), 253. https://doi.org/10.2307/3172528

Das, M., Saha, V., & Roy, A. (2021). Inspired and engaged: Decoding MASSTIGE value in engagement. *International Journal of Consumer Studies*, ijcs.12726. https://doi.org/10.1111/ijcs.12726

De Botton, A. (2004). *Status anxiety* (1. publ). Hamish Hamilton.

De Caigny, A., Coussement, K., & De Bock, K. W. (2018). A new hybrid classification algorithm for customer churn prediction based on logistic regression and decision trees. *European Journal of Operational Research*, *269*(2), 760–772. https://doi.org/10.1016/j.ejor.2018.02.009

Dietterich, T. G. (2000). An Experimental Comparison of Three Methods for Constructing Ensembles of Decision Trees: Bagging, Boosting, and Randomization. *Machine Learning*, *40*(2), 139–157. https://doi.org/10.1023/A:1007607513941

Drèze, X., & Nunes, J. C. (2009). Feeling Superior: The Impact of Loyalty Program Structure on Consumers' Perceptions of Status. *Journal of Consumer Research*, *35*(6), 890–905. https://doi.org/10.1086/593946

Dubois, B., Czellar, S., & Laurent, G. (2005). Consumer Segments Based on Attitudes Toward Luxury: Empirical Evidence from Twenty Countries. *Marketing Letters*, *16*(2), 115–128. https://doi.org/10.1007/s11002-005-2172-0

Dubois, D., Jung, S., & Ordabayeva, N. (2021). The psychology of luxury consumption. *Current Opinion in Psychology*, *39*, 82–87. https://doi.org/10.1016/j.copsyc.2020.07.011

Džeroski, S., & Ženko, B. (2004). Is Combining Classifiers with Stacking Better than Selecting the Best One? *Machine Learning*, *54*(3), 255–273. https://doi.org/10.1023/B:MACH.0000015881.36452.6e

Escalas, J. E., & Bettman, J. R. (2003). You Are What They Eat: The Influence of Reference Groups on Consumers' Connections to Brands. *Journal of Consumer Psychology*, *13*(3), 339–348. https://doi.org/10.1207/S15327663JCP1303_14

Gao, L., Wheeler, S. C., & Shiv, B. (2009). The "Shaken Self": Product Choices as a Means of Restoring Self-View Confidence. *Journal of Consumer Research*, *36*(1), 29–38. https://doi.org/10.1086/596028

Gelbrich, K. (2011). I Have Paid Less Than You! The Emotional and Behavioral Consequences of Advantaged Price Inequality. *Journal of Retailing*, *87*(2), 207–224. https://doi.org/10.1016/j.jretai.2011.03.003

Gentina, E., Shrum, L. J., & Lowrey, T. M. (2016). Teen attitudes toward luxury fashion brands from a social identity perspective: A cross-cultural study of French and U.S. teenagers. *Journal of Business Research*, *69*(12), 5785–5792. https://doi.org/10.1016/j.jbusres.2016.04.175

Goor, D., Ordabayeva, N., Keinan, A., & Crener, S. (2020). The Impostor Syndrome from Luxury Consumption. *Journal of Consumer Research*, *46*(6), 1031–1051. https://doi.org/10.1093/jcr/ucz044

Hagtvedt, H., & Patrick, V. M. (2009). The broad embrace of luxury: Hedonic potential as a driver of brand extendibility. *Journal of Consumer Psychology*, *19*(4), 608–618. https://doi.org/10.1016/j.jcps.2009.05.007

Han, Y. J., Nunes, J. C., & Drèze, X. (2010). Signaling Status with Luxury Goods: The Role of Brand Prominence. *Journal of Marketing*, *74*(4), 15–30. https://doi.org/10.1509/jmkg.74.4.015

Heine, K., & Phan, M. (2011). Trading-Up Mass-Market Goods to Luxury Products. *Australasian Marketing Journal*, *19*(2), 108–114. https://doi.org/10.1016/j.ausmj.2011.03.001

Homburg, C. (2017). *Marketingmanagement: Strategie - Instrumente - Umsetzung - Unternehmensführung* (6. Auflage). Springer Gabler.

Homburg, C., Wieseke, J., & Hoyer, W. D. (2009). Social Identity and the Service-Profit Chain. *Journal of Marketing*, *73*(2), 38–54. https://doi.org/10.1509/jmkg.73.2.38

Hoyer, W. D., MacInnis, D. J., & Pieters, R. (2013). *Consumer behavior* (6th edition). South Western Cengage Learning.

Hung, K., Huiling Chen, A., Peng, N., Hackley, C., Amy Tiwsakul, R., & Chou, C. (2011). Antecedents of luxury brand purchase intention. *Journal of Product & Brand Management*, *20*(6), 457–467. https://doi.org/10.1108/10610421111166603

Husic, M., & Cicic, M. (2009). Luxury consumption factors. *Journal of Fashion Marketing and Management: An International Journal*, *13*(2), 231–245. https://doi.org/10.1108/13612020910957734

James, G., Witten, D., Hastie, T., & Tibshirani, R. (2021). *An introduction to statistical learning: With applications in R* (Second edition). Springer. https://doi.org/10.1007/978-1-0716-1418-1

Jensen Schau, H., & Gilly, M. C. (2003). We Are What We Post? Self-Presentation in Personal Web Space. *Journal of Consumer Research*, *30*(3), 385–404. https://doi.org/10.1086/378616

Kapferer, J.-N. (1997). Managing luxury brands. *Journal of Brand Management*, *4*(4), 251–259. https://doi.org/10.1057/bm.1997.4

Kapferer, J.-N., Klippert, C., & Leproux, L. (2014). Does luxury have a minimum price? An exploratory study into consumers' psychology of luxury prices. *Journal of Revenue and Pricing Management*, *13*(1), 2–11. https://doi.org/10.1057/rpm.2013.34

Kappes, H. B., Gladstone, J. J., & Hershfield, H. E. (2021). Beliefs about Whether Spending Implies Wealth. *Journal of Consumer Research*, *48*(1), 1–21. https://doi.org/10.1093/jcr/ucaa060

Khamitov, M., Wang, X. (Shane), & Thomson, M. (2019). How Well Do Consumer-Brand Relationships Drive Customer Brand Loyalty? Generalizations from a Meta-Analysis of Brand Relationship Elasticities. *Journal of Consumer Research*, *46*(3), 435–459. https://doi.org/10.1093/jcr/ucz006

Kivetz, R., & Simonson, I. (2002). Earning the Right to Indulge: Effort as a Determinant of Customer Preferences toward Frequency Program Rewards. *Journal of Marketing Research*, *39*(2), 155–170. https://doi.org/10.1509/jmkr.39.2.155.19084

Kroeber-Riel, W., & Gröppel-Klein, A. (2019). *Konsumentenverhalten* (11. Auflage). Verlag Franz Vahlen.

Kumar, A., Kabra, G., Mussada, E. K., Dash, M. K., & Rana, P. S. (2019). Combined artificial bee colony algorithm and machine learning techniques for prediction of online consumer repurchase intention. *Neural Computing and Applications*, *31*(S2), 877–890. https://doi.org/10.1007/s00521-017-3047-z

Kumar, A., Paul, J., & Starčević, S. (2021). Do brands make consumers happy?- A masstige theory perspective. *Journal of Retailing and Consumer Services*, *58*, 102318. https://doi.org/10.1016/j.jretconser.2020.102318

Kumar, A., Paul, J., & Unnithan, A. B. (2020). 'Masstige' marketing: A review, synthesis and research agenda. *Journal of Business Research*, *113*, 384–398. https://doi.org/10.1016/j.jbusres.2019.09.030

Lay, A., & Furnham, A. (2019). A New Money Attitudes Questionnaire. *European Journal of Psychological Assessment*, *35*(6), 813–822. https://doi.org/10.1027/1015-5759/a000474

LeDell, E., Gill, N., Aiello, S., Fu, A., & Candel, A. (2022). *H2o: R Interface for the "H2O" Scalable Machine Learning Platform. R package version 3.36.1.2*. https://CRAN.R-project.org/package=h2o

Lee, S., Baumgartner, H., & Winterich, K. P. (2018). Did They Earn It? Observing Unearned Luxury Consumption Decreases Brand Attitude When Observers Value Fairness. *Journal of Consumer Psychology*, *28*(3), 412–436. https://doi.org/10.1002/jcpy.1028

Lemmens, A., & Croux, C. (2006). Bagging and Boosting Classification Trees to Predict Churn. *Journal of Marketing Research*, *43*(2), 276–286. https://doi.org/10.1509/jmkr.43.2.276

Lessmann, S., Baesens, B., Seow, H.-V., & Thomas, L. C. (2015). Benchmarking state-of-the-art classification algorithms for credit scoring: An update of research. *European Journal of Operational Research*, *247*(1), 124–136. https://doi.org/10.1016/j.ejor.2015.05.030

Liaw, A., & Wiener, M. (2002). *Classification and Regression by randomForest* (2(3); R News, pp. 18–22). https://CRAN.R-project.org/doc/Rnews/

Lilhore, U. K., Simaiya, S., Prasad, D., & Verma, D. K. (2021). Hybrid Weighted Random Forests Method for Prediction & Classification of Online Buying Customers. *Journal of Information Technology Management*, *13*(2). https://doi.org/10.22059/jitm.2021.310062.2607

Mandel, N., Rucker, D. D., Levav, J., & Galinsky, A. D. (2017). The Compensatory Consumer Behavior Model: How self-discrepancies drive consumer behavior. *Journal of Consumer Psychology*, *27*(1), 133–146. https://doi.org/10.1016/j.jcps.2016.05.003

Maslow, A. H. (1943). A theory of human motivation. *Psychological Review*, *50*(4), 370–396. https://doi.org/10.1037/h0054346

Mazodier, M., & Merunka, D. (2012). Achieving brand loyalty through sponsorship: The role of fit and self-congruity. *Journal of the Academy of Marketing Science*, *40*(6), 807–820. https://doi.org/10.1007/s11747-011-0285-y

McFerran, B., Aquino, K., & Tracy, J. L. (2014). Evidence for two facets of pride in consumption: Findings from luxury brands. *Journal of Consumer Psychology*, *24*(4), 455–471. https://doi.org/10.1016/j.jcps.2014.03.004

McFerran, B., & Argo, J. J. (2014). The Entourage Effect. *Journal of Consumer Research*, *40*(5), 871–884. https://doi.org/10.1086/673262

Miguéis, V. L., Camanho, A. S., & Borges, J. (2017). Predicting direct marketing response in banking: Comparison of class imbalance methods. *Service Business*, *11*(4), 831–849. https://doi.org/10.1007/s11628-016-0332-3

Morgeson, F. V., Hult, G. T. M., Mithas, S., Keiningham, T., & Fornell, C. (2020). Turning Complaining Customers into Loyal Customers: Moderators of the Complaint Handling–Customer Loyalty Relationship. *Journal of Marketing*, *84*(5), 79–99. https://doi.org/10.1177/0022242920929029

Morhart, F., Wilcox, K., & Czellar, S. (Eds.). (2020). Luxury and environmental responsibility. In *Research Handbook on Luxury Branding*. Edward Elgar Publishing. https://doi.org/10.4337/9781786436351.00031

Mummendey, H. D. (1997). *Die Ausführung unterschiedlicher Arten selbstschädigenden Verhaltens zum Zwecke der Eindruckssteuerung*. Universität Bielefeld. https://pub.uni-bielefeld.de/record/2466638#apa

Naimi, A. I., & Balzer, L. B. (2018). Stacked generalization: An introduction to super learning. *European Journal of Epidemiology*, *33*(5), 459–464. https://doi.org/10.1007/s10654-018-0390-z

OED. (n.d.). *Oxford English Dictionary. "machine, n".* Oxford University Press. https://www.oed.com/view/Entry/111850

Park, C. W., Macinnis, D. J., Priester, J., Eisingerich, A. B., & Iacobucci, D. (2010). Brand Attachment and Brand Attitude Strength: Conceptual and Empirical Differentiation of Two Critical Brand Equity Drivers. *Journal of Marketing*, *74*(6), 1–17. https://doi.org/10.1509/jmkg.74.6.1

Paul, J. (2015). Masstige Marketing Redefined and Mapped: Introducing a Pyramid Model and an MMS Measure. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.2677495

Paul, J. (2018). Toward a masstige theory and strategy for marketing. *European J. of International Management*, *1*(1), 1. https://doi.org/10.1504/EJIM.2018.10012543

Porter, M. E. (1980). *Competitive strategy: Techniques for analyzing industries and competitors*. Free Press.

Purohit, S., & Radia, K. N. (2022). Conceptualizing masstige buying behavior: A mixed-method approach. *Journal of Business Research*, *142*, 886–898. https://doi.org/10.1016/j.jbusres.2022.01.023

R Core Team. (2021). *R: A language and environment for statistical computing. R Foundation for Statistical Computing*. https://www.R-project.org/

Richins, M. L., & Dawson, S. (1992). A Consumer Values Orientation for Materialism and Its Measurement: Scale Development and Validation. *Journal of Consumer Research*, *19*(3), 303. https://doi.org/10.1086/209304

Rucker, D. D., & Galinsky, A. D. (2008). Desire to Acquire: Powerlessness and Compensatory Consumption. *Journal of Consumer Research*, *35*(2), 257–267. https://doi.org/10.1086/588569

Sarstedt, M., & Mooi, E. (2019). *A Concise Guide to Market Research: The Process, Data, and Methods Using IBM SPSS Statistics*. Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-662-56707-4

Schnebelen, S., & Bruhn, M. (2018). An appraisal framework of the determinants and consequences of brand happiness. *Psychology & Marketing*, *35*(2), 101–119. https://doi.org/10.1002/mar.21073

Shukla, P. (2012). The influence of value perceptions on luxury purchase intentions in developed and emerging markets. *International Marketing Review*, *29*(6), 574–596. https://doi.org/10.1108/02651331211277955

Silverstein, M. J., & Fiske, N. (2003). Luxury for the masses. *Harvard Business Review*, *81*(4), 48–57. WorldCat.org.

Soh, C. Q. Y., Rezaei, S., & Gu, M.-L. (2017). A structural model of the antecedents and consequences of Generation Y luxury fashion goods purchase decisions. *Young Consumers*, *18*(2), 180–204. https://doi.org/10.1108/YC-12-2016-00654

Solomon, M. R. (2013). *Consumer Behavior: Buying, Having, and Being* (10th edition). Pearson Education Limited.

Statista. (2022a). *LVMH: Umsatz von LVMH weltweit in den Jahren 2006 bis 2021*. https://de.statista.com/statistik/daten/studie/218059/umfrage/gesamtumsatz-der-lvmh-gruppe/

Statista. (2022b). *Luxusgüter: Umsatz mit Luxusgütern weltweit in den Jahren 1996 bis 2021*. https://de.statista.com/statistik/daten/studie/76801/umfrage/umsatz-der-luxusgueterindustrie-weltweit-seit-1995/

Statista. (2022c). *Hermès: Umsatz von Hermès weltweit in den Jahren 2004 bis 2021*. https://de.statista.com/statistik/daten/studie/262781/umfrage/umsatz-von-hermes-weltweit/

Stokburger-Sauer, N., Ratneshwar, S., & Sen, S. (2012). Drivers of consumer–brand identification. *International Journal of Research in Marketing*, *29*(4), 406–418. https://doi.org/10.1016/j.ijresmar.2012.06.001

Sun, G., D'Alessandro, S., & Johnson, L. (2014). Traditional culture, political ideologies, materialism and luxury consumption in China: Traditional culture, ideology in China, luxuries. *International Journal of Consumer Studies*, *38*(6), 578–585. https://doi.org/10.1111/ijcs.12117

Tin Kam Ho. (1998). The random subspace method for constructing decision forests. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *20*(8), 832–844. https://doi.org/10.1109/34.709601

Truong, Y., McColl, R., & Kitchen, P. J. (2009). New luxury brand positioning and the emergence of Masstige brands. *Journal of Brand Management*, *16*(5–6), 375–382. https://doi.org/10.1057/bm.2009.1

Truong, Y., McColl, R., & Kitchen, P. J. (2010). Uncovering the relationships between aspirations and luxury brand preference. *Journal of Product & Brand Management*, *19*(5), 346–355. https://doi.org/10.1108/10610421011068586

United Nations Statistics Division. (n. d.). *Standard Country or Area Codes for Statistical Use*. https://unstats.un.org/unsd/methodology/m49/

van der Laan, M. J., Polley, E. C., & Hubbard, A. E. (2007). Super Learner. *Statistical Applications in Genetics and Molecular Biology*, *6*(1). https://doi.org/10.2202/1544-6115.1309

van Kleef, G. A., Heerdink, M. W., Cheshin, A., Stamkou, E., Wanders, F., Koning, L. F., Fang, X., & Georgeac, O. A. M. (2021). No guts, no glory? How risk-taking shapes dominance, prestige, and leadership endorsement. *Journal of Applied Psychology*, *106*(11), 1673–1694. https://doi.org/10.1037/apl0000868

Varmuza, K., Ping, H., & Fang, K.-T. (2003). Boosting applied to classification of mass spectral data. *Journal of Data Science*, *1*, 391–404.

Viaene, S., Derrig, R. A., Baesens, B., & Dedene, G. (2002). A Comparison of State-of-the-Art Classification Techniques for Expert Automobile Insurance Claim Fraud Detection. *Journal of Risk & Insurance*, *69*(3), 373–421. https://doi.org/10.1111/1539-6975.00023

Wickham, H. (2016). *ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag New York.

Wiedmann, K.-P., Hennigs, N., & Siebels, A. (2009). Value-based segmentation of luxury consumption behavior. *Psychology and Marketing*, *26*(7), 625–651. https://doi.org/10.1002/mar.20292

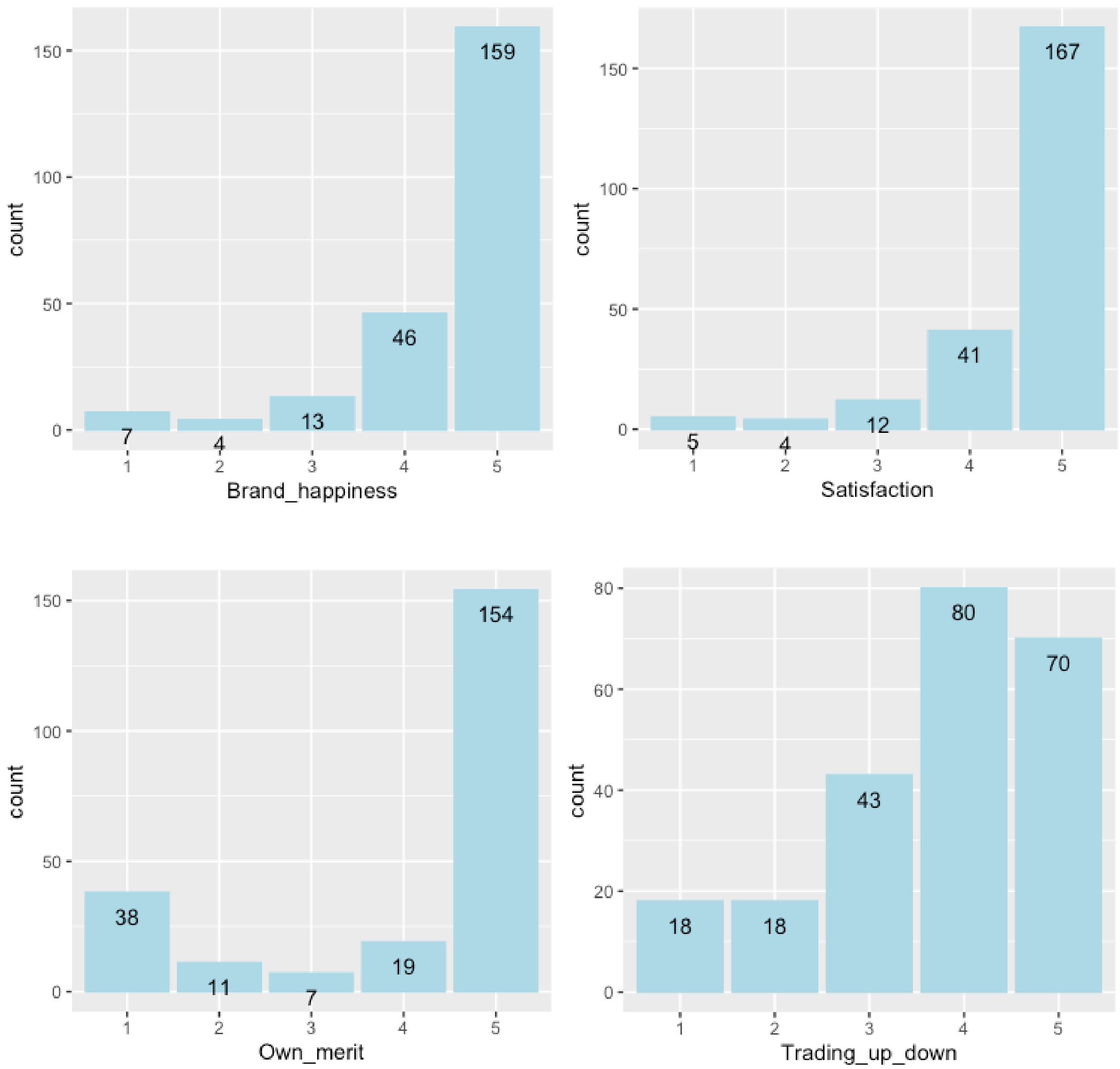
Wiedmann, K.-P., Hennings, N., & Siebels, A. (2007). Measuring Consumers' Luxury Value Perception: A Cross-Cultural Framework. *Academy of Marketing Science Review*, *2007*, 1. International Bibliography of the Social Sciences (IBSS); ProQuest One Business.

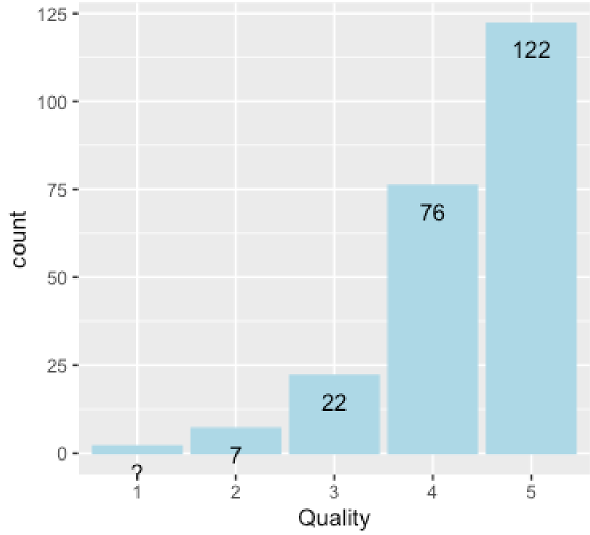
Wolpert, D. H. (1992). Stacked generalization. *Neural Networks*, *5*(2), 241–259. https://doi.org/10.1016/S0893-6080(05)80023-1

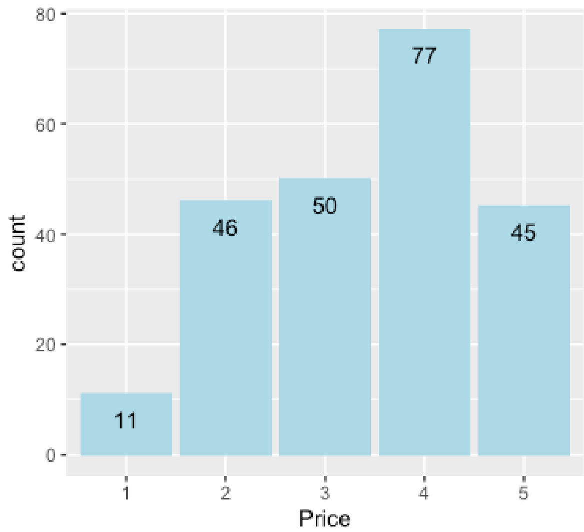
Zhang, L., & Zhao, H. (2019). Personal value vs. luxury value: What are Chinese luxury consumers shopping for when buying luxury fashion goods? *Journal of Retailing and Consumer Services*, *51*, 62–71. https://doi.org/10.1016/j.jretconser.2019.05.027

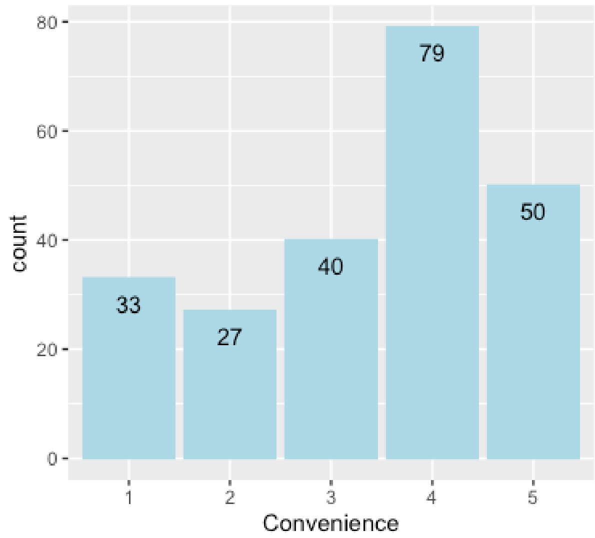
Appendix

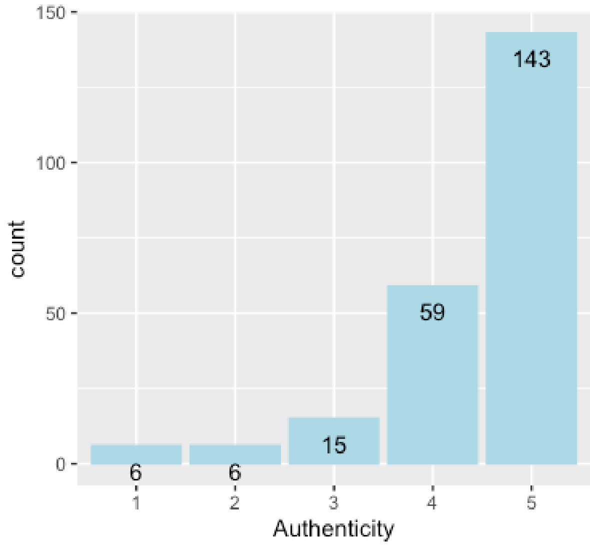
Appendix 1: Distribution per item of the questionnaire among participants on a 5-point Likert scale without dependent variable (iPhone ownership)

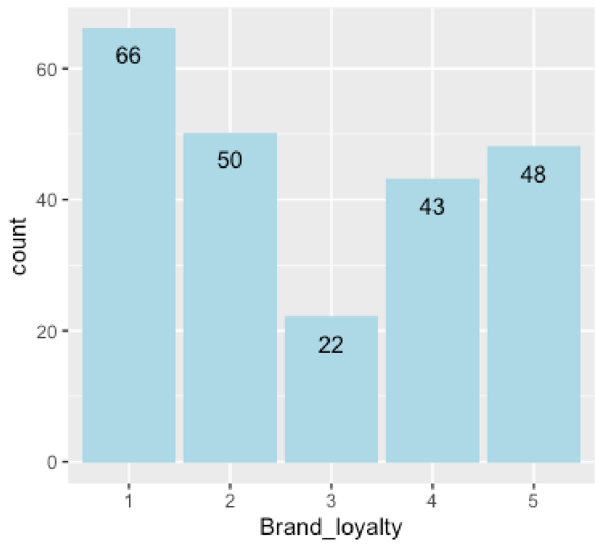


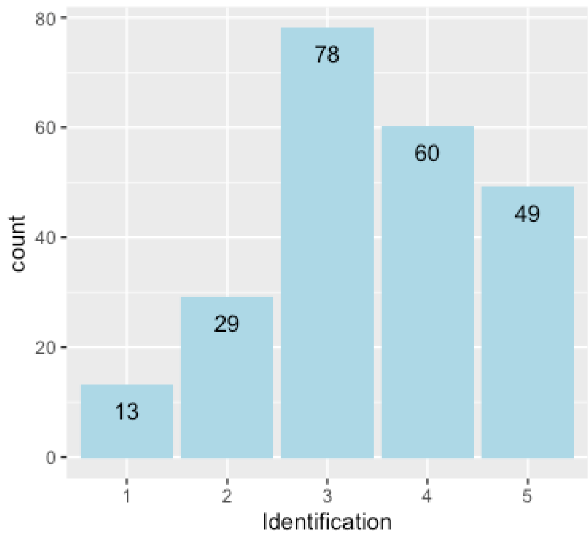

125
100
75
50
25
0
count
?
7
22
76
122
1
2
3
4
5
Quality

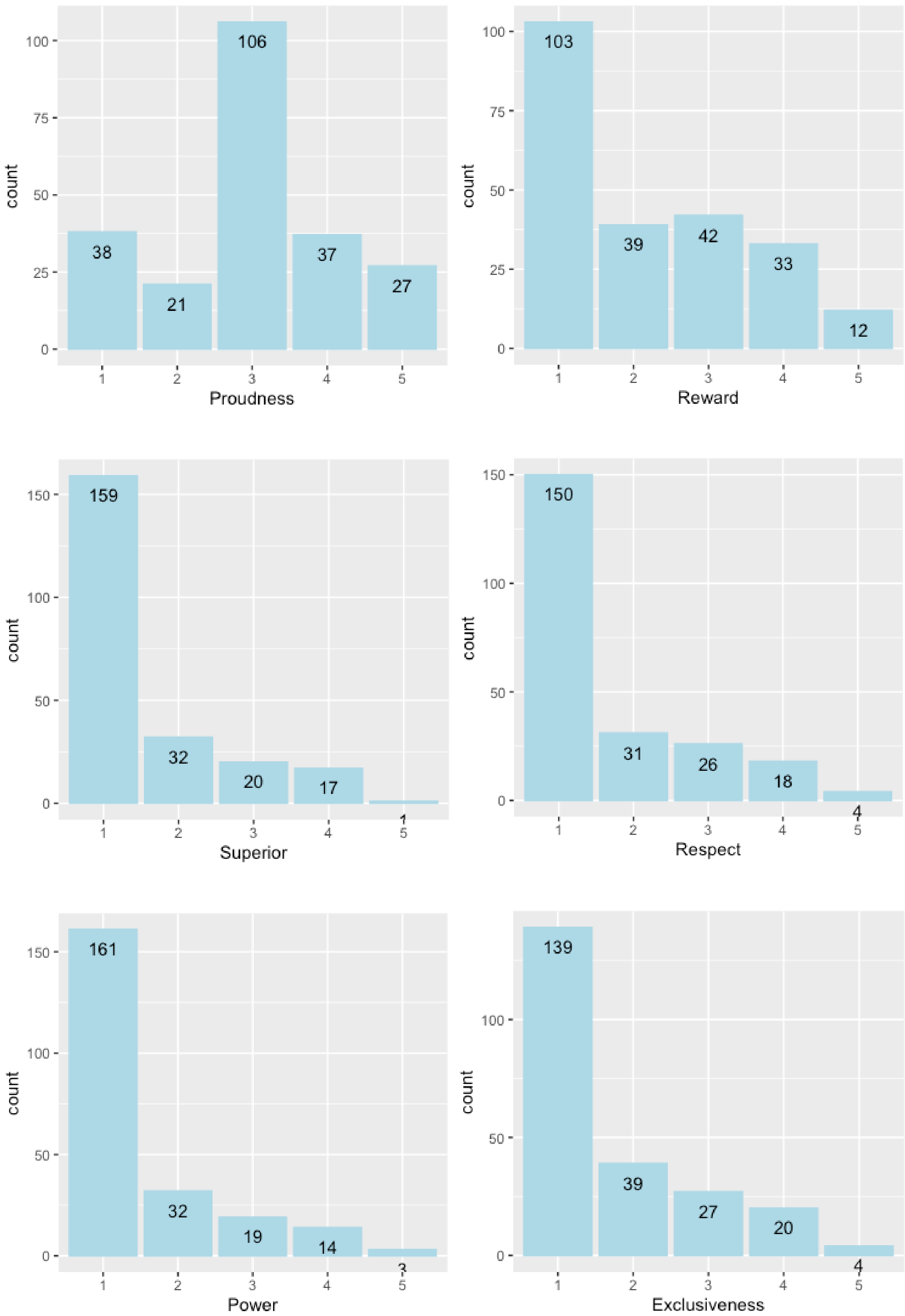

80
60
40
20
0
count
11
46
50
77
45
1
2
3
4
5
Price

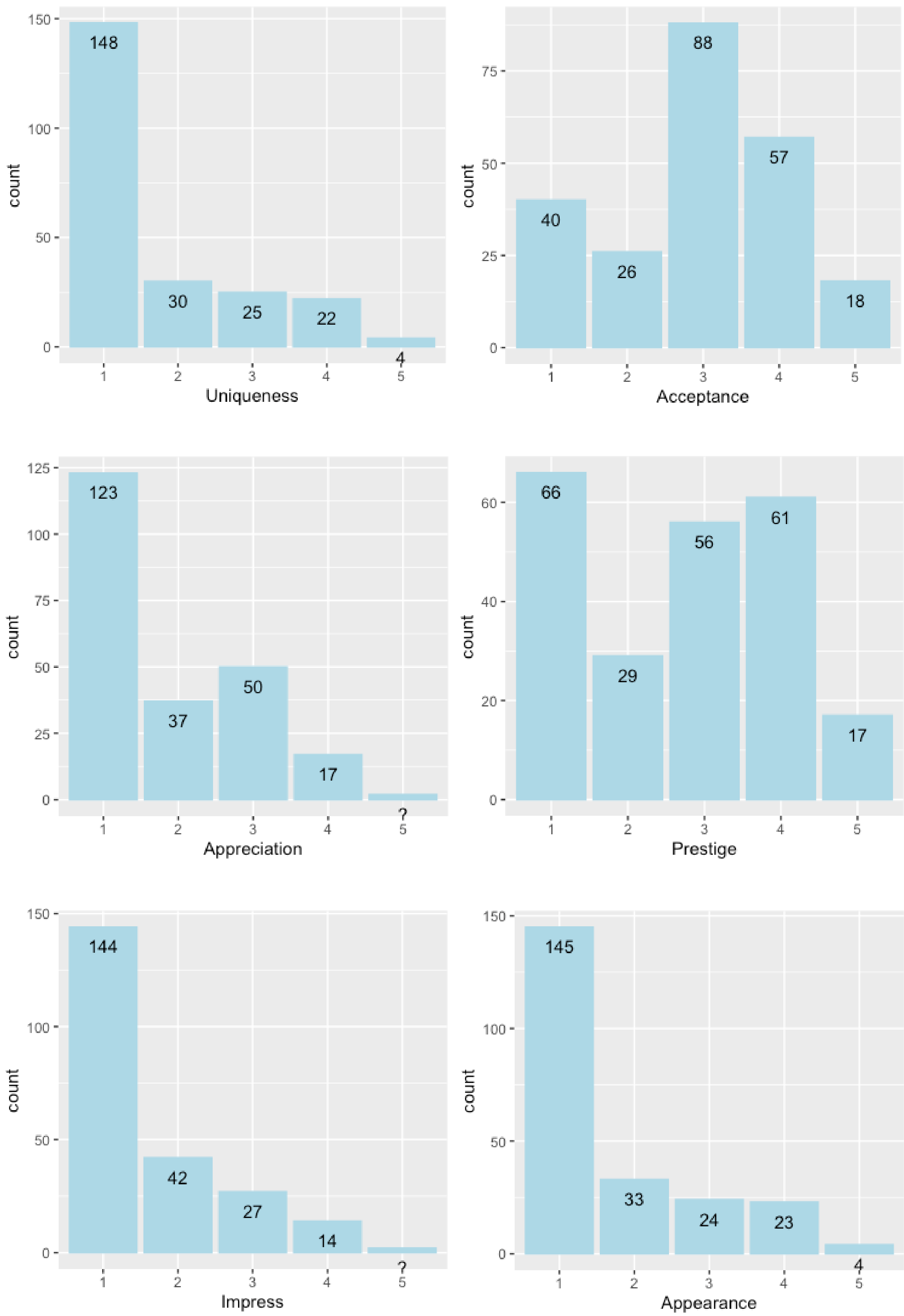

80
60
40
20
0
count
33
27
40
79
50
1
2
3
4
5
Convenience

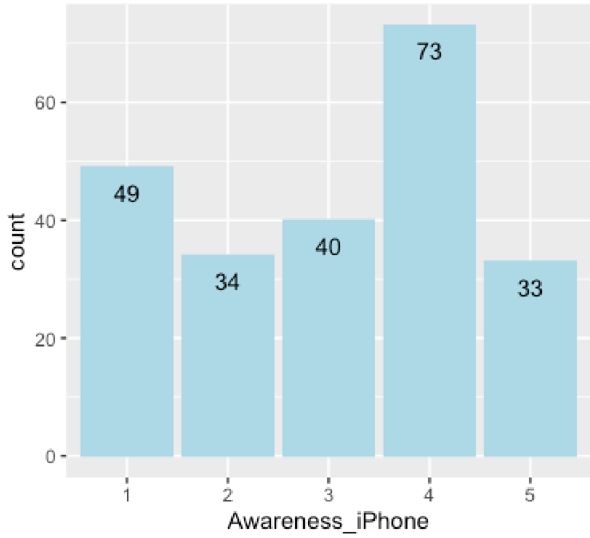

150
100
50
0
count
6
6
15
59
143
1
2
3
4
5
Authenticity

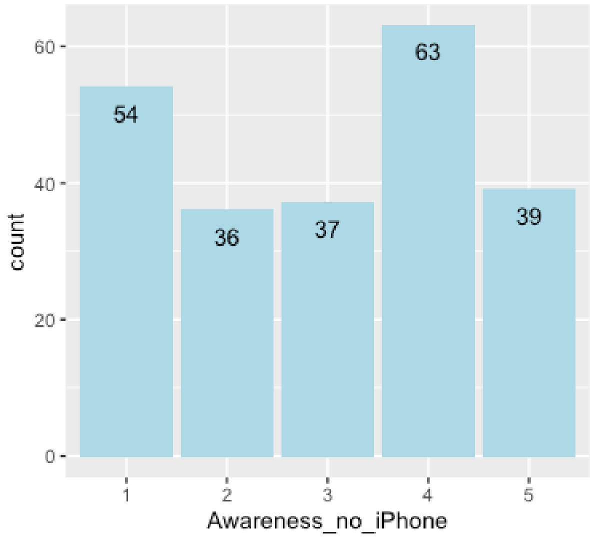

60
40
20
0
count
66
50
22
43
48
1
2
3
4
5
Brand_loyalty

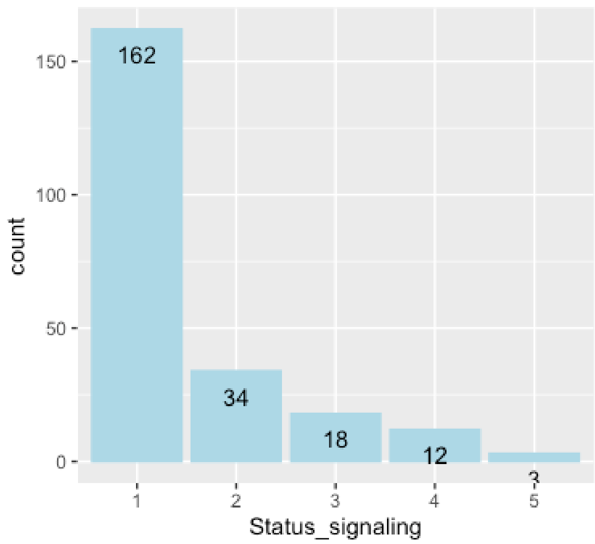

80
60
40
20
0
count
13
29
78
60
49
1
2
3
4
5
Identification

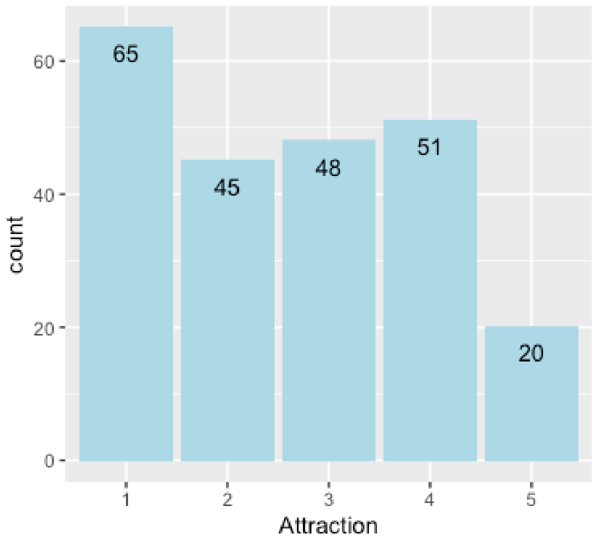

count
Proudness
106
38
21
37
27
Reward
103
39
42
33
12
Superior
159
32
20
17
1
Respect
150
31
26
18
4
Power
161
32
19
14
3
Exclusiveness
139
39
27
20
4

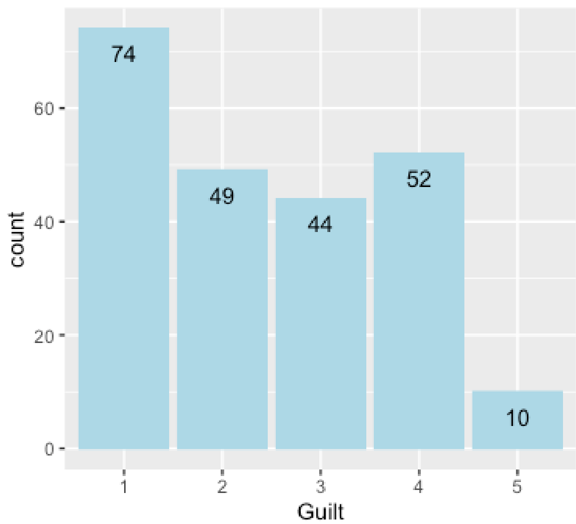

count
Uniqueness
148
30
25
22
4
Acceptance
40
26
88
57
18
Appreciation
123
37
50
17
Prestige
66
29
56
61
17
Impress
144
42
27
14
Appearance
145
33
24
23
4

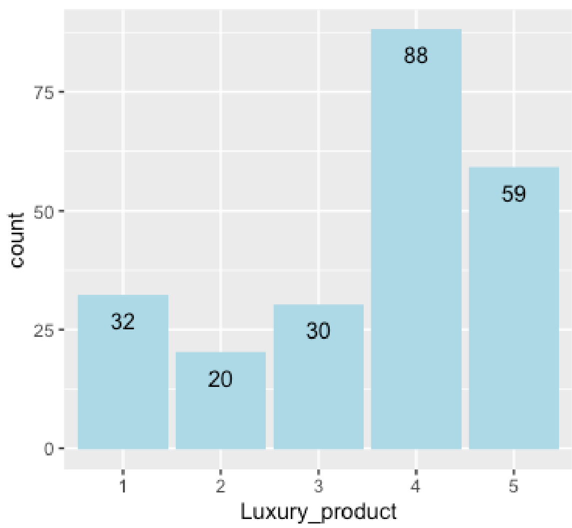

count
0
20
40
60
49
34
40
73
33
1
2
3
4
5
Awareness_iPhone

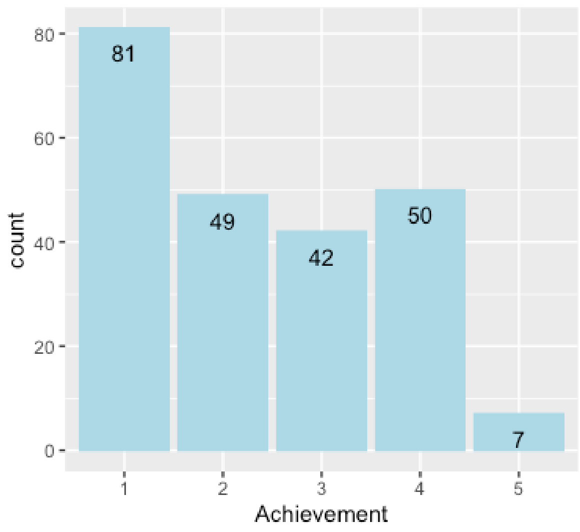

count
0
20
40
60
54
36
37
63
39
1
2
3
4
5
Awareness_no_iPhone

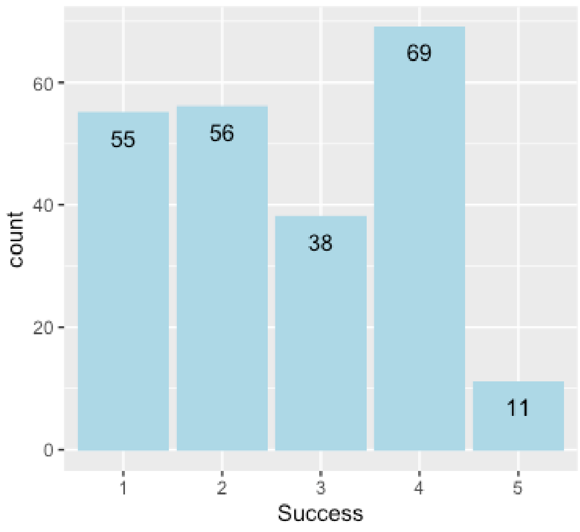

count
0
50
100
150
162
34
18
12
3
1
2
3
4
5
Status_signaling

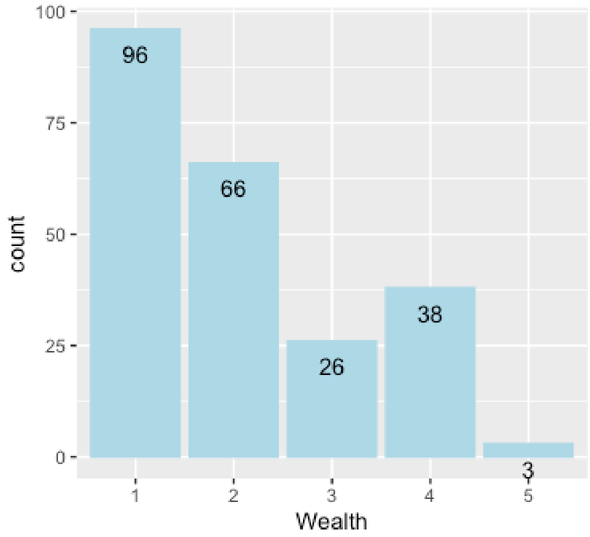

count
0
20
40
60
65
45
48
51
20
1
2
3
4
5
Attraction

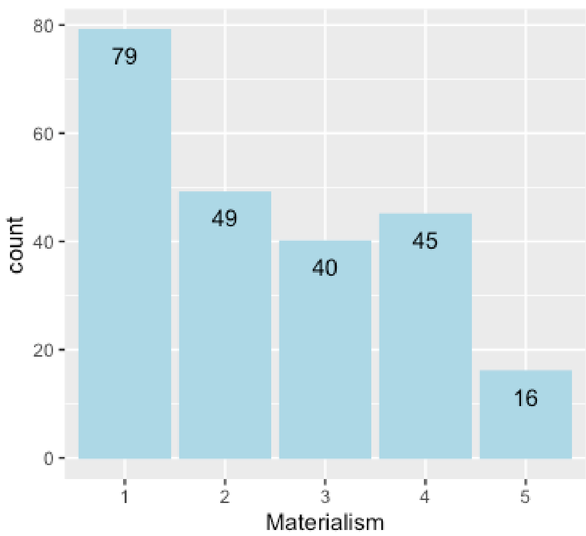

count
0
20
40
60
74
49
44
52
10
1
2
3
4
5
Guilt

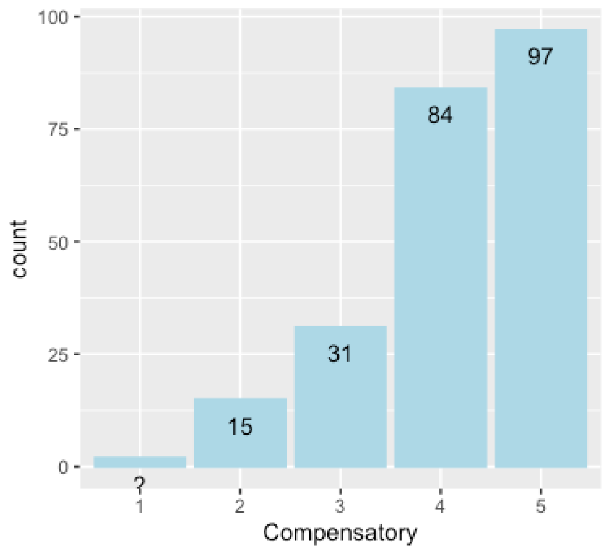

count
0
25
50
75
32
20
30
88
59
1
2
3
4
5
Luxury_product


count
Achievement
81
49
42
50
7
0
20
40
60
80
1
2
3
4
5

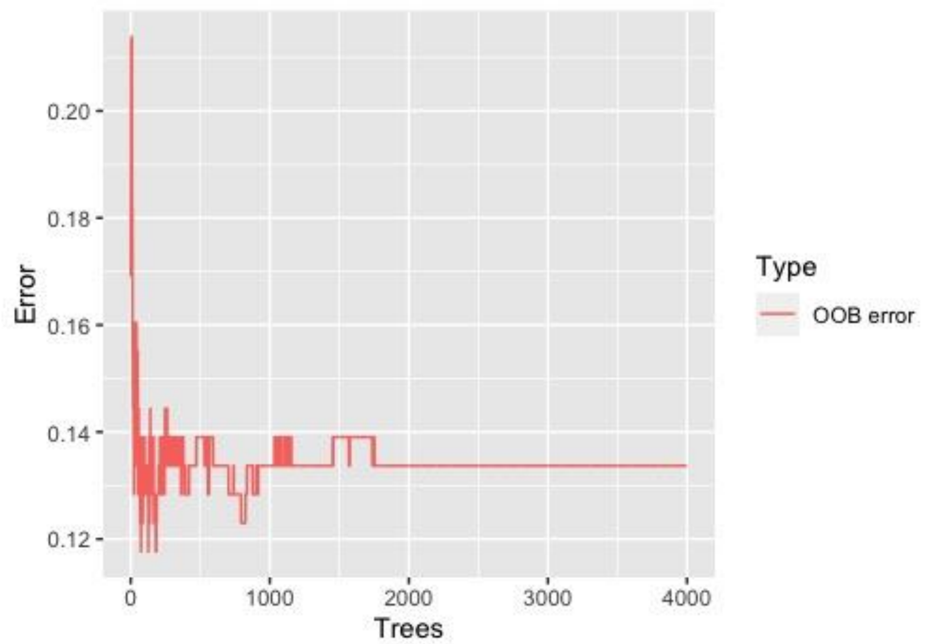

count
Success
55
56
38
69
11
0
20
40
60
1
2
3
4
5


count
Wealth
96
66
26
38
3
0
25
50
75
100
1
2
3
4
5

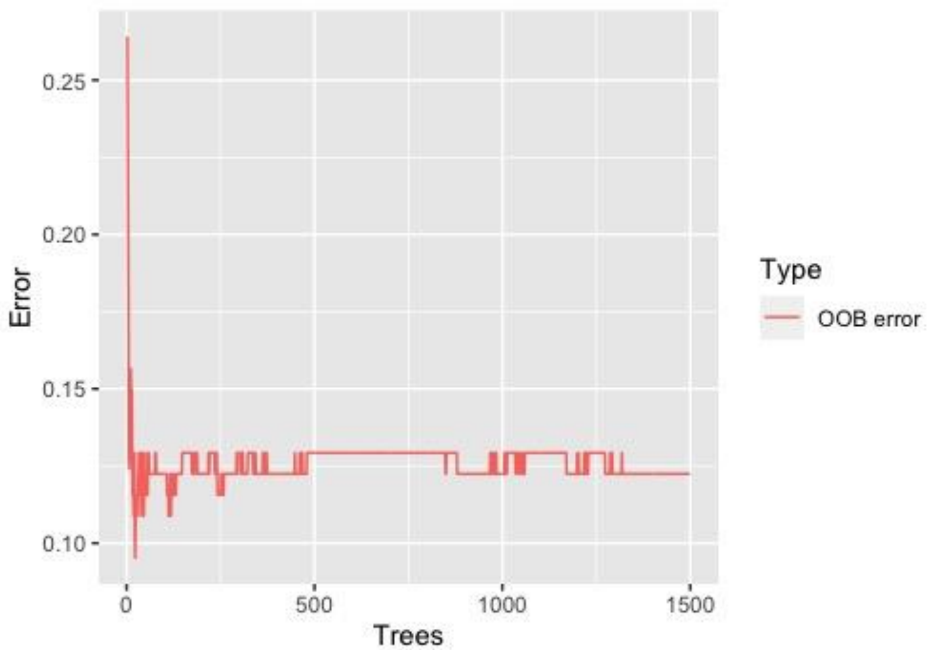

count
Materialism
79
49
40
45
16
0
20
40
60
80
1
2
3
4
5


count
Compensatory
2
15
31
84
97
0
25
50
75
100
1
2
3
4
5

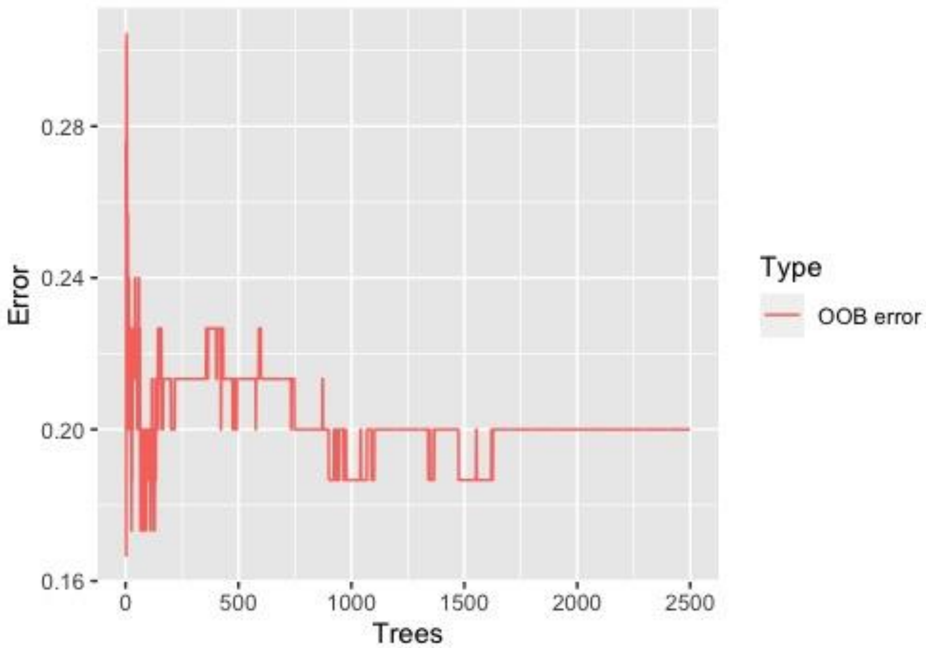
Appendix 2: OOB error rates to determine number of trees (ntrees) within the different RF models



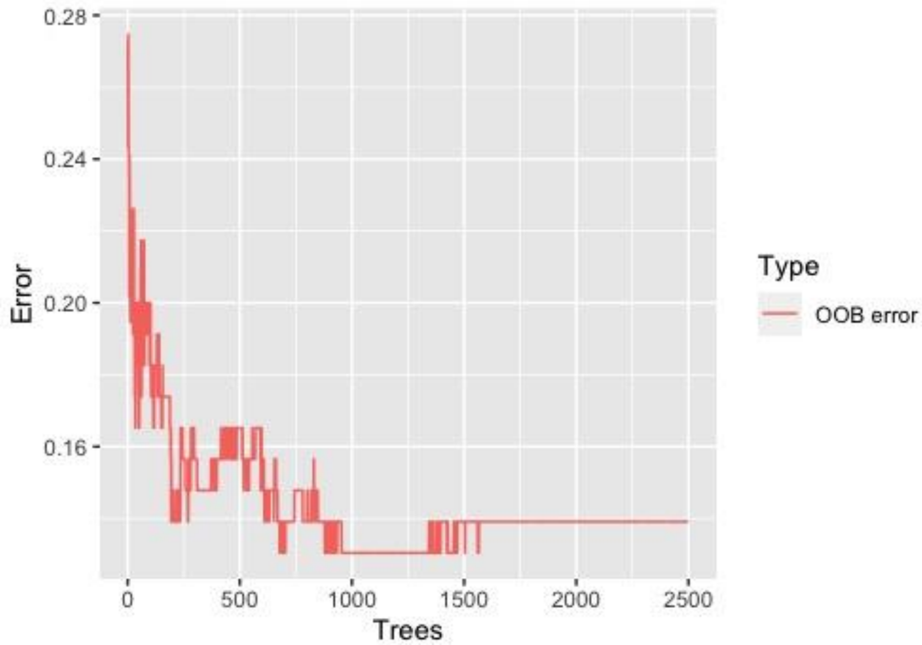
OOB error rates for global analysis
indicating stability after around 2000 trees



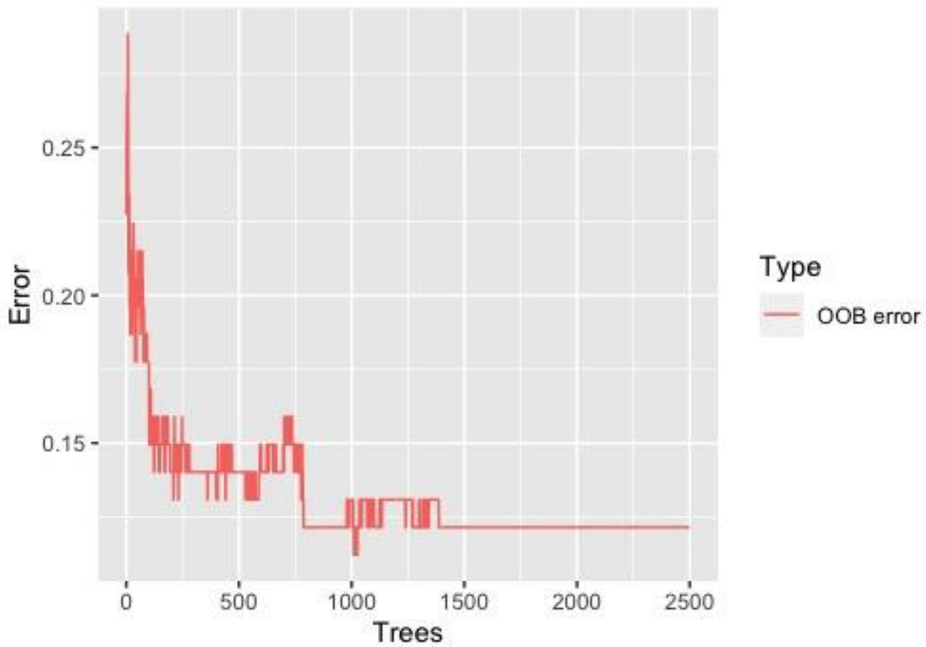
OOB error rates for local analysis cluster 1
indicating stability after around 1500 trees



OOB error rates for local analysis cluster 2 indicating stability after around 2000 trees



OOB error rates for local analysis female cluster indicating stability after around 2000 trees



OOB error rates for local analysis male cluster indicating stability after around 1500 trees