MODUL UNIVERSITY
PRIVATE UNIVERSITY VIENNA

Identification of the most influential alpine ski resort characteristics

---

Master Thesis submitted in fulfillment of the Degree

Master of Science

in Management

Submitted to Dr. Christian Weismayer

Yaroslav Martyniuk

1721519

Vienna, 7th December 2022

AFFIDAVIT

I hereby affirm that this Master's Thesis represents my own written work and that I have used no sources and aids other than those indicated. All passages quoted from publications or paraphrased from these sources are properly cited and attributed.

The thesis was not submitted in the same or in a substantially similar version, not even partially, to another examination board and was not published elsewhere.

Date	Signature

ABSTRACT

This study aimed to find out which aspects of alpine ski resorts have the most impact on customers' overall satisfaction with the use of multiple linear regression. As a result, 18 aspects of the alpine ski resort have a significant impact and, therefore, can predict overall satisfaction. Furthermore, using hierarchical cluster analysis, this research divides customers into segments based on their satisfaction levels with different characteristics. While the results of multiple linear regression analysis provided an informative number of the variables that are significantly good at predicting overall satisfaction, cluster analysis of the customers based on their satisfaction with the variables from the multiple linear regression analysis did not prove the presence of multidimensional differences, which means the grouping was done based on the overall satisfaction level, rather than on the differences between the individual level of satisfaction with different variables. A cluster analysis of the characteristics with significant predictability of overall satisfaction was also performed, grouping the attributes according to the distances between customer evaluations. This resulted in the identification of the three themes important to the customers.

Keywords: cluster analysis, multiple linear regression, ski resort characteristics.

ACKNOWLEDGMENTS

This section is a special acknowledgment of the author's gratefulness for the people and companies helping in preparation for and during the research:

Dr. Christian Weismayer

Modul University Vienna GmbH

Klaus Grabler

Manova GmbH

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

WKÖ – Austrian Chamber of Commerce

KPI – Key Performance Indicator

NPS – Net Promoter Score

GIS – Geographical Information Systems

MLR – Multiple Linear Regression

IBM – International Business Machines

UNIX – Uniplexed Information Computing System

1 INTRODUCTION

The alpine ski business is a challenging environment that keeps ski resorts under much pressure (Vanat, 2020). Even though there is a lot of competition and risk in the alpine skiing industry, it needs a long-term growth strategy to overcome specific difficulties. (Vanat, 2020) A lack of snow, inadequate snow depth, and early snow melting are only a few ways climate change may impact this branch, and there are other potential pitfalls made by other crises (Campos Rodrigues et al., 2018; Elsasser & Messerli, 2001; Haugom & Malasevska, 2019). Next, alpine skiing's popularity changes significantly depending on the client's free time, weather conditions at the destination, and the price of the whole holiday, including lift, accommodation, equipment, and other along-appearing expenses (Haugom & Malasevska, 2019). Lastly, the sector has trouble appealing to people in their 20s and 30s, who have different ways of spending money and a high demand for different kinds of entertainment (Vanat, 2020). In order to deal with the problems listed above and weigh the effects of varying long-term and short-term solutions, a better overview of the things that are important to future and current skiers should be given (Haugom & Malasevska, 2019). As mentioned, the qualities of a ski resort significantly affect the level of interest people have in going alpine skiing (Haugom & Malasevska, 2019). As a result, ski resorts must live up to the skier's high standards (Haugom & Malasevska, 2019). This often results in specific characteristics attracting specific customer groups, and knowing the appealing attributes requires research like this one (Gössling et al., 2012).

1.1 Context and Previous Research

Previous research in the alpine ski field confirms that the ski resort's characteristics have a crucial influence on ski resort attractiveness, along with various uncontrolled parameters (Haugom & Malasevska, 2019). Gössling and colleagues in their study created a figure visualizing the impact of different parameters on destination attractiveness, as presented in Figure 1. (Gössling et al., 2012).

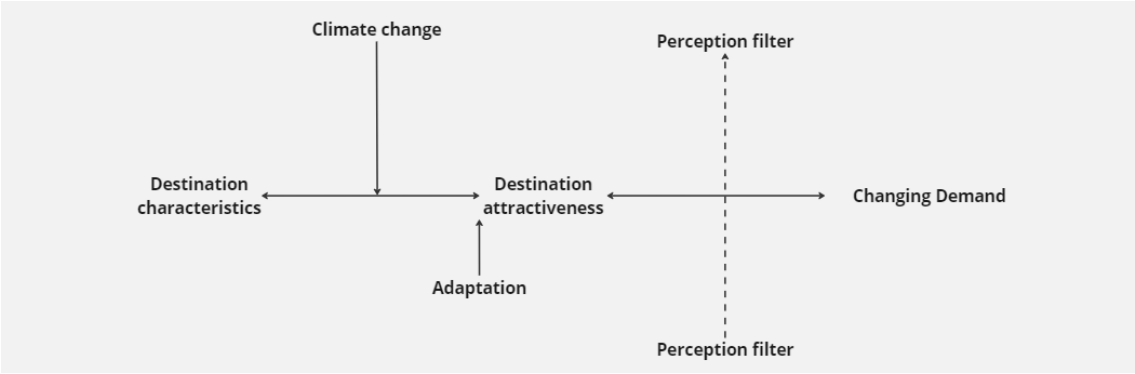


FIGURE 1: THE ROLE OF PERCEPTIONS IN DEFINING DESTINATION ATTRACTIVENESS (GÖSSLING ET AL., 2012)

Haugom and Malasevska adjusted this figure for the purpose of their alpine ski resort study, as visualized in Figure 2 (Haugom & Malasevska, 2019). Figure 2 is more applicable for the alpine ski studies and does not require any additional adjustments, as it is already concerned with specific uncontrolled factors such as weather, ski ability, and pandemic, as well as controlled factors, such as ski resort characteristics (Haugom & Malasevska, 2019).

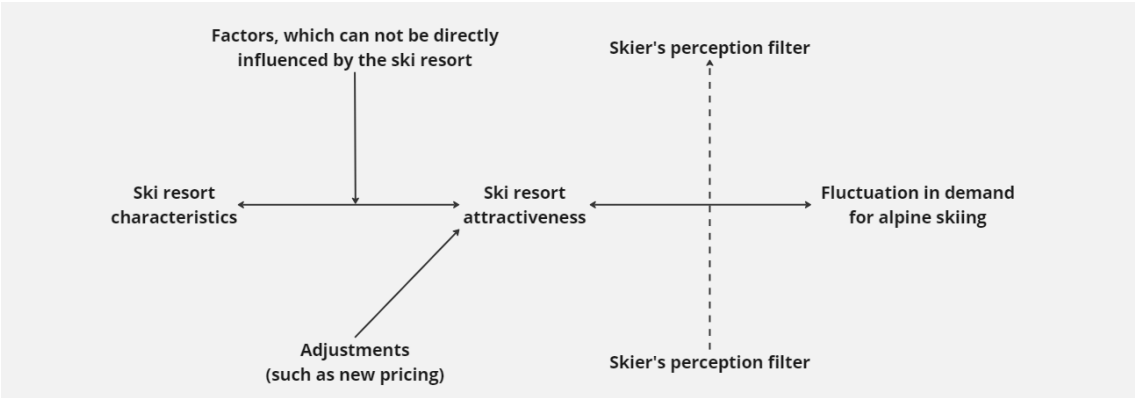


FIGURE 2: DEFINING SKI RESORT ATTRACTIVENESS THROUGH CUSTOMER'S PERCEPTION (GÖSSLING ET AL., 2012; HAUGOM & MALASEVSKA, 2019)

Figures 1 and 2 represent the foundation of most of the market research studies of alpine skiing resorts. Part of the studies focuses on uncontrolled factors, trying to find out how those influence the demand and profits of ski resorts. (Falk, 2015; Falk and Vieru, 2016; Holmgren and McCracken, 2014). As a result of those studies, the customers' leisure time and weather-related parameters were considered the essential uncontrolled parameters influencing the demand for ski resorts (Haugom & Malasevska, 2019). The next part of the studies focuses on the ski resort characteristics and their influences along with different customer segmentation approaches (Park & Yoon, 2009; Jang et al., 2002). The literature review chapter about similar studies inspects the customer segmentation studies in the alpine ski market more closely. Other studies look at how ski resorts can adjust the controlled parameters to unfavorable uncontrolled ones to maximize the ski resort attractiveness (Haugom & Malasevska, 2019). Furthermore, one of

the best mitigators of uncontrolled parameters is found to be the price (Haugom & Malasevska, 2019). Further studies in market analysis of the alpine ski have done the conjoint analysis to identify the relative importance of the ski resort parameters (Carmichael, 1996; Siomkos, Vasiliadis, and Lathiras, 2006). Carmichael, in her research, has done a conjoint analysis of ski resort parameters to improve the data-gathering process for the further creation of customer groups (Carmichael, 1996). Using the conjoint analysis, Siomkos and his colleagues discovered that the price of the lift and gastronomy, as well as accessibility parameters, have higher relative importance than other parameters studied, such as services and premises quality (Siomkos, Vasiliadis, and Lathiras, 2006). The research done by Won and colleagues found that snow quality is an equally important parameter for all types of ski resort guests (Won, Bang and Shonk, 2008). In contrast, other parameters' importance varies between snowboarders and skiers (Won, Bang and Shonk, 2008). While not all the possible study examples in the field of the alpine ski market are mentioned above, all of them strive to provide the data necessary for the creation of alpine ski business strategies, which in turn motivated the author to conduct this research.

1.2 Structure of The Thesis

There are five main chapters: introduction, literature review, methodology, results and discussion, and conclusion. Each chapter has a short foreword and afterword connecting the chapters for a better read-flow between them. The introduction chapter consists of context and previous research subchapters, the structure of the thesis, and general background, including the problem identification and objective development section. Next comes the literature review, which overviews the following topics: customer satisfaction, satisfaction theories and models, customer segmentation, similar studies, Austrian winter tourism, instruments, intended software usage, customer persona as a design tool, conceptual framework, and hypothesis development. The next part is methodology. It is concerned with research approaches, designs, and worldviews, followed by a selection of methodology, data acquirement, source and reliability, relevant survey parameters, limitations, ethical considerations, and finally, data analysis. The following major part is results and discussion, and it is divided into two subchapters discussing multiple linear regression and cluster analysis results, respectively. The final chapter is a conclusion. It elaborates on the topics of contribution to knowledge and implications for relevant stakeholders and also suggests the future development of the research.

This introduction is finished with a general background, including information about the market that should benefit from this research. Finally, the research's problem identification and objective are provided before continuing with the literature review.

1.3 General Background

Rural municipalities of Austria rely heavily on revenue from tourists, and many small towns in rural areas base their economies mainly on the hospitality business (Steiger & Scott,

2020). Nevertheless, after the United States and France, Austria is the third biggest destination for skiers and snowboarders per capita globally (Vanat, 2020). According to Steiger and Scott, Austrian ski resort visitors spend € 7.9 billion per year (including direct, indirect, and induced impacts) on their mountain trips, based on the data from 2018 provided by the Austrian Chamber of Commerce (Steiger & Scott, 2020). The data as of the season 2020/21 indicates expenditure of € 11.2 billion (including cable cars, accommodation, gastronomy, sports retail, and transport) (Wirtschaftskammer Österreich, 2022). Thus, the skiing branch provides 2.65% of jobs in Austria (Statistik Austria, 2018b, cited by Steiger & Scott, 2020; Wirtschaftskammer Österreich, 2022). The strategic planning and continued investments of Austria's domestic lift firms are crucial to the prosperity of the country's winter tourist industry (Wirtschaftskammer Österreich, 2022). More than 9 billion euros have been invested by Austrian ropeways since 2000 (6 billion of which in the last ten years) to modernize and improve the comfort of facilities and to improve snowmaking technology to meet the rising expectations of winter tourists (Wirtschaftskammer Österreich, 2022). An example of such expenditures in season 2020/2021 can be seen in Figure 3: € 147 million were spent on safety, quality, and comfort, € 65 million on snowmaking, and € 211 million on other amenities (Wirtschaftskammer Österreich, 2022). More information on that topic is provided in the Austrian winter tourism chapter.

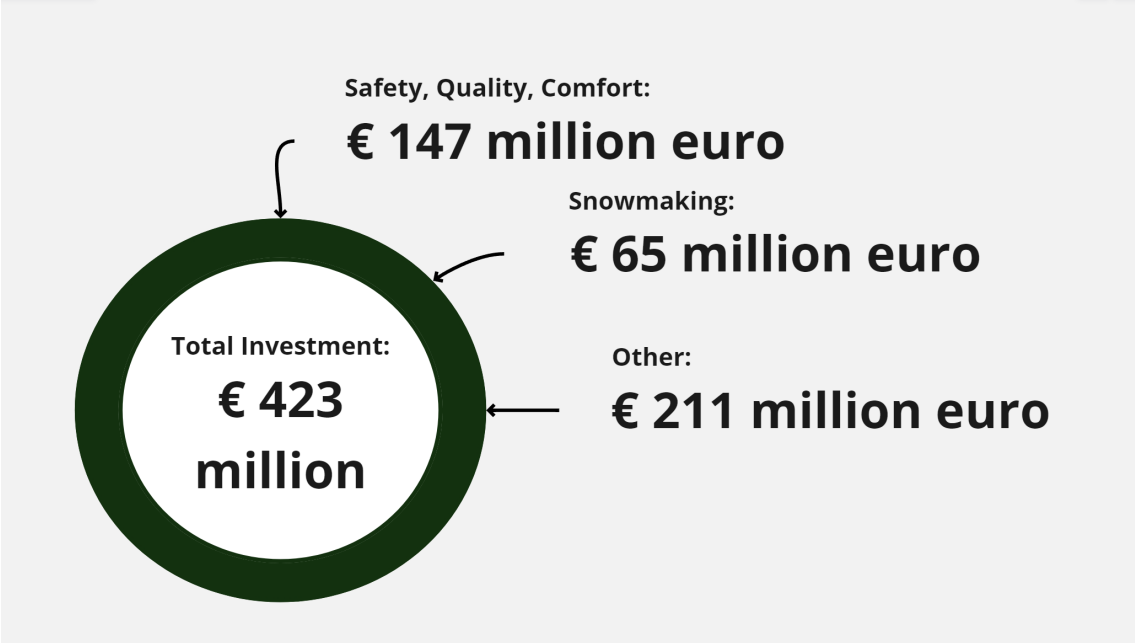


FIGURE 3: INVESTMENTS IN SEASON 2020/2021 (WIRTSCHAFTSKAMMER ÖSTERREICH, 2022).

While investors are considering whether they should invest more in snowmaking during the ongoing crises (e.g., COVID-19 or climate change) or switch the direction of their investments to non-snow influenced alternatives, the ski businesses in Austria require reliable sources of information, so they can make their investment and advertising decisions (Steiger & Scott, 2020).

1.3.1 Problem Identification and Objective Development

For the provision of such data, the author narrowed in on a problem statement, an aim, and some research questions to investigate. Alpine ski businesses base their advertising and investment decisions on their business KPIs (Key Performance Indicators). The problem is, they do not know for sure which ones are the key performance indicators without conducting such a study on a regular basis, as the indicators can change with time. This research aims at providing factual data regarding the alpine ski businesses' most influential characteristics, which can reliably assist in making investing and advertising decisions, and also at providing an algorithm for further repetition of this study when required after a long period of time or after a sudden market change.

The best way to reach the goal of the study is to answer the following research question: “Which characteristics of the skiing area influence the customers’ overall satisfaction with their holiday?” Along with secondary ones: “What are the “a-posteriori” groupings of customers based on their satisfaction?” and “What are the “a-posteriori” groupings of the most influential characteristics based on customer satisfaction?” The hypothesis of this study tests the starting hypothesis suggesting that “There are no characteristics of the alpine ski resort that can predict the overall satisfaction of the customer of the skiing area in Austria.” And if the significance tests will prove $p<0.05$, the starting hypothesis will be rejected, and the new hypothesis that “There are characteristics of the alpine ski resort that can predict the overall satisfaction of the customer of the skiing area in Austria” will be accepted. The hypothesis development is easier to understand after the main concepts, scales, and intended techniques are presented in the literature review. Therefore, hypothesis development is a part of the literature review.

2 LITERATURE REVIEW

As an introduction to the literature review, its structure is provided. It starts with defining customer satisfaction, including several subchapters on its importance, effects, measurements, KPIs, and scales, followed by applicable theoretical models, paradigms, and theories. The following chapter elaborates on the topic of customer segmentation. Furthermore, the next one is about similar studies. Before switching the focus to more technical things, a short background about Austrian winter tourism is given. The following part of the literature review discusses the techniques and instruments planned to be used. After that, the hypotheses are presented. The next part explains the software usage required for implementing the techniques while testing the hypotheses. The last three parts of the literature review are the customer persona as a design tool, the conceptual framework of the research, and the hypothesis development.

2.1 Customers' Satisfaction

2.1.1 Defining Customers' Satisfaction

There is no unique definition of satisfaction, which the researchers agreed upon and used in their studies. As Oliver addressed in 1997, the term is so simple that people assume they know it but struggle to define it when asked to do so (Oliver, 1997). Since it is often believed that satisfaction has already been defined, most studies ignore definitional issues in favor of evaluating models of customer contentment (Giese & Cote, 2000). Therefore, it is important to consider the work of Peterson and Wilson from 1992, which suggested that the lack of standardization in the definition is the main description of all satisfaction studies (Peterson & Wilson, 1992). Further evidence of the instability of the definition lies in the fact that researchers still debate whether customer satisfaction is a continuous process or a simple outcome (Giese & Cote, 2000). Most definitions offered by researchers suggest that customer satisfaction is an output of evaluation processes (Giese & Cote, 2000).

In order to provide the reader with an explanation of customer satisfaction, the framework of different definitions offered by different researchers is provided next.

The article "Defining Consumer Satisfaction" by Giese and Cote provides a framework consisting of twenty different definitions in twenty studies (Giese & Cote, 2000). Those studies were done by different authors and sometimes the same authors but in different circumstances, which still does not guarantee that they covered all the existing interpretations (Giese & Cote, 2000). The following definitions are considered helpful in this particular research. Also, the outcomes of the work done by Giese and Cole are discussed at the end.

Robert A. Westbrook and Michel D. Reilly talk about the emotional response of the customer after the purchase (Giese & Cote, 2000; Westbrook & Reilly, 1983). According to their work written in 1983, the satisfaction of the customer can be described as a signal response on the emotional level to the service or product the customer bought (Westbrook & Reilly, 1983). This emotional response evaluates all the aspects, such as the place, the employees, the value of the service/product overall, and situationally other parameters (Westbrook & Reilly, 1983). The following interpretation also relies on customers' emotions but leads to different conclusions. Richard L. Oliver, in his journal article in 1992, focused on product attributes during consumption and researched contentment as an emotion, and finally, he concluded that contentment is not an isolated consumer emotion but rather a phenomenon that exists with other such feelings (Giese & Cote, 2000; Oliver, 1992). In 1997 when considering services and products during their consumption Oliver mentioned that the client's fulfillment, more specifically over- or under-fulfillment during the consumption, creates a judgment in the client's mind regarding it being a satisfying or dissatisfying process (Oliver, 1997; Matsuoka, 2022). The following explanation of customer satisfaction was offered even earlier, in 1977. This alternative definition of customer satisfaction by Hunt describes it as an appraisal of whether or not the experience met expectations (Hunt, 1977). Such evaluations are done during the consumption of the product or service (Giese & Cote, 2000).

Furthermore Hunt's definition was expanded by Tse and Wilton by stating that client contentment is the activity by which users reply to the assessment of the apparent disparity between their pre-consumption assumptions and the package's overall performance (Tse & Wilton, 1988). Oliver added on that aspect that it is a value judgment (Oliver, 1997). Siskos and Grigoroudis combined those two mentioned aspects in their book from 2010, saying that one-half of this formulation highlights the visual, cognitive, and emotional process that leads to happiness, while the other describes satisfaction as an end outcome or as an end-state arising from the consumer experience (Grigoroudis & Siskos, 2010). Churchill and Surprenant offer another alternative definition. They identify the four main motivational factors of satisfaction: evaluation of competence, expectations, subtractive and subjective disconfirmation (Churchill & Surprenant, 1982; Matsuoka, 2022). They build their definition of satisfaction of customer around comparing the costs and the rewards achieved through the exchange, much as price divided by value equation (Giese & Cote, 2000).

Summarizing the multiple definition options of satisfaction, including the ones mentioned above, Giese and Cote concluded that satisfaction in the literature is understood as some expression of emotion, thought, or will in response to a researcher's questioning about a product/service or its qualities before, during, or after the judging process, after using the product, after a period of prolonged use, or at any other moment, the researcher deems appropriate (Giese & Cote, 2000).

2.1.1.1 Customers Dissatisfaction

After so much is said about customer satisfaction, it is essential to mention the other extreme – dissatisfaction. According to Oliver, both levels of contentment and discontentment reflect the prevailing emotional environment (Oliver, 1992). Herzberg's motivation-hygiene theory underlies the concept's two-dimensionality; therefore, it is crucial to evaluate both sides of the scale during the research (Oliver, 1980).

2.1.2 Measuring Customer Satisfaction

These days, most companies place a high value on evaluating consumer happiness (Hill,2006). Over the last several decades, the value of happy customers has grown, prompting many businesses to explore making customer satisfaction measurement a core KPI metric (Grigoroudis & Siskos, 2010). According to multiple specialists in the area, customer satisfaction as a KPI helps employees to motivate themselves to perform better and to keep the resulting measurements high, thus keeping the overall company performance at a significantly good level (Gerson, 1993; Hill, 2006; Wild, 1980). Nonetheless, the process of gathering customer satisfaction data requires time and consistency, and the final results of this process should help improve different divisions of the business (Grigoroudis & Siskos, 2010). For example, it can help determine a company's present position in the market in relation to its competitors or outline previously unseen market opportunities, apart from the obvious help in identifying customer needs and wants (Grigoroudis & Siskos, 2010).

Thus, it is important to pay attention to the measures and scales of customer satisfaction during the research. Therefore, the following parts speak about measuring the KPIs of customer satisfaction and scales used in this process.

2.1.3 KPIs for Measuring Customers' Satisfaction

Cleave identified 6 KPIs for measuring the customers' satisfaction as follows: customer satisfaction score, customer effort score, net promoter score (NPS), SERVQUAL, customer retention rate, and churn rate. All of those can be measured differently (Cleave, 2019). Customer satisfaction scores can be measured using the Likert, Semantic Differential, or slider scales (more on those in the Scales chapter) (Cleave, 2019; Sauro, 2019). The customer effort score measures the user-friendliness of the service and can use the same tools used for overall satisfaction (Cleave, 2019). The NPS is trying to count the number of clients who actively share how good the service is through word of mouth and recommend it to their friends and family (Cleave, 2019). The customer retention rate measures the ability of the business to retain its customers. SERVQUAL metric is a combination of service and quality measured through a measurement model including reliability, assurance, tangibles, empathy, and responsiveness (Cleave, 2019; Parasuraman & Zeithaml, 1988).

Finally, the churn rate metric is a metric that must be remembered to have a negative effect, meaning the more extensive the number, the worse the situation is. It identifies how many customers were lost during the specific cycle (Cleave, 2019). All those measurements require a tool to conduct those measurements, and such as with temperature and Celsius, Fahrenheit, and Kelvin scales, Satisfaction has different ways and tools to be measured.

2.1.4 Scales Used in Satisfaction Studies

Scales of client satisfaction can be divided into single and multiple-factor scales (Danaher & Haddrell, 1996). While using the simple one-factor scale can seem easy (overall satisfaction with a product or service with 2-9 points on a scale from being very satisfied to very dissatisfied), the outcome of such an approach may result in the researcher knowing if the customer is satisfied or not but not knowing for what particular reason (Danaher & Haddrell, 1996; Oliver, 1977). Therefore, the multi-item scale is used to ask respondents about their overall opinion on the service and other critical product or service characteristics (Danaher & Haddrell, 1996). Because of its informativeness, customer satisfaction researchers tended to use a multi-item approach to increase the reliability of their studies (Bearden & Teel, 1983; Churchill & Surprenant, 1982; Danaher & Mattsson, 1994; Oliver, 1980).

Danaher and Haddrell identified a variety of different scales used in consumer research and grouped them into three categories: performance, disconfirmation, and satisfaction, as seen in Figure 4 (Danaher & Haddrell, 1996)

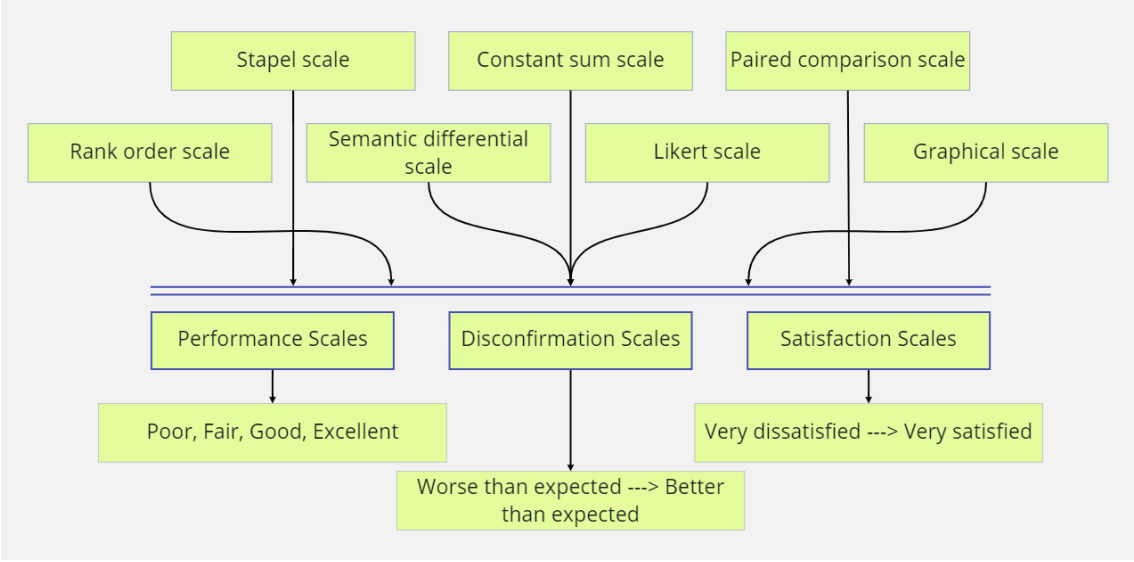


FIGURE 4: SCALES USED IN CONSUMER RESEARCH, SORTED (DANAHER & HADDRELL, 1996).

Next, the overview of the three created by Danaher and Haddrell categories is given instead of discussing each scale individually, and at the end, the individual scales which are interesting for this particular research are presented with some examples (Danaher & Haddrell, 1996).

As visualized in Figure 4, the performance scales are concerned with the level of performance described as excellent for the best possible and poor for the worst possible. Disconfirmation scales measure customers' expectations compared to the service provided, with a scale starting at worse than expected and ending at better than expected (Danaher & Haddrell, 1996). The satisfaction scale is similar to the disconfirmation scale in the visualization but measures an entirely different thing: customer satisfaction (Danaher & Haddrell, 1996). The first logical approach to measure customer satisfaction from the first view would be to choose satisfaction scales directly. Nonetheless, Devlin and colleagues and Rust and colleagues identified three reasons for choosing disconfirmation scales for customer satisfaction measurements (Devlin et al., 1993; Rust et al., 1994).

The first reason is that they use the well-known disconfirmation paradigm in the context of consumer happiness (Cadotte et al., 1987). A simple and straightforward inquiry is created instead of the one proposed in the SERVQUAL assessment as a two-factor assessment (Parasuraman & Zeithaml, 1988). Simply put, respondents give the service a rating based on how well it meets their needs (Danaher & Haddrell, 1996).

The next thing that makes this scale stand out is the fact that it can be quantitatively shown that questions involving comparison to expectations correlate more strongly with customer retention than quality or satisfaction scale-based questions (Rust et al., 1994).

The disconfirmation scale's ability to significantly lessen the imbalance in the evaluated customers' perceptions is the third and final appealing characteristic (Danaher & Haddrell, 1996). This is derived from the fact that the customer who rated the quality as excellent will not necessarily say that the service was delivered at a greater level than expected (Danaher & Haddrell, 1996).

For those reasons, the priority when choosing a scale for this research is given to disconfirmation scales followed by satisfaction and productivity scales.

The two scales discussed as helpful for this research are the Likert scale and the semantic differential scale. Both can be useful for gauging how people feel about the research's subject (Øvad, 2020)

2.1.4.1 Likert Scale

Users' approval or disapproval of a statement is measured using a Likert scale (Likert, 1932). It might be a single statement (item) or a series of items (questionnaire) (Likert, 1932). In each scenario, respondents are polled on how strongly they feel about a set of claims (Øvad, 2020). The collected data may infer a general feeling about one or more related assertions (Øvad, 2020; Likert, 1932).

2.1.4.2 Semantic Differential Scale

On a bipolar adjective scale, the participant's position toward a statement is measured using a semantic differential scale. To make the most use of the scale, the opposite adjectives should be used as the endpoint names of the scale (Øvad, 2020).

2.1.5 Customers' Satisfaction Importance

After identifying what customer satisfaction is, talking about its measurements, KPIs, and scales used to measure customer satisfaction, it is time to outline why it is so important.

According to the client-centric philosophy and the basic principles of constant development of contemporary businesses, client satisfaction is among the most critical difficulties facing businesses of all sizes and industries (Grigoroudis & Siskos, 2010). Measuring it is one of the five primary roles of organizational studies since it facilitates learning, assessment, and progress. (Massnick, 1997). The value of happy clients to businesses has grown substantially over time (Grigoroudis & Siskos, 2010). Since measuring customer satisfaction effectively, directly, meaningfully, and objectively reflects the desires and demands of consumers in a manner that is hard to fake, it has replaced all other forms of feedback as the gold standard (Gerson, 1993).

Consequently, client happiness may serve as both a minimum need and an aspirational goal for every company (Gerson, 1993). Every day, more and more businesses are using customer satisfaction as their primary KPI to ensure they focus on their customers' needs (Grigoroudis & Siskos, 2010). Nevertheless, it is not easy to sustain the enthusiasm of an organization on something as nebulous and theoretical as ensuring that customers are satisfied (Grigoroudis & Siskos, 2010). As a result, customer satisfaction must be recast in terms of a set of metrics directly related to people's jobs or elements that both employees and users can interpret or control (Deschamps and Nayak, 1995). In addition, measuring customer happiness boosts morale by rewarding workers at every level of the service chain for their efforts. An increase in productivity and performance is a direct result of using satisfaction assessment in this manner (Hill, 2006; Wild, 1980). The main focus of client behavioral research is how the client acts after the exchange is over, which shows how important it is to find out how happy the client is (Hill, 2006; Kotler, 1994).

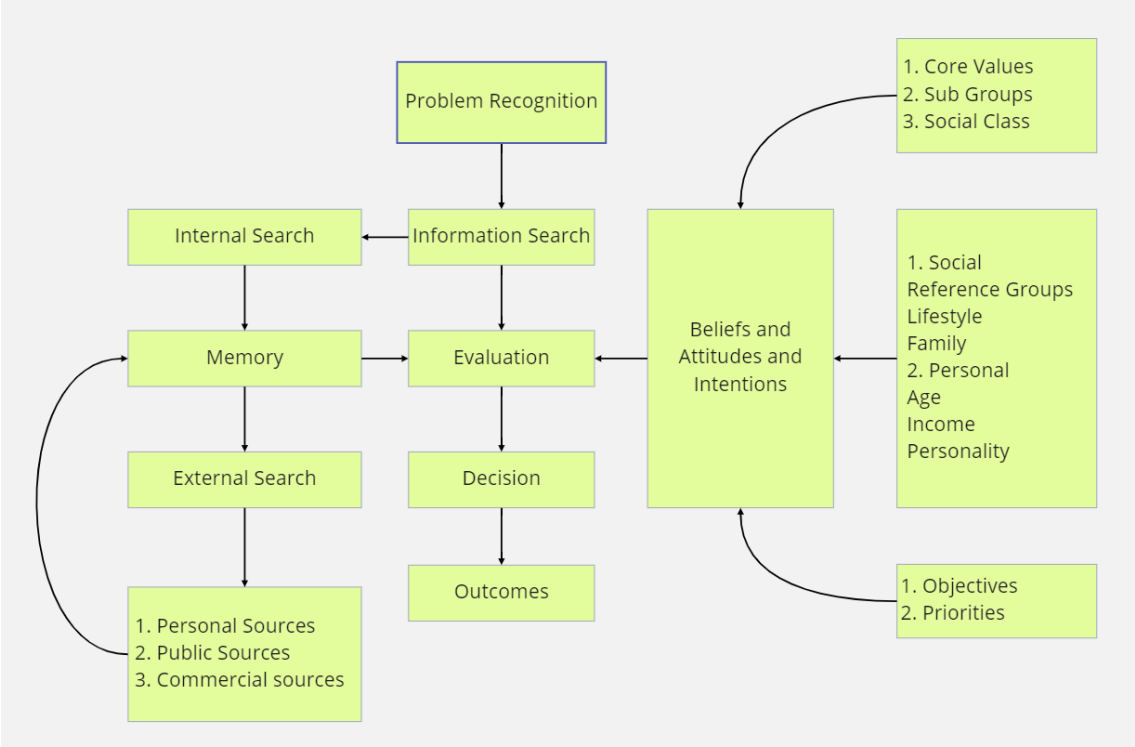


FIGURE 5: THE INDIVIDUAL BUYING DECISION PROCESS (GRIGOROUDIS & SISKOS, 2010; HILL, 2006).

In particular, previously mentioned studies endeavored to assess the outcomes of services or goods used and the impact of such utilization on buyers' post-purchase behavior, as shown in Figure 5 (Grigoroudis & Siskos, 2010; Hill, 2006).

2.1.6 Effects of Customers' Satisfaction

It was not confirmed nor denied in the study by Roger Hallowell that there is a strong relationship between customer loyalty and customer satisfaction (Hallowell, 1996). While the correlation showed a positive relation, the variance of 40% had to be considered, which means that the test results are not reliable enough to suggest a causal effect. Therefore, no solid outcomes could be shown by the study regarding their influence on each other. Nonetheless, those two terms were considered connected and advised for future research (Hallowell, 1996). Increasing the loyalty of customers is important not to lose them as increasing one's customer stock is extremely cost-intensive, it is cheaper to hold customers compared with gaining new ones (Hill, 2006).

2.1.6.1 Customer Retention

According to Hill's research, bringing in a new client takes far more effort than maintaining an old one. (Hill, 2006). A growing number of businesses' "loyalty programs" in recent years result from the fact that the expense of acquiring a new client far outweighs the value of retaining an existing one (Hill, 2006). Unsurprisingly, happy clients are the best loyalty strategy (Hill, 2006). No matter what the name of the incentive given to the customer is, loyalty systems merely play

a supporting role in achieving this goal by performing well what matters most for the buyer. The primary rationale for conducting customer satisfaction surveys is to offer businesses data that will aid in making choices that will increase user happiness and, by extension, customer retention (Hill, 2006). Management is doubtful to make intelligent choices leading to the necessary incremental increases in client retention if customer satisfaction is not measured accurately (Hill, 2006).

2.1.6.2 Customer Loyalty

According to the latest research, customer loyalty is a complex concept that depends on many factors, including buyers' intents, opinions, and evaluations of the seller's efficiency (Palmatier et al., 2006). According to Kumar and Shah, there are two types of client loyalty: behavioral loyalty and attitude loyalty (Kumar & Shah, 2004). The observable purchase behavior of customers is the foundation of the behavioral part of customer loyalty, which is based on the frequency of the exchanges, repeat purchase behavior, purchase sequence, and other characteristics of purchase behavior (Kumar & Shah, 2004). Confidence in the firm over time is the essence of attitude loyalty (Kumar & Shah, 2004). Those academics who focus on consumer happiness and product or service quality say that attitude loyalty is not always present at the same time as behavioral loyalty (Frank et al., 2014). This situation can happen because many other variables contribute to retention (Fornell, 1992; Matsuoka, 2022). Because consumers' spending habits are sometimes prompted by convenience rather than commitment, researchers of the hospitality branch stress the significance of differentiating between behavioral and attitude loyalty (Kandampully & Suhartanto, 2003). The attitude loyalty not only causes the repetitive use of services or purchase of goods, but it also, as mentioned in work by Zeithaml and colleagues in 1996, increases the chances of successful further recommendations, which brings in new customers who are so hard to acquire through other advertising practices (Hill 2006; Zeithaml et al., 1996). Also, the same loyal customer intentions were confirmed by Kumar and Shah (Kumar & Shah, 2004). The best technique to gauge consumer loyalty has yet to be discovered, but studies have shown that consecutive purchases and the number of word-of-mouth promotions are potentially reliable markers (Kandampully & Suhartanto, 2003; Lentz et al., 2021).

2.2 Satisfaction Theories, Models, and Paradigms

The models and theories considered before starting the research are the following:

> **Maslow's theory of human motivation** explains the basics of human motivations, which are crucial for the further development of the other satisfaction theories and models (Gawel, 1996).
>
> **Herzberg's motivation-hygiene theory** is the theory that divides the factors which theoretically should have motivational outcomes into two parts – hygiene factors and actual motivation factors (Gawel, 1996; Lundberg et al., 2009).

Kano model is the theory that professor Noriaki Kano provided in the 1980s (Mkpojiogu & Hashim, 2016). It is based on Herzberg's two-factor theory but tries not to repeat its mistakes, for which it was widely criticized (Matzler et al., 2004).

The disconfirmation paradigm is concerned about the expectations about a product or service (Cassidy-Smith, Baumann and Boudreaux, 2007). If the expectations are met, this leads to satisfaction (Cassidy-Smith, Baumann and Boudreaux, 2007). If the expectations are not met, this leads to dissatisfaction (Cassidy-Smith, Baumann and Boudreaux, 2007).

Both theories, a model and a paradigm, are further discussed in the next subchapter.

2.2.1.1 Maslow's Theory of human motivation

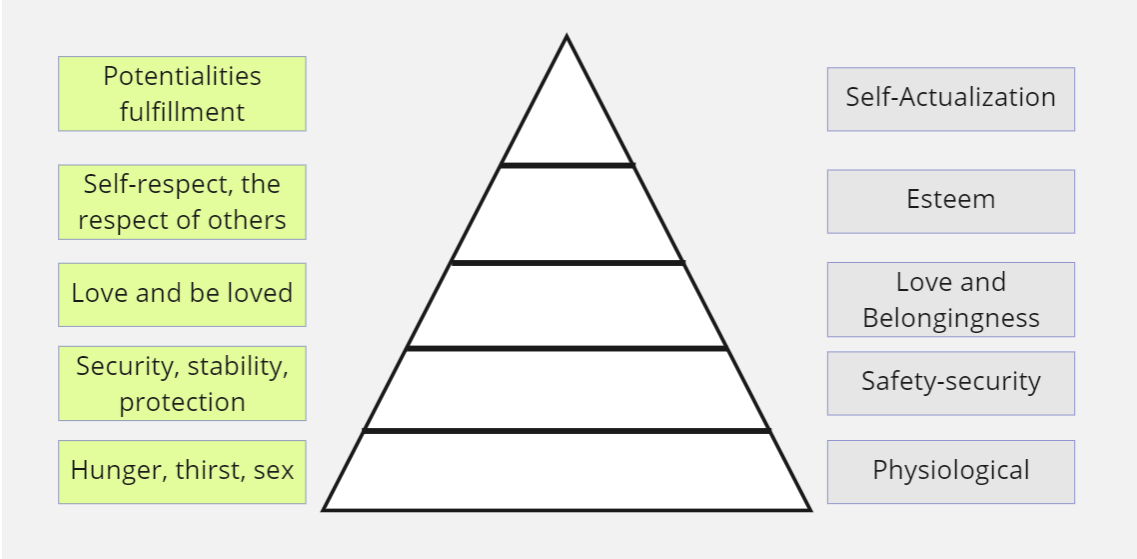


FIGURE 6: MASLOW'S HIERARCHY OF NEEDS (GAWEL, 1996; TAORMINA & GAO, 2013).

Maslow's theory of human motivation is based on his pyramid of hierarchy of needs (Figure 6), where Maslow states that people tend to satisfy their needs starting from physiological needs, followed by safety-security needs, followed by love and belongingness needs, after that esteem needs and finishing with self-actualization needs (Gawel, 1996). It is important to mention that the fulfillment of the different levels does not have to be finished to 100% to start fulfilling the next level (Maslow, 1954). According to the theory, it is expected that the lowest level will have the highest fulfillment rate and the highest level – the lowest fulfillment rate (Maslow, 1954). Arbitrary example: physiological - 90%, safety - 70%, belongingness - 50%, esteem - 30 %, self-actualization - 10% (Maslow, 1954).

2.2.1.2 Herzberg's Theory of Motivation

Herzberg's motivation-hygiene theory is a two-dimensional paradigm. One dimension is concerned with the factors whose absence excludes the possibility of satisfaction and can only increase dissatisfaction, and those factors are called hygiene factors (Gawel, 1996). Another dimension is looking into the motivation factors that might increase satisfaction if the hygiene factors are present at the level at which motivation factors can work (Gawel, 1996). According to Kurt Matzler and colleagues, Herzberg's two-factor theory received much criticism regarding its reproduction inability, its failure to include situational and interpersonal differences, and the simplicity of the model or its oversimplification (Matzler et al., 2004). It is a prevalent satisfaction theory, but given the parameters above, it might not be the best option to base the research on. The next model discussed, based on this theory, is a more progressive and reliable approach to making conclusions in this research.

2.2.1.3 Kano Model

Kano's Customer Satisfaction Model is a theory that compares pre-purchase expectations with actual feelings after the purchase is made (Oliver, 1980). Suppose the pre-purchase expectations were right on point. In that case, the customer is indifferent or moderately satisfied. If pre-purchase expectations are not met, the customer gets dissatisfied, and only if the pre-purchase expectations are lower than the actual product quality the client is satisfied (Matzler et al., 2004). The Kano model has three different aspects of quality, and the factors can dynamically switch from one factor to another throughout time. These are excitement, basic, and performance factors (Oliver, 1980). Basic factors are the musts – taken for granted or pre-required. Without those, satisfaction processes are not working, and customers can only get dissatisfied (Matzler et al., 2004). Performance aspects can satisfy and dissatisfy customers in case performance is good or bad (Matzler et al., 2004). Finally, excitement factors are unexpected, which should increase customer satisfaction if present but will not decrease it if not offered (Matzler et al., 2004).

2.2.1.4 Disconfirmation Paradigm

The Disconfirmation Paradigm states that people's impressions of customer experiences may be broken down into two categories: confirmation and disconfirmation (Cassidy-Smith et al., 2007). When a service is provided as expected, this means confirmation of the awaited experiences, leading to satisfaction (Cassidy-Smith et al., 2007). According to the disconfirmation paradigm, dissatisfaction results when customers' expectations are not satisfied (Cassidy-Smith et al., 2007). Disconfirmation occurs when current findings fall short of expectations (Cassidy-Smith et al., 2007). When the service is regarded as worse than anticipated, this is called negative disconfirmation, and the opposite is called positive disconfirmation (Cassidy-Smith et al., 2007). Maximum happiness is attained when the consumer feels a positive disconfirmation. (Cassidy-Smith et al., 2007)

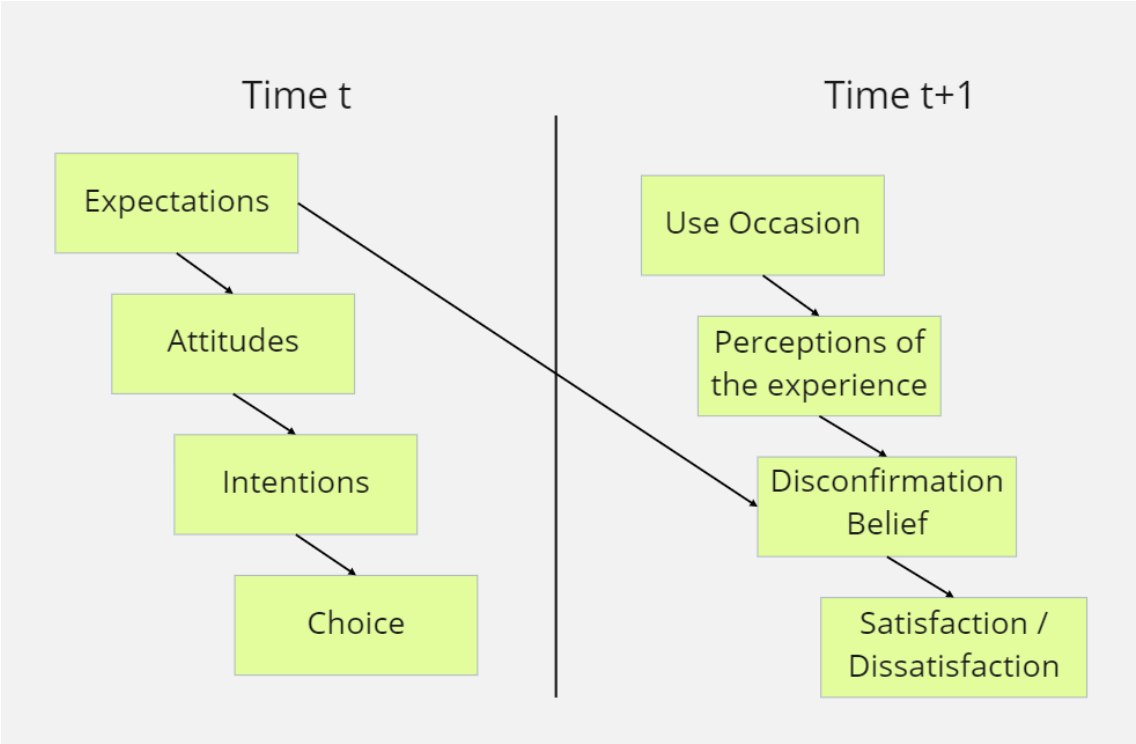


FIGURE 7: CONCEPTUAL MODEL OF DISCONFIRMATION OF EXPECTATIONS PROCESS (CADOTTE ET AL., 1987).

Cadotte and colleagues offered a conceptual model of the disconfirmation paradigm, which can be seen in Figure 7 (Cadotte et al., 1987). It is divided into two timeframes: "t," and "t+1." That gives a visual representation of how expectations lead to choice in "t" and how those exact expectations affect disconfirmation belief in "t+1", which later results in rather a dissatisfaction or satisfaction (Cadotte et al., 1987). In other words, the specific product or service is evaluated in timeframe "t," and expectations about its usage or purchase are formed. This leads to the creation of personal attitudes towards this good or service, which leads to intentions and, finally, to the choice of whether to buy it or not. After the moment in time "t+1," a transaction happens, and the client evaluates the experience of the product purchased. This experience and the expectations from the past lead to disconfirmation beliefs, potentially having three outcomes: better than expected, exactly as expected, and worse than expected. That, in turn, leads to high satisfaction, reasonable satisfaction, and dissatisfaction, respectively.

Mazanec, in his study from 2007, provided an excellent example of the use of the excitement, essential, and performance factors provided by Kano in the ski domain (Mazanec, 2007). His work researched the effects of the singular aspects of the alpine ski resorts in Austria on overall satisfaction (Mazanec, 2007). It concentrated on eight aspects of the 27 Austrian resorts and, as a result, highlighted three of them: gastronomy, slopes, and location, as the aspects affecting overall satisfaction the most (Mazanec, 2007).

2.3 Customer Segmentation

After explaining so much about customer satisfaction, it is important to mention that not all customers have the same expectations and that, for the best outcome, it is necessary to treat them differently (Formica & Uysal, 2001). Customer segmentation aids in the processes of identifying groups of similar customers and determining the best ways to work with them (Konu et al., 2011; Singh Minhas & Jacobs, 1996).

The purpose of market segmentation is to categorize customers into similar subsets based on some criterion or set of factors (Konu et al., 2011; Middleton & Clarke, 2012; Wedel & Kamakura, 2002). Tsiotsou identified four advantages of conducting customer segmentation in his work in 2006 (Tsiotsou, 2006). First, marketers may narrow their focus using the information gleaned from their created segments (Tsiotsou, 2006). Second, it aids in the creation of more targeted marketing strategies to better meet the requirements of a specific market niche (Tsiotsou, 2006). Next, the process of differentiating goods is simplified by segmentation, and last, market opportunities and risks may be better identified with the help of customer segmentation (Tsiotsou, 2006). The most common market segmentation methods include target audiences' demographic, geographical, behavioral, and psychological characteristics (Konu et al., 2011). Statistics readily available from government agencies make demographic and geographical parameters a standard option for segmentation (Dolnicar et al., 1999; Hudson, 2000; Konu et al., 2011; Yan et al., 2007). Since the introduction of GIS (Geographical Information Systems), which allows for handling a massive amount of data on households in particular locations, it has become more common to segment customers based on geo-demographic parameters (Musyoka et al., 2007). According to Tsiotsou, benefits, frequency of usage, and loyalty are only a few factors that may be used for behavioral segmentation (Tsiotsou, 2006). It is generally agreed that benefit segmentation is a subset of behavioral segmentation (Singh Minhas & Jacobs, 1996). It has been argued that this method helps create destination subsegments since it may be used to learn about the expectations and experiences of different types of travelers (Ahmed et al., 1997). Markets may be broken down using psychographic segmentation to focus on subsets of consumers with similar preferences but distinct personalities (Konu et al., 2011). According to Zografos and Allcroft, psychographic segmentation is a technique used to assess customers' values, preferences, and passions (Zografos & Allcroft, 2007). Those could be evaluated by looking at a person's religious and political values, behavioral traits, and the things people like doing in their spare time (Konu et al., 2011; Zografos & Allcroft, 2007).

Segmentation techniques may be classified as either "a priori" or "a posteriori" (Dolnicar & Leisch, 2003b; Formica & Uysal, 2001; Tsiotsou, 2006). When the criteria factor is already determined, “a priori” segmentation may be employed (Tsiotsou, 2006). In this context, demographic information, like age, and gender, can be used. Due to its reliance on empirical evidence, “a

posteriori" segmentation is used when no explicit knowledge of relevant categories exists (Formica & Uysal, 2001). To rephrase, "a posteriori" segmentation relies on specific data parameters (Tsiotsou, 2006).

2.3.1 Customer Segmentation in Ski Domain

There are various possibilities for customer segmentation in the ski domain. The following examples in Table 1 are only a small excerpt of the much wider variety. Studies have been conducted utilizing motives, value propositions, and features of destination choices as a foundation for segmentation (Jang et al., 2002; Park & Yoon, 2009). In their research, Mills and colleagues segmented customers into two groups based on how much their expenses are: vast and low expenses groups (Mills et al., 1986). Perdue segmented customers of one ski area by the frequency of visits of international and local or regional tourists (Perdue, 2004). Dolnicar and Leisch used psychographic and behavioral segmentation in their research about Austria (Dolnicar & Leisch, 2003a). They divided customers based on their travel motives and the activities they performed during this trip (Dolnicar & Leisch, 2003a). Tsiotsou used visit frequency to segment customers in his research (Tsiotsou, 2006). The research done by Konu and colleagues segmented customers by their destination choice (Konu et al., 2011). The work of Füller and Matzler analyzed the variations in happiness among lifestyle categories (Füller & Matzler, 2008). A short description of outcomes is presented in the next chapter, talking about similar studies.

2.4 Similar Studies

The following Table 1 represents eight similar studies which were segmenting the tourists traveling in different countries or regions. The customer segmentation column was discussed in the previous chapter. The first two columns identify that the samples were done from the winter ski resort visitors of such countries as Austria, Lapland (Finland), Switzerland, Italy, Japan, Texas (USA), Korea, Colorado (USA), and Greece (Konu et al., 2011). The final segments identified varied from two to six (Konu et al., 2011). The outcomes of Jang and colleagues divided the travel seekers into three groups novelty, family, and escape seekers, which have distinctively different views about traveling (Jang et al., 2002). Park and Yoon looked into motivation factors for tourism and distinguished four groups: passive travelers, travelers interested in learning something new, travelers wanting a piece of everything, and tourists aiming at the unity of a family (Park & Yoon, 2009). Another research done by Füller and Matzler resulted in grouping ski resort visitors into five types: family, not family, demanding, intellectual, and sporty (Füller & Matzler, 2008). The study by Dolnicar and Leish grouped the customers in two ways, based on their motive and based on their preferred pastime, which, when combined, resulted in 5 vacation styles (Dolnicar & Leisch, 2003a). First, one aimed at sledding and having a good time, the next group aimed at healing and calming themselves, and next, they differentiated between pure and moderate culture tourists (Dolnicar & Leisch, 2003a). Finally, they separated the group, loving snowboarding, and clubs (Dolnicar & Leisch, 2003a). Tsiotsou 2006 segmented customers by their visits,

whether they were weekly or monthly visitors (Tsiotsou, 2006). Konu with colleagues concluded the research by having customers of the ski resort in Finland segmented into six groups: want-it-all, cross-country, all-but-downhill, sporty, and relaxation seekers. The last group is the passive tourists (Konu et al., 2011).

Country or Region	Tourism Market	Number of Segments	Customer Segmentation by	Authors
Austria	Winter tourists in Austria	5	travel motives and the activities they performed during this trip	(Dolnicar & Leisch, 2003a)
Austria, Switzerland, Italy	Ski area customers	5	lifestyle	(Füller & Matzler, 2008)
Japan	Japanese outbound travel market	3	benefits	(Jang et al., 2002)
Texas, USA	Texans visiting skiing places	2	expenses	(Mills et al., 1986)
Korea	Tourists visiting rural tourism villages in Korea	4	motives, value propositions, and features of destination choices	(Park & Yoon, 2009)
Colorado, USA	Ski destination visitors	2	geographical, visit frequency	(Perdue, 2004)
Greece	Ski resort customers	2	visit frequency	(Tsiotsou, 2006)
Lapland, Finland	Ski resort visitors	6	destination choice	(Konu et al., 2011)

TABLE 1: SEGMENTATION STUDIES IN THE SKI DOMAIN (KONU ET AL., 2011).

2.4.1 Space for Potential Further Generalization of Outcomes

After analyzing similar articles and the yearly mountain tourism report from Vanat, it has to be said that the generalization of the outcomes of the research done in Austria is expected to be possible only to some degree (Vanat, 2020). Austria's neighboring countries, such as Switzerland and Italy, might benefit from using the outcome of this research more than other countries. However, they should still be aware that the results might differ significantly due to external parameters. Further generalization in other countries such as the United States of America, France, Japan, or Finland is encouraged only after similar research is conducted by local skiing areas with local customers.

After having the research done, other ideas are discussed in the conclusion chapter. The next chapter gives an overview of Austrian winter tourism and briefly touches on some delicate points regarding skiing tourism in Austria.

2.5 Austrian Winter Tourism

As mentioned by Vanat in his report from 2020, Austria is the only country with as many as 15 resorts that accumulate 1,000,000 tourist visits throughout the winter season (Vanat, 2020). Austria is followed by France, with 13 resorts, and Italy is having already only 7 (Vanat, 2020). Also, according to Vanat's report, Austria, Liechtenstein, and Switzerland are the three countries where more than 30% of the population participates in skiing (Vanat, 2020). The following countries on the list are Finland, Norway, and the Czech Republic, with 21 to 23% of the population interested in such activity (Vanat, 2020).

The next admirable milestone is that Austria is among the first three countries with about 3,000 lifts each (Vanat, 2020). In other countries, the United States of America and France, it is quite challenging to locate so many lifts with such density in Austria to compete with countries much larger in size (Vanat, 2020). This brings us to the following subchapter.

2.5.1 Tourism Spatial Distribution

Austria's most attractive tourist spots are highly geographically concentrated (Pröbstl-Haider et al., 2021). Some tourist hotspots see as many as one thousand visitors per capita annually (Prettenthaler & Formayer, 2011). Tyrol and Salzburg are home to several of Austria's most tourist-heavy, high-overnight-stay municipalities (Pröbstl-Haider et al., 2021). About a third of Austria's overnight stays in 2018 were in Tyrol, while 19.7% were in Salzburg (Statistics Austria, 2019). Upon closer inspection, significant discrepancies may also be seen at the local level (Price et al., 2011). Overcrowding, gentrification, and the disruption of social life, particularly in rural regions, are all possible outcomes in tourist hotspots during the season (Pröbstl-Haider et al., 2021).

2.5.2 Challenges

The winter tourism sector in Austria is facing some challenges. The ongoing crises include the COVID-19 pandemic, climate change, the Ukraine crisis, and the USA and China trade war (Murphy, 2022; Pröbstl-Haider et al., 2021; Wirtschaftskammer Österreich, 2022; Yang et al., 2022). While some of the mentioned crises have a direct impact, such as COVID-19 and climate change crisis, the other two, while not having direct contact with the industry, still hugely affect it indirectly (Murphy, 2022; Yang et al., 2022.) Some responsibilities of the winter tourism sector in Austria are also to mitigate the impacts caused by the number of CO2 and equivalent emissions caused (Pröbstl-Haider et al., 2021). Therefore, climate change for the alpine ski industry means high investments in snowmaking and the implementation of various eco-friendly practices (Pröbstl-Haider et al., 2021).

While the crises the Austrian tourism sector faces are not the direct interest of this research, it is essential to underline the unstable situation, which encourages precise and argumentative decisions on the ski business management (Pröbstl-Haider et al., 2021; Wirtschaftskammer Österreich, 2022).

2.5.2.1 Covid Crisis Outcomes

The Austrian Chamber of Commerce on their website provides data regarding the season 2020/2021. The first season affected by COVID-19 directly, and its impacts are enormous (Wirtschaftskammer Österreich, 2022). The cash turnover decreased by 90 % compared to the previous year. For the industry that provided 125,900 jobs, which according to the data from 2018, approximated 2.65% of the jobs in Austria, is quite a considerable difference (Steiger & Scott, 2020; Wirtschaftskammer Österreich, 2022). This information only describes one season. Nonetheless, the future seasons are expected to show the winter tourism businesses' rehabilitation; coming to the pre-covid years' rates will take some time (Pröbstl-Haider, Mostegl, et al., 2021).

2.5.2.2 Investments Changes in Winter Tourism Area

While the investment numbers for the previous decades and more narrowed examples of investments in season 2020/2021 (Figure 3) were already mentioned and discussed in the general background chapter, this short part going right after crisis outcomes underlines the fact that due to the crises the ski business becomes more and more volatile and as Steiger and Scott mentioned in their article the investors might seek other more profitable or less risky investment options in the future (Steiger & Scott, 2020).

2.6 Instruments

This part explains the techniques and instruments used in this research. The following subchapters explain what linear regression is and what multiple linear regression is and gives an overview of cluster analysis and its two main types, hierarchical and non-hierarchical cluster analysis techniques.

2.6.1 What is Linear Regression?

Predicting one variable's values from knowing the values of another is the purpose of linear regression analysis (IBM, 2022). There are two types of factors independent and dependent. With the help of an independent variable, researchers attempt to predict the dependent one (IBM, 2022). One or several predictive variables are used to make a formula that allows forecasting as close as possible to the actual data (IBM, 2022). The most important aspect of linear regression is the fact that it is considered an easy and reliable method for forecasting in different areas such as business, social studies, and environmental studies. Because of this, this model

has become popular in academic and business studies (IBM, 2022). Multiple linear regression (MLR) is a more powerful version of simple linear regression, where there are multiple instead of one predictor variable. Marta Flores-Sosa, in her work, mentioned that multiple linear regression is among the most popular instruments in academic research (Flores-Sosa et al., 2022). When conducting multiple regression, multiple equations are created to explore if there is a connection between independent and dependent parameters (Flores-Sosa et al., 2022). Multiple regression, in other words, is an extended version of the linear regression technique (Flores-Sosa et al., 2022).

2.6.2 What is Cluster Analysis?

According to Everitt and colleagues, the term cluster analysis is seen nowadays as a method that probes data for clusters (Everitt et al., 2011). However, nonetheless, it is important not to forget that there is a possibility that cluster analysis can return a researcher with an outcome where no clustering is justified. Therefore no strikingly different groups are detected (Everitt et al., 2011). International Business Machines Corporation (IBM), the company which created and still develops statistical software SPSS in their database, describes cluster analysis as an exploratory instrument for natural “a posteriori” grouping (Everitt et al., 2011; IBM, 2021). There are two ways to conduct cluster analysis, often referred to as hierarchical clustering and non-hierarchical clustering (Kassambara, 2017). Both are further discussed in the next subchapters.

2.6.2.1 Hierarchical clustering

In the first step of hierarchical cluster analysis, individual objects are placed into their own groups (IBM, 2021b). Entities are united into a single tree structure by relaxing the criteria by which they are divided at each level of the assessment, starting with the initial separation of the objects by combining the two most identical ones, and continuing the process until the tree is complete (IBM, 2021b). The exact process can be backward described as all variables being in one group and then divided until the point where each object is in a single group. The range is considered a primary factor used to form groups. In general, it makes sense for nearby things to be grouped together in the same cluster (IBM, 2021b). In contrast, objects separated by a considerable distance should be placed in separate clusters (IBM, 2021b). The groups formed from a particular dataset are determined by the values supplied for the three main settings: method of clustering, measure, and standardization (IBM, 2021b). The principles for cluster creation are laid forth by the chosen method of clustering. There are three approaches to measuring how far the items are from each other (IBM, 2021b). First, the nearest objects can be paired and made sure they are in one segment, or the furthest away items can be identified and made sure they are for sure in different segments (IBM, 2021b). Alternatively, a mix of those methods can be used to create the third one. The measure setting is set in place to provide the formula to calculate the distance between items (IBM, 2021b). The Euclidean measure uses a straight line, the

binary expects only two values, and the interval measure looks for a particular scale. The assumption of continuous numeric values in count measurements leads to choosing a count measure (IBM, 2021b). And the last one is standardization, which is used to organize and make variables with different scales comparable and, therefore, useful (IBM, 2021b). The clustering procedure in this research is done using Euclidean distance and Ward's linkage method. Among the literature researched for this study, the mentioned distance and linkage methods are most popular and considered the ones with the most reliable data outcomes.

2.6.2.2 Non-Hierarchical Clustering

K-means cluster analysis is a method for categorizing data based on a collection of factors, with the goal of classifying instances into a definite number of clusters whose features are unknown at this time (IBM, 2021a). It appears to be more useful in situations when a massive number of instances need to be sorted, and the number of groups is known beforehand. The excellent work of the cluster analysis is achieved through a minimal number of clusters while still providing relevant information regarding the researched subjects (IBM, 2021a). When looking for the right cluster analysis approach in this research, this method was not chosen because there is no "a-priori" knowledge about the number of clusters. And the dendrogram in the hierarchical clustering approach was considered more informative and valuable for the selection of the number of clusters.

2.7 Software Usage

R is considered the most useful statistical environment for this research, and the RStudio interface is chosen for its user-friendly interface. Both software products are presented in the following subchapters. Also, the complimentary software used is mentioned as a separate subchapter.

2.7.1 R as a Programming Language for Statistical Computing

R provides a free and user-friendly environment for those interested in doing analysis and creating visual representations of statistical data. It can be compiled and executed on many UNIX systems (UNIX abbreviation comes from UNICS – UNiplexed Information Computing System) and computers from Apple and Windows systems from Microsoft (R Core Team, 2019; *UNIX Full Form*, 2020).

R is both a programming language and an environment for computationally and graphically intensive statistical work (The R Foundation, 2019). It is very similar to another programming language called S, to the point that majority of commands written for S are executed without additional edits in R (The R Foundation, 2019). Nonetheless, some critical differences exist, and R is not absolutely backward compatible with S (The R Foundation, 2019). The R has a library of sta-

tistical instruments for different types of modeling, analysis, and tests with possibilities for variable visualizations of results (The R Foundation, 2019). Also, R has an open-source platform for other developers, allowing them to enhance the standard functions by creating their own additional libraries (The R Foundation, 2019). The extraordinary General Public License of R is provided for free in the form of the source code and, as mentioned above, is operatable with a variety of operating systems.

The developers insist on R being an environment because it is a product that covers all researcher's needs instead of only providing a library of statistical tools.

2.7.2 RStudio as an Integrated Development Environment for R

RStudio, or as it is often called RStudio IDE (Integrated Development Environment), is an even more advanced environment that is used with such programming languages as R and Python (RStudio, 2019). While having a free version of the product RStudio team also develops a commercial edition which allows for easy cooperation of researchers on the projects. Including the ability to execute commands online through anything having internet access. The system which J. J. Alaire, the CEO of RStudio, describes in his foreword to the annual report for 2021 is reminding of the circular motivation system (RStudio, 2021). The team develops open-source software that creates an interest in commercial users who pay to use it (RStudio, 2021). That money is then used to develop open-source software, which is a base for commercial software, which creates even more commercial interest in the product, and so on and so forth (RStudio, 2021). The RStudio team insists that they work for the creation of a solid open-source product that will be available to the broad community (RStudio, 2021). They hope it will unite existing researchers and introduce new people to statistical research through the use of their product (RStudio, 2021). Millions of individuals across the world already utilize RStudio's open-source software. Furthermore, thousands of people and companies decided to pay for the commercial product.

Therefore, in the context of this research, R is the core of it, and the RStudio environment is used mainly because of how easy and fast it is to implement the needed techniques within the environment. The next chapter focuses more on the complimentary software and also elaborates on the R libraries used.

2.7.3 Complimentary Tools

For the creation of figures other than the output of the RStudio, the Miro is used. It is an online platform that allows people alone or in groups to create nice notes and figures similar to the real whiteboard (Miro, 2022).

For some data organizing, easier changes of the variable's names, and for the creation of tables, Microsoft Excel is used.

Also, a variety of additional R packages are used for the purposes of statistical research and better visualization. Those are briefly mentioned in the next subchapter.

2.7.3.1 R Packages Used

For the purposes of this research, the libraries “haven,” “dplyr,” “ggplot2,” “AMR,” “labelled,” “sjplot,” “arsenal,” “cluster,” “dendextend,” and “factoextra” are used. Library “haven” is used to read the data from the “.sav” file (Wickham et al., 2019). Library “dplyr” is a library necessary for additional data manipulations (Wickham et al., 2020). Library “ggplot2” is used for visualization purposes (Wickham, Chang, et al., 2019). Library “AMR” allows grouping variables such as age and transforming them from numerical form to groups (Berends et al., 2022). Library “labelled” is used to quickly change labels (Larmarange et al., 2022). Library “sjplot” is used to visualize the semantic differential scale (Lüdecke, 2021). Library “arsenal” is used for the creation of the summary of the descriptive statistics (Heinzen et al., 2021). Library “cluster” is used for clustering purposes (Maechler et al., 2019), and library “dendextend” is used for the external functionality of the dendrograms in R (Galili et al., 2021). Library “factoextra” helps visualize multivariate data analysis (Kassambara & Mundt, 2020).

2.8 Customer Persona as a Design Tool

The multidimensional differences between clusters of customers may be better interpreted using a customer persona tool. Many different types of experts may participate in service design thanks to design tools like client personas, which are widely utilized in the procedure of refining and creating new goods and services (Broberg et al., 2011; Mackrill et al., 2017). Stereotypical users' behaviors, wants, motivations, qualities, and constraints may be better understood with the help of customer personas (Cooper, 1999; Goodwin, 2009). The development team's ability to empathize with actual users is greatly enhanced by creating at least a small number of personas (Bradley et al., 2021). When it comes to prioritizing the designing solutions around the demands of a particular user group, personas may be a useful, practical tool (Miaskiewicz & Kozar, 2011). The final suggested offer should attract the most significant number of buyers (Cooper, 1999). Therefore, developers must provide room for enough personalization so the good or service may expand and morph to fit the buyer's requirements (Cooper, 1999).

After the customers are clustered, this tool should help differentiate customers of the different groups if applicable.

2.9 Conceptual Framework

This research is developed in such a manner that it can be easily repeated with data from other years, locations, or both. Therefore, the conceptual framework in Figure 8 is an example of a complete cycle from customers’ feedback to investments and advertisements impacts. Figure 8

represents the development of the variables with an influence of the theories, models, and paradigms discussed in the literature review on the factors on the left side and the use of different statistical tools with those variables in the center and on the right side. For easier understanding, different types of components such as impacts, variables, theories, comments, and tools have their own color identified in the right part of the figure. Also, on the right, the four types of arrows are explained: the usual line arrow represents the development or transformation of the variable or influence of the theory or the model. The dotted arrow leads to comments. The arrow with the word “Predicts” stays between independent and dependent variables. The dashed line represents the long-term impact. The shortening IND.V and D.V. stay in independent and dependent variables boxes, respectively.

From looking at Figure 8, it can be seen that the disconfirmation paradigm influences the transformation of the variables from the customers’ emotional feedback regarding their expectations to the customers’ satisfaction factors. Then those factors are influenced by Kano's three-factor model, which leads to the creation of the three groups of variables: Basic factors, Performance Factors, and Excitement factors. Maslow’s theory of human motivation, with its pyramid of needs, explains the fundamental factors not considered in this research. The primary focus is on the performance and excitement factors influencing customer satisfaction. This influence and prediction possibilities are then tested with the help of multiple linear regression. The outcomes of the multiple linear regression are the factors that significantly influence overall satisfaction. Those variables and respondents are further “a posteriori” grouped in a data-driven manner through cluster analysis. The outcome of this research has the potential to influence investment and advertisement decisions, which in turn later should affect the customers’ feedback and, over time, change the system, which requires the repetition of the study with the new data in the future.

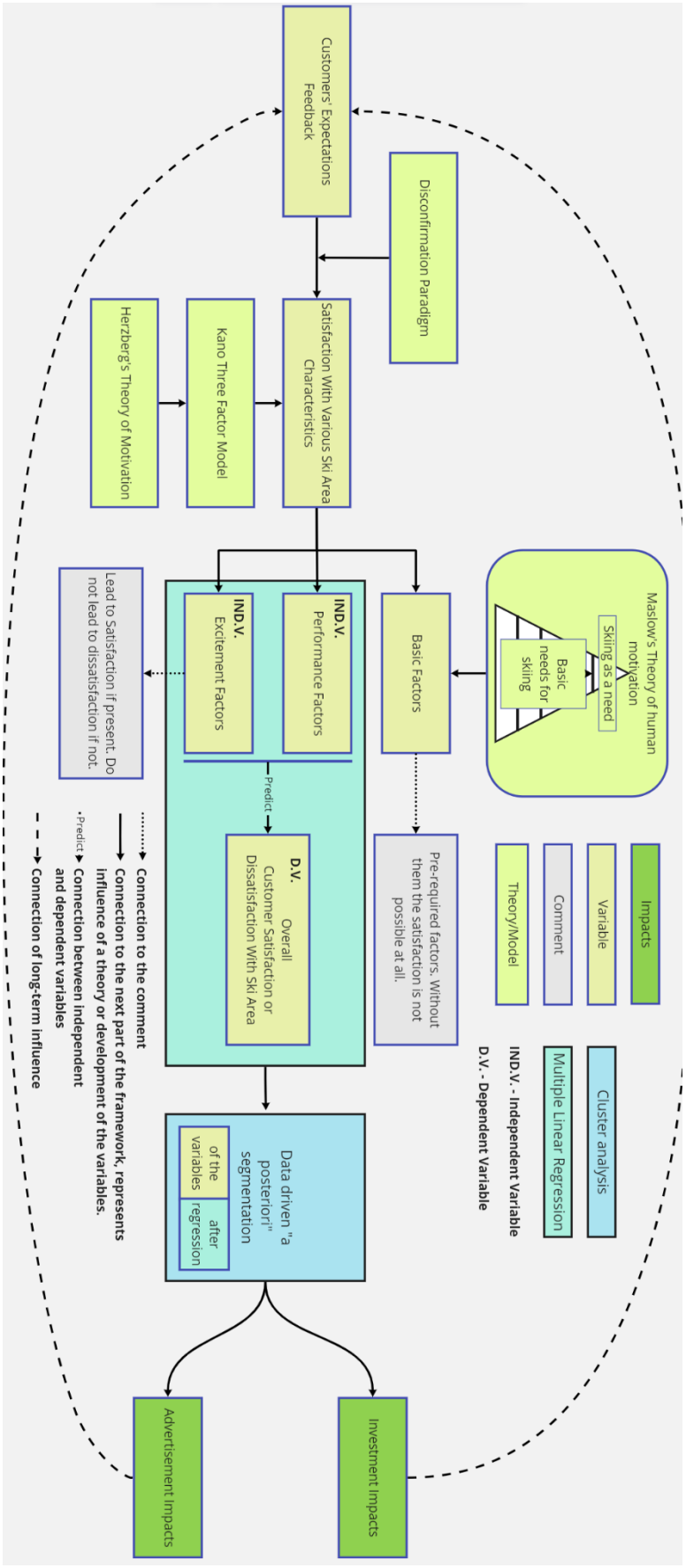


FIGURE 8: CONCEPTUAL FRAMEWORK.

2.10 Hypothesis Development

After the background and the tools intended for use are provided, it is time to present the hypotheses. In order to identify the most influential characteristics, the author should question the equality of importance of alpine ski resort characteristics. Therefore, the author suggests that they are not all equally predicting the customers' overall satisfaction with their holiday. And also, not all of the characteristics are useful for such prediction. This process leads to the development of the following hypothesis:

H0: There is no characteristic that can predict the overall satisfaction of the customer of the alpine skiing area in Austria.

H1. A specific characteristic of the alpine skiing area in Austria can predict the customer's overall satisfaction with the skiing area.

Instead of the words “specific characteristic” in H1, there should be actual variables from the data. The dataset has 47 useful satisfaction variables apart from the one which is overall satisfaction. The hypothesis is provided in this form to omit the unnecessary cumbersome presentation of the H1-H47.

Among the tested variables are snow conditions, space on the slopes, gastronomy overall, gastronomy space, offers for kids, ski schools, and others, better presented in the “descriptive statistics of the sample” chapter.

The second part of the research presents cluster analysis, and this technique does not test for significant differences; therefore, there are no hypotheses tests in the second part. The second part aims to group the respondents and characteristics based on customers’ answers using cluster analysis as a desired “a posteriori” segmentation technique.

The literature review chapter explained the main aspects of this research, provided the theoretical background, developed the main research question and aim of the research, discussed instruments and ways of their implementation, provided an outline of relevant theories, models, and paradigms, made the reader aware of the winter tourism situation in Austria, reviewed the similar studies from different parts of the world and provided an overview on the topic of customer satisfaction and customer segmentation. The author tried to write about all aspects necessary for the reader to understand this research without requiring additional prior research.

The following chapter elaborates more on the topic of methodology.

3 METHODOLOGY

The methodology chapter can be divided into three parts. It first provides some general methodology-related information overview and then, based on this information, outlines the choices done for this research. The second part focuses on the data, its source and reliability, limitations and ethical considerations, and the research instrument used for data collection. Data analysis is the third and final part, not counting the conclusion. The chapter conclusion adds everything and prepares the reader for the following chapter.

3.1 Research Approaches, Designs, Methods, and Worldviews

This subchapter outlines the methods used when deciding which research approach is appropriate. Based on this chapter, the following explains the decisions explicitly made for this research.

Research approaches are strategies and processes for a study that cover the phases from basic ideas to precise techniques of data collection, processing, and evaluation (Creswell & Creswell, 2018). These procedures may also be called methodologies (Creswell & Creswell, 2018). A variety of factors influence the choice of the research approach. The researchers' prior knowledge, the topic or issue being studied, and the intended recipients of the study all have a role in the methodology used (Creswell & Creswell, 2018). Which method should be employed to investigate a subject as a whole constitutes the ultimate choice. Theoretical and philosophical assumptions of the author of the study, the chosen design of the research, and the tools used for the conduction of the study, including the data gathering and evaluation, all rely on the three main capstones: research approach, research design and research method (Creswell & Creswell, 2018).

The choice of the right approach is highly influenced by three factors: worldview, design, and research methods, as shown in a framework created by Creswell & Creswell in Figure 9 (Creswell & Creswell, 2018). Also, the framework identifies the sub-variables inside each factor.

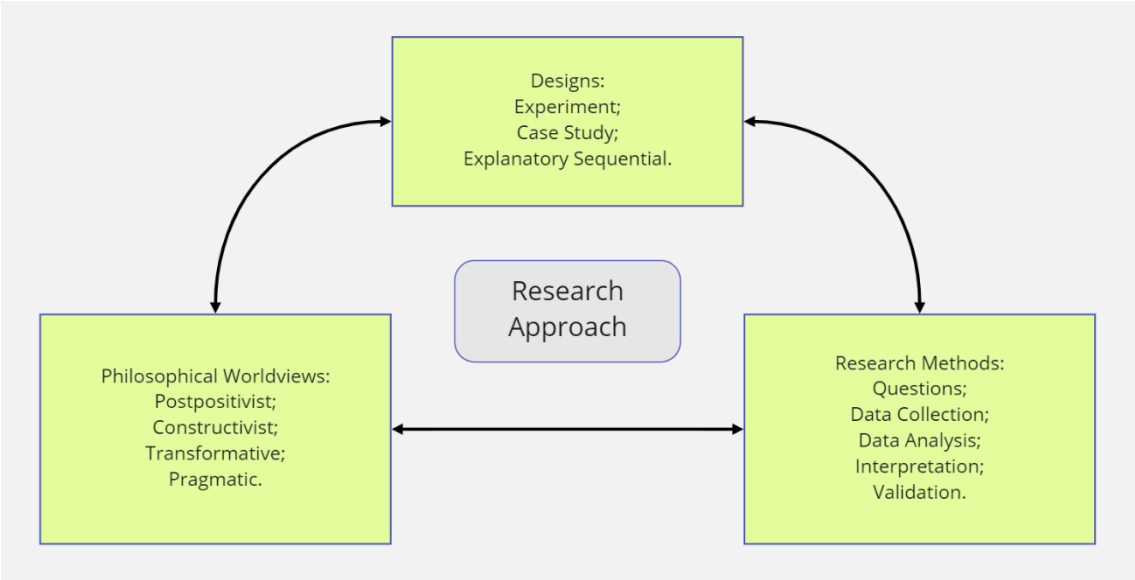


FIGURE 9: A FRAMEWORK FOR RESEARCH—THE INTERCONNECTION OF WORLDVIEWS, DESIGN, AND RESEARCH METHODS (CRESWELL & CRESWELL, 2018).

3.1.1 Research Approaches

There are three research approach options: quantitative, qualitative, and mixed (Creswell & Creswell, 2018).

3.1.1.1 Quantitative Research Approach

Studying the interplay of several elements, quantitative research is a method for putting theoretical hypotheses to the test in an unbiased manner (Creswell & Creswell, 2018). These factors data are gathered through the use of specific methods, allowing then quantitative data to be examined by statistical methods (Creswell & Creswell, 2018). The study's complete written form is bound to a predetermined format that includes an introductory part, literature and theory review, methodology, findings, and commentary (Creswell & Creswell, 2018). For quantitative research, assumptions about hypotheses testing are made with the use of deductive logic (Creswell & Creswell, 2018). These also include precautions against bias, putting controls for substitutes or other hypothetical explanations, and having the ability to generalize and repeat the results just as they are (Creswell & Creswell, 2018).

3.1.1.2 Qualitative Research Approach

Qualitative research seeks to learn how different people and communities interpret and react to a social or human issue (Creswell & Creswell, 2018). The emergence of questions and techniques characterizes the study, the collection of data in the context of interest, the inductive logic, and the development of analysis from specifics to overarching subjects, and the results are bound to the scientist's assumptions to conclude the meaning of the researched subjects

(Creswell & Creswell, 2018). A malleable outline is used for the final version of the study (Creswell & Creswell, 2018). Those that partake in this line of questioning advocate for a research methodology that respects inductive reasoning, emphasizes the significance of describing a problem's diversity, and prioritizes expressing personal opinions and interpretations (Creswell & Creswell, 2018).

3.1.1.3 Mixed Research Approach

This type of research requires the collection of both qualitative and quantitative data, which sometimes is required to fit the developed framework, the worldview assumptions, or due to the study specifications (Creswell & Creswell, 2018). Basically, the premise of mixed research is that one can learn more by combining qualitative and quantitative data than possible through the use of qualitative or quantitative (Creswell & Creswell, 2018).

For this research, the quantitative method is preferred. The unbiased manner of the hypothesis testing and the practicality of the quantitative data-gathering methods allow this research to look for causalities between characteristics compared. Which is the main reason this research started. The alternative approach, which is the qualitative approach, was considered, but it was concluded that it could be of use only as a further development after the quantitative research is done.

3.1.2 Philosophical Worldviews

Philosophical worldviews have an enormous impact on research (Creswell & Creswell, 2018; Slife & Williams, 1995). While the prior identification of the author's worldview helps the reader understand the choice of research approach, it is a common situation when authors keep the philosophical ideas unspecified (Slife & Williams, 1995). While there is a much greater variety of beliefs available, the author sticks to the most common ones used in research literature according to Creswell and Creswell: postpositivist, constructivist, transformative, and pragmatic worldviews (Creswell & Creswell, 2018). Their central beliefs and ideas are presented in Table 2.

Postpositivism	Constructivism
Determination	Understanding
Reductionism	Multiple participant meaning
Empirical observation and measurement	Social and historical construction
Theory verification	Theory generation
Transformative	**Pragmatism**
Political	Consequences of actions
Power and justice-oriented	Problem-centered
Collaborative	Pluralistic
Change-oriented	Real-world practice-oriented

TABLE 2: FOUR WORLDVIEWS (CRESWELL & CRESWELL, 2018).

3.1.2.1 Postpositivism

Postpositivism is a decisive research theory that holds that some causes almost certainly influence effects and outcomes. Therefore, postpositivists' concerns mirror those of situations where it is necessary to isolate and weigh the factors that contribute to the results, as is the case in experimental settings (Creswell & Creswell, 2018). This worldview is very appealing for the conduction of quantitative research (Creswell & Creswell, 2018).

3.1.2.2 Constructivism

Constructivism which also can be referred to as interpretivism is another worldview to consider. According to Creswell & Creswell, the roots of constructivism were established in such works as "*The Social Construction of Reality,"* written by Luckmann and Berger in 1967, and "*Naturalistic Inquiry*" written by Lincoln and Guba in 1985 (Berger & Luckmann, 1967; Creswell & Creswell, 2018; Lincoln & Guba, 1985). Their ideas were further adjusted by other authors nowadays, but they all hold assumptions that constructivism is evolving around the complexity of views each individual has about a particular thing or event (Creswell & Creswell, 2018). This worldview is appealing to qualitative research more than to quantitative, as the questions tend to be open-ended and wide, and the whole idea of the worldview is not to narrow down any questions as in quantitative research (Creswell & Creswell, 2018).

3.1.2.3 Transformative Worldview

With a Transformative Worldview, researchers take a stand for marginalized groups by choosing a specific qualitative framework (e.g., indigenous peoples, women, racial groups, and people with disabilities, among others) and using that framework to promote social change and equality (Mertens, 2010).

3.1.2.4 Pragmatic Worldview

The pragmatist worldview or philosophical stance is one that is the result of what has happened instead of a predetermined set of assumptions (as in postpositivism). It is essential to focus on practical uses and effective approaches to issues. Researchers put more weight on the study subject and use many strategies to get to the bottom of it (Creswell & Creswell, 2018). This perspective is good for the mixed-method approach due to its flexibility it can work with qualitative and quantitative approaches without contradicting itself (Creswell & Creswell, 2018).

According to what is written above, the postpositivist worldview is the one the author thinks is the best for conducting this research. As this method uses the quantitative approach, all the postpositivism ideas and notions perfectly comply with this research's aims. Alternatively, the author considered the pragmatic worldview due to its flexibility, but as this is not mixed method research and the strict ideas of pragmatism comply perfectly with the aims of the study, the postpositivist worldview is favored.

3.1.3 Research Designs

Quantitative Designs	Qualitative Designs	Mixed Method Designs
• Experimental designs • Nonexperimental designs, such as surveys • Longitudinal designs	• Narrative research • Phenomenology • Grounded theory • Ethnographies • Case Study	• Convergent • Explanatory sequential • Exploratory sequential • Complex designs

TABLE 3: RESEARCH DESIGNS (CRESWELL & CRESWELL, 2018).

After discussing the worldviews, the different types of designs are discussed. The research designs can be harshly divided into types according to the research approach used: qualitative, quantitative, and mixed ones, as can be seen in Table 3 (Creswell & Creswell, 2018). Nonetheless, those three groups shape the types of designs inside them, and the actual number of existing designs got very big (Creswell & Creswell, 2018). Especially in the last two decades with the popularization of open-source statistical software. Creswell and Creswell, in their book, provide several examples, which can also be seen in Table 3 (Creswell & Creswell, 2018). Those are experimental, non-experimental, longitudinal designs for the quantitative group, narrative research, phenomenology, grounded theory, ethnographies, and case study for the qualitative designs group (Creswell & Creswell, 2018). Convergent, explanatory sequential, exploratory sequential, and complex designs form mixed method designs (Creswell & Creswell, 2018). When the researcher chooses one of the mentioned designs for a study, the design sets some requirements and, in some cases, also provides an outline or a list of actions that are supposed to be done by the researcher to successfully conduct the study (Creswell & Creswell, 2018). This knowledge and algorithms are based on previous research; therefore, these structures are more trustworthy (Creswell & Creswell, 2018). Nonetheless, it is up to the researcher how to shape

the research, and while the implementation of some new steps in the design will definitely require more time to explain the theoretical background and reliability, it is sometimes required for the needs of the research (Creswell & Creswell, 2018).

After describing the research design's core ideas and mentioning the quantitative approach and postpositivism worldview as the main underlying aspects of this research, the research design for this research is one of the quantitative designs, specifically the nonexperimental survey design.

3.1.4 Research Methods

The research methods are explicitly the ways of obtaining and analyzing the needed research information (Creswell & Creswell, 2018). Those ways highly relate to the chosen research design.

For example, qualitative methods of obtaining the information are pre-determined, use close-ended questions, and are usually concerned about performance, attitude, census, or observational data (Creswell & Creswell, 2018). This data is then analyzed in statistical ways, and the interpretations are based on a statistical manner (Creswell & Creswell, 2018). On the other hand, qualitative methods are practically the very opposite. Instead of the pre-determined techniques, they require new and emerging methods, instead of closed questions, open-ended questions are used, and the data gathered is usually audiovisual, such as interviews or observation data (Creswell & Creswell, 2018). The analysis includes text, audio, video, and image analysis, and the patterns and themes are being interpreted as a result. The mixed research methods include both of the mentioned above and only their sequence changes according to the needs of the researchers (Creswell & Creswell, 2018).

This paragraph summarizes the core ideas forming the foundation of this research. This research uses the quantitative approach as research is going to work with a bulk numerical dataset. The study maintains the set of postpositivist worldview believers because it is tightly linked to the cause-effect relationship. It uses the non-experimental fixed strategy, more specifically survey research, because it is one of the best ways to gather data for a quantitative research approach. As already mentioned above, the alternatives were considered, but the parameters mentioned were the perfect fit, and therefore the choices regarding the fundamental aspects of this research are final.

3.2 Data Acquirement, Source, and Reliability

While usually, this part of the research should explain the type of sampling methods used and develop the questionnaire, in this research, this process is substituted with the description of the dataset and survey, which the author got from, a firm that specifies in field of market research and professionally provides such services as monitoring, benchmarking, quality management and other services concerned about transforming data into knowledge.

Nonetheless, this dataset is technically secondary data because it was collected not specifically for the purposes of this study but for the purposes of conducting similar studies and was developed in a shape that accommodates the needs of this research perfectly. Therefore, the straightforward alternative approach, which would include the development of the questionnaire, approval of the questionnaire by the institutional research board, development of the connections with ski areas, and finally, conduction of the data gathering, was not undertaken for a couple of reasons. First, the amount of time and resources available for the researcher is not enough to gather as much data. The time needed to establish a working relationship with each ski area to provide the dataset equal in the number of participants potentially takes a couple of years. Also, the provision of the motivation to participate to the customer is then solely of the researcher. This leads to quite a big required budget for conducting such research, and that is in case if the personnel of the ski areas cooperates and asks customers to participate in this survey. The second reason is the COVID-19 pandemic's influence. The researched businesses suffered a lot from it, so the data available from the pandemic-influenced years might be unreliable. The third reason is that ski resorts are unlikely to cooperate with a researcher when they already have similar data gathered through other sources, such as a firm that provided the dataset. They are more likely to deny data gathering and not burden customers with a variety of similar questionnaires.

3.3 Relevant Survey Parameters

The original survey example is partly provided in Appendix 1 (some of the parts not relevant to the research were blurred to omit the easiness of reproduction of work done by the providers of the data). The survey consisted of 29 questions of various types, including multiple choice with one possible answer, multiple choice with more than one possible answer, a couple of open-ended questions, Likert scaled questions asking clients to show whether they approve or not some of the aspects, and ten multi-item satisfaction oriented questions with the semantic differential scale with around 50 characteristics total included in them. The survey asked the customer for permission to use their answers for the conduction of such experiments. The researcher did not receive any private information about ski businesses or their customers.

The population of the study is all the customers of the ski area. The sample consists of more than 48 thousand respondents, but it still has to be cleaned. The sampling design has a single-stage design.

Participation in the survey was allowed to the ski area visitors older than 13 years old (starting from 14) and as a motivation to participate – each person who submitted it had a chance to win some prize. Therefore, it was purposive non-probability sampling. Random sampling would not be the best idea here, as with random sampling, it might occur that participants did not experience the service, and therefore purposeful sampling is chosen to omit the decrease in participation rates and potentially biased data.

It was mentioned that the data source company had gathered such data for more than a decade. Therefore, the previous years can be seen as a pilot test for the dataset used in this research now. This also describes this survey as a longitudinal one.

The variables which are studied in this research are semantic differential scaled with a range of 1 to 6, and with the use of fascial expressions. Where one corresponds to a happy smiley face and means “extremely inspiring,” and six corresponds to an unhappy smiley face and means “rather disappointing.”

In order to convert the customer’s semantic description to the satisfaction variable, the disconfirmation paradigm is then used to assign the markings extremely satisfied to 1 and highly dissatisfied to 6.

3.3.1 Question Examples

1. Ski resort characteristics:

	1	2	3	4	5	6
General Ski Resort	😊	🙂	😐	😑	😕	☹
Gastronomy Atmosphere	😊	🙂	😐	😑	😕	☹
Gastronomy general	😊	🙂	😐	😑	😕	☹
"General Atmosphere Winter Experience"	😊	🙂	😐	😑	😕	☹
General Location	😊	🙂	😐	😑	😕	☹
General Accommodation	😊	🙂	😐	😑	😕	☹
Ski Resort Panorama	😊	🙂	😐	😑	😕	☹
Apres Ski	😊	🙂	😐	😑	😕	☹
Services Ski Schools	😊	🙂	😐	😑	😕	☹
"Gastronomy Value for Money"	😊	🙂	😐	😑	😕	☹
General Hiking	😊	🙂	😐	😑	😕	☹
General Slopes	😊	🙂	😐	😑	😕	☹
Lift Access	😊	🙂	😐	😑	😕	☹
General Children Offers	😊	🙂	😐	😑	😕	☹
Slopes Space	😊	🙂	😐	😑	😕	☹
General Rodelbahn	😊	🙂	😐	😑	😕	☹

Slopes Safety 

2. What is your education level?
   - Lower secondary education
   - Upper secondary education
   - Tertiary education

3. How old are you? ___

4. Gender:
   - Male
   - Female
   - Other

3.4 Limitations

The data gathered only in a variety of Austrian ski areas might still mean there are other resorts that do not work the same way as the others. Considering the similarities, this is very unlikely but has to be underlined. Also, the fact that only Austrian ski areas were used as a source of data limits the generalization possibilities of this study for other countries. The dataset represents data from the season 2019/2020. Therefore, the data is limited to only this season. While it is more than enough for the purposes of this research, it does not necessarily represent the situation right now. Another constraint is that only guests over the age of 14 were allowed to participate in the survey.

Another limitation can appear through the use of multiple linear regression. While it still finds the most influential characteristics for the majority of customers, some minorities might be missed. For example, customers traveling by public transportation or customers traveling with pets might have different preferences, but because of the very small amounts of those, their opinions will be left out during the linear regression, which means the alpine ski resorts covering this niche might be misguided by the results of this research and may need to conduct their own research. Also, there is a slight chance that there are characteristics that influence customer satisfaction and are not yet mentioned in the survey, which might limit the reliability of the study. The nonprobability sampling used in this research also limits its results because it can be assumed that those people who did not have time to participate or just did not want to or liked to for any reason to participate might have different opinions, which might affect the results significantly. Another part of the research that limits its reliability might be the violation of the linear regression assumptions, with the scale being ordinal in this research.

This research was based on secondary data. While it is already mentioned that it was collected for similar research purposes, it still limits the author's ability to add other variables or remove unnecessary ones.

3.5 Ethical Considerations

While dealing with a sample consisting of more than 48 thousand participants, it is vital to have strict research ethics. The author got data without any personal information of the participant as well as no identification of the ski area which was rated. The information complimenting the satisfaction variables included them being the data from the season 2019/2020 and that the ski resorts are from Austria. This ensures private information safety for the ski resorts and customers. Also, it is essential to mention that participation in the survey was voluntary. While customer satisfaction research benefits both the customer and the company, participation in the lottery to win some special price was offered to the participants. During the analysis of the data, there was no advantage given to any variables by the author. While all the information necessary for the study was mentioned in the whole volume, some aspects not used in the research were not mentioned on purpose, as it makes the data description cumbersome while unnecessarily oversharing the intellectual property of the data provider. The use of the data provided by the customers is explained and identified to them before participation. All questions in the survey are in clear, straightforward, and appropriate language. All the esteemed authors' works that inspired different parts of this thesis are appropriately cited. There is no conflict of interest between the author and the topic or participants of the research.

3.6 Data Analysis

To be able to run the multiple linear regression and other analyses, the dataset first had to be cleaned. For these purposes, the rows and columns with more than 50% of the data missing are excluded from the research. After this, the software allowed to perform multiple linear regression-related commands. After cleaning, the 48 semantic differential variables were left.

Also, while speaking about the satisfaction variables author is referring to the variables on a semantic differential scale from 1 to 6 where 1 is “extremely inspiring,” and 6 is “rather disappointing,” that in this case is interpreted as a scale from 1 to 6 where 1 is extremely satisfied, and 6 is highly dissatisfied according to disconfirmation paradigm.

This part is further divided into three subparts that correspond with the three phases of the research. Descriptive analysis was done first to explore the dataset, followed by multiple linear regression to limit the number of the variables only to ones with significant influence or, in other words, high predictability of the general satisfaction. Next, based on linear regression output from the second stage, the respondents and the variables are grouped through cluster analysis.

3.6.1 Descriptive Statistics of the Sample

This part starts with presenting the sample profile. The three frequency charts are provided next. Figure 10, Figure 11, and Figure 12 visualize the gender, age, and education frequency, respectively.

As can be seen in Figure 10, the sample consists of 34% female and 56.2% male participants. 0.1% decided on “other” gender identification, and 9.6% of the participants did not mention their gender at all.

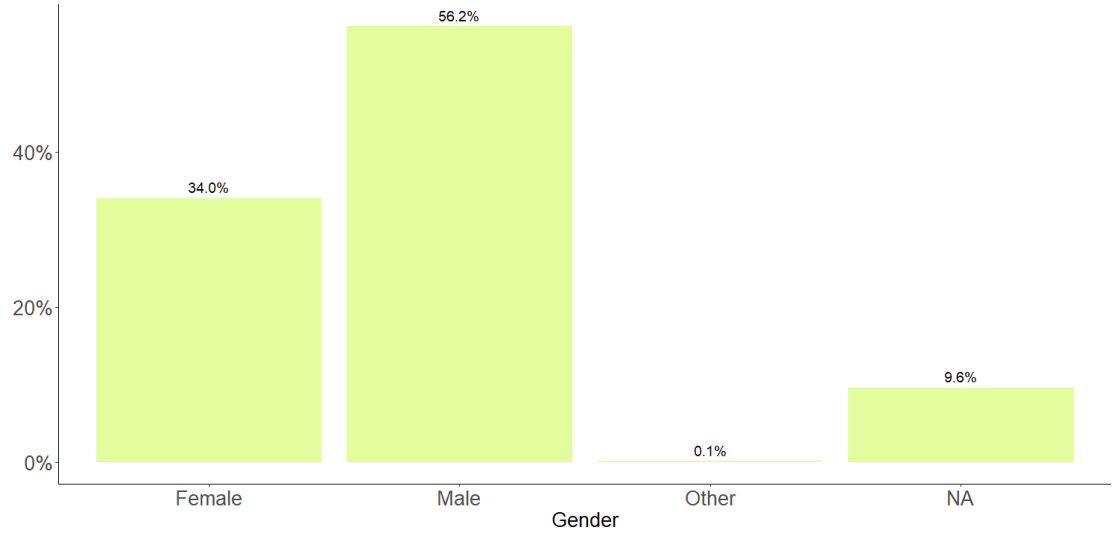


FIGURE 10: GENDER FREQUENCY.

Figure 11 showcases the age groups of the participants. Not to forget that participation was granted only starting from the age of 14. Therefore the 0-13 age group is not considered. The age group 14-25 provided 10.43% of the feedback, the age group 26-45 provided 35.67% of the feedback, and the age group, 45-65, provided the most significant share of the feedback, 39.63%. The group consisting of respondents older than 65 provided 3.85% of the feedback. 10.42% of the respondents did not identify their age.

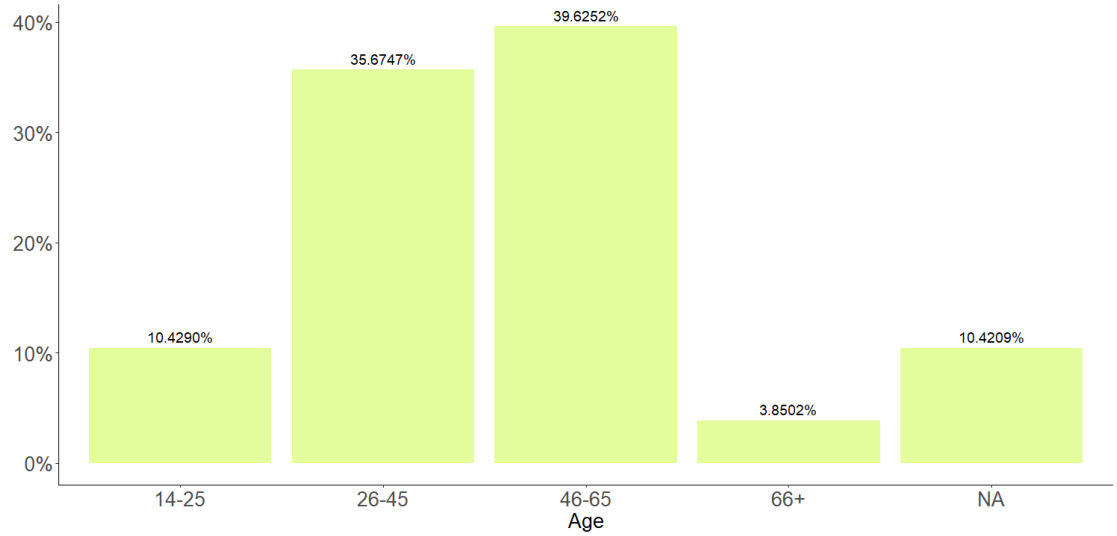


FIGURE 11: AGE FREQUENCY.

Figure 12 reflects the education-level overview of the sample. 24.8% of respondents identified themselves as graduates of lower secondary education, 45.8% identified themselves as graduates of tertiary education, 18.5% identified themselves as upper secondary education graduates, and 10.9% did not mention their education level.

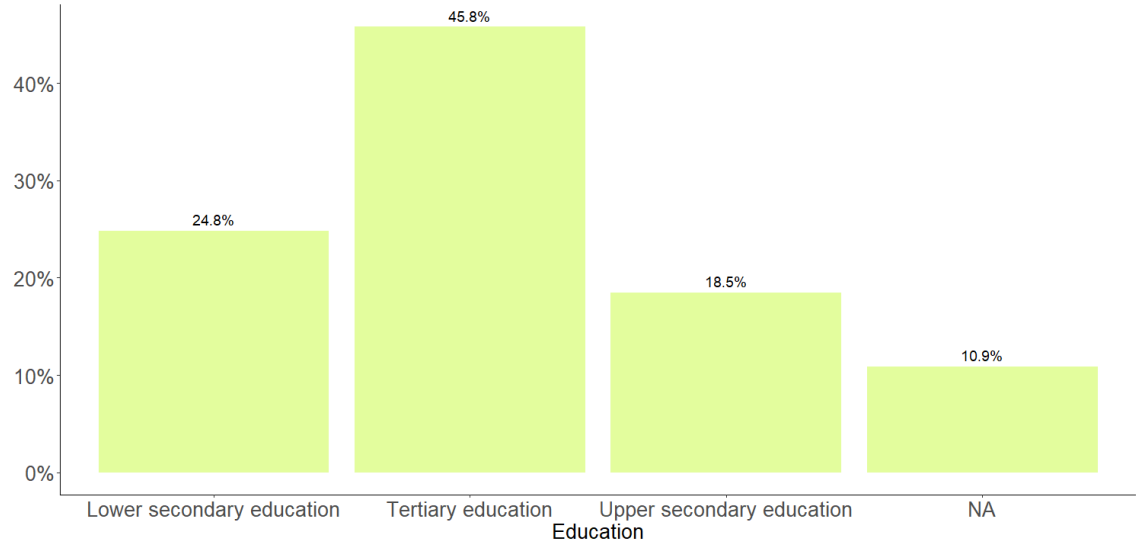


FIGURE 12: EDUCATION FREQUENCY.

Table 4 showcases the mean and standard deviation of the 48 semantic differential scaled variables with a range from 1 to 6. Out of those variables, 47 are the independent variables, and 1 is the dependent variable. While the names of the variables with not significant p-values ($p<.05$ soft green, $p<.1$ soft orange, the dependent variable is in dark green in Table 4) in the second phase are taken out of the table, their mean and standard deviation are left in place for the comparison purposes.

Mean (SD)	1.705 (0.856)	Mean (SD)	1.809 (0.869)	Mean (SD)	1.910 (0.892)
Lift_Access		**Services_Ski_Schools**		**Gastronomy_Atmosphere**	
Mean (SD)	1.756 (0.852)	Mean (SD)	1.711 (0.934)	Mean (SD)	1.817 (0.861)
Mean (SD)	1.972 (1.028)	Mean (SD)	1.595 (0.821)	Mean (SD)	2.224 (1.056)
				Gastronomy_Waiting_Time	
Mean (SD)	1.833 (0.960)	Mean (SD)	1.665 (0.939)	Mean (SD)	2.068 (0.945)
Mean (SD)	1.554 (0.766)	Mean (SD)	1.958 (0.854)	Mean (SD)	1.751 (0.780)
General_Ski_Resort		**General_Hiking**		**Gastronomy_Value_for_Money**	
Mean (SD)	1.613 (0.768)	Mean (SD)	2.071 (1.047)	Mean (SD)	2.390 (1.061)
		General_Children_Offers			
Mean (SD)	1.744 (0.845)	Mean (SD)	1.912 (0.956)	Mean (SD)	1.691 (0.819)
				General_Offers	
Mean (SD)	1.722 (0.864)	Mean (SD)	2.095 (1.067)	Mean (SD)	1.780 (0.730)
		General_Rodelbahn		**General_Atmosphere Win-ter_Experience**	
Mean (SD)	1.834 (0.879)	Mean (SD)	1.999 (1.132)	Mean (SD)	1.697 (0.782)
				Apres_Ski	
Mean (SD)	1.677 (0.900)	Mean (SD)	2.004 (1.050)	Mean (SD)	2.006 (1.071)
Ski_Resort_Panorama					
Mean (SD)	1.354 (0.659)	Mean (SD)	2.024 (1.063)	Mean (SD)	1.924 (0.881)
General_Slopes				**General_Location**	
Mean (SD)	1.677 (0.769)	Mean (SD)	2.101 (1.088)	Mean (SD)	1.811 (0.825)
				General_Accommodation	
Mean (SD)	1.801 (0.942)	Mean (SD)	1.923 (0.902)	Mean (SD)	1.535 (0.772)
				General_Overall	
Mean (SD)	1.693 (0.808)	Mean (SD)	1.902 (0.877)	Mean (SD)	1.631 (0.713)
Mean (SD)	1.961 (0.957)	Mean (SD)	1.948 (0.926)	Mean (SD)	2.221 (1.184)
Slopes_Safety				**Slopes_Space**	
Mean (SD)	1.697 (0.807)	Mean (SD)	1.904 (0.882)	Mean (SD)	1.993 (0.984)

TABLE 4: DESCRIPTIVE STATISTICS.

p<0.05; p<0.1;p>0.1; dependent variable

The original dataset has shortened German variable names, which the author substitutes with full English names to ease the outputs' reading. The underscore between single words of the variable name is a software requirement. Therefore, while speaking about the "Gastronomy Atmosphere" factor, the “Gastronomy_Atmosphere” variable is discussed, and similarly with others.

For a better understanding of the reader and easier comparison of the factors, Table 5 is created by taking the 19 factors from Table 4 and sorting them by mean from lowest to highest.

Name	Mean	Standard Deviation
Ski_Resort_Panorama	1.354	0.659
General_Accommodation	1.535	0.772
General_Ski_Resort	1.613	0.768
General_Overall	1.631	0.713
General_Slopes	1.677	0.769
General_Atmosphere Winter_Experience	1.697	0.782
Slopes_Safety	1.697	0.807
Services_Ski_Schools	1.711	0.934
Lift_Access	1.756	0.852
General_Offers	1.780	0.730
General_Location	1.811	0.825
Gastronomy_Atmosphere	1.817	0.861
General_Children_Offers	1.912	0.956
Slopes_Space	1.993	0.984
General_Rodelbahn	1.999	1.132
Apres_Ski	2.006	1.071
Gastronomy_Waiting_Time	2.068	0.945
General_Hiking	2.071	1.047
Gastronomy_Value_for_Money	2.390	1.061

TABLE 5: TABLE 4 SORTED AND CLEANED. DESCRIPTIVE STATISTICS.

Factors with $p<0.05$ are highlighted with white background. Factors with $0.05 <p<0.1$ are highlighted with light orange background. Overall satisfaction factor with dark green.

At first, the three unique variables will be presented to the reader. Those are highlighted in color in Table 5. The dependent variable “General_Overall” has a mean of 1.631 with a standard deviation of 0.713. The other two variables highlighted are the variables with their $0.05<p<0.1$, which do not show any strange behavior and are located closer to the middle of Table 5. The factor "General Rodelbahn" (where rodelbahn is a gravity-driven rollercoaster) has a mean of 1.999 and a standard deviation of 1.132, and the "Slopes Safety" factor has a mean of 1.697 and a standard deviation of 0.807. It is also worth mentioning that the “General_Rodelbahn” factor has the highest standard deviation in Table 5. The other 16 variables are presented according to their means, starting with the most satisfied and descending further. The "Ski Resort Panorama" factor has a mean of 1.354 and a standard deviation of 0.659. The next closest is "General Accommodation." This factor has a mean of 1.535 and a standard deviation of 0.772. The next closest is "General Ski Resort," and it has a mean of 1.613 and a standard deviation of 0.768.

Next goes the “General_Overall” factor, but it is already mentioned. The following one is “General_Slopes,” and it has a mean of 1.677 and a standard deviation of 0.769. The subsequent factor is the "General Atmosphere Winter Experience," and it has a mean of 1.697 and a standard deviation of 0.782. The next closest by mean in descending order factor is "Services Ski Schools," and it has a mean of 1.711 and a standard deviation of 0.934. They are followed by the "Lift Access" factor, which has a mean of 1.756 and a standard deviation of 0.852. Next goes the "General Offers" factor, it has a mean of 1.780 and a standard deviation of 0.730. Followed by the "General Location" and "Gastronomy Atmosphere" factors, which are very close and have a mean of 1.811, a standard deviation of 0.825, a mean of 1.817, and a standard deviation of 0.892, respectively. The next is the “General Children Offers” factor which has a mean of 1.912 and a standard deviation of 0.956. and the “Slopes Space” factor has a mean of 1.993 and a standard deviation of 0.984. Those are the last two factors before the threshold of 2, excluding the already mentioned parameter with its $p>0.05$. The next four variables are the ones having the highest range between them, and the “Ski_Resort_Panorama” factor, having the lowest mean of 1.354. Those are the "Après Ski" factor having a mean of 2.006 and a standard deviation of 1.071. The "Gastronomy Waiting Time" factor, having a mean of 2.068 and a standard deviation of 0.945. The "General Hiking" factor, having a mean of 2.071 and a standard deviation of 1.047. And finally, the one having the highest mean, the "Gastronomy Value for Money" factor, having a mean of 2.390 and a standard deviation of 1.061. Having the essential descriptive statistic aspects outlined, it is time to move to the multiple linear regression procedure.

3.6.2 Multiple Linear Regression

Before coming to the multiple linear regression, the assumptions and decisions of the author should be clarified. The central assumption of multiple linear regression is the interval scale of both independent and dependent variables. In this case, we have the range from 1 to 6. Nonetheless, it cannot be guaranteed that the distance between 1 and 2 is equal to the distance between 4 and 5. Therefore, the scale of the data is ordinal. This violates the assumption of multiple linear regression. Nevertheless, due to the interpretation of ordinal regression being not user-friendly and cumbersome, in this research, the author assumes that the intervals of the scale are equal. Figures 13 to 16 are provided for a better overview, showcasing information on multiple regression assumptions regarding the dataset.

According to Figure 13 and Figure 14, it can be seen that the range of 1 to 6 is not covered equally and the parameters 1, 2, and 3 are much better represented than 4, 5, and 6 because the lines on the left side are much better outlined then on the right. In Figure 13, the dots in the optimal world should be spread equally between the minimum and maximum range and not follow the lined patterns as we see in this case. Also, the red line should be a straight horizontal line exactly at zero. What comes to Figure 14, there is a similar situation. In this case, dots are creating the two curved lines starting at the integers and going left and right, but should normally by equally

spread between minimum and maximum range, and the red line should be horizontal in the middle of the graph.

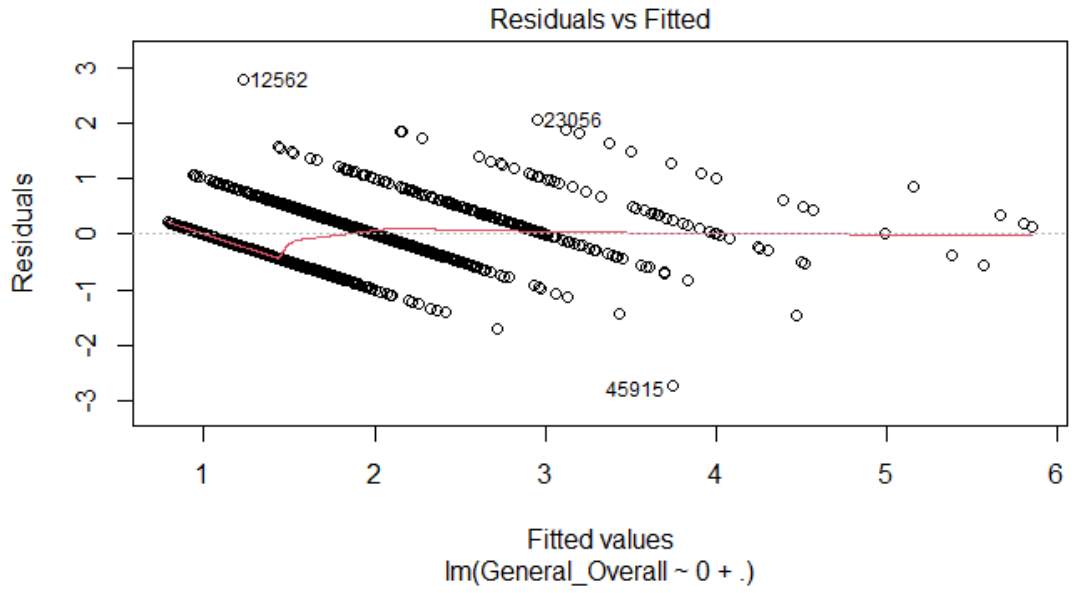


FIGURE 13: LINEARITY.

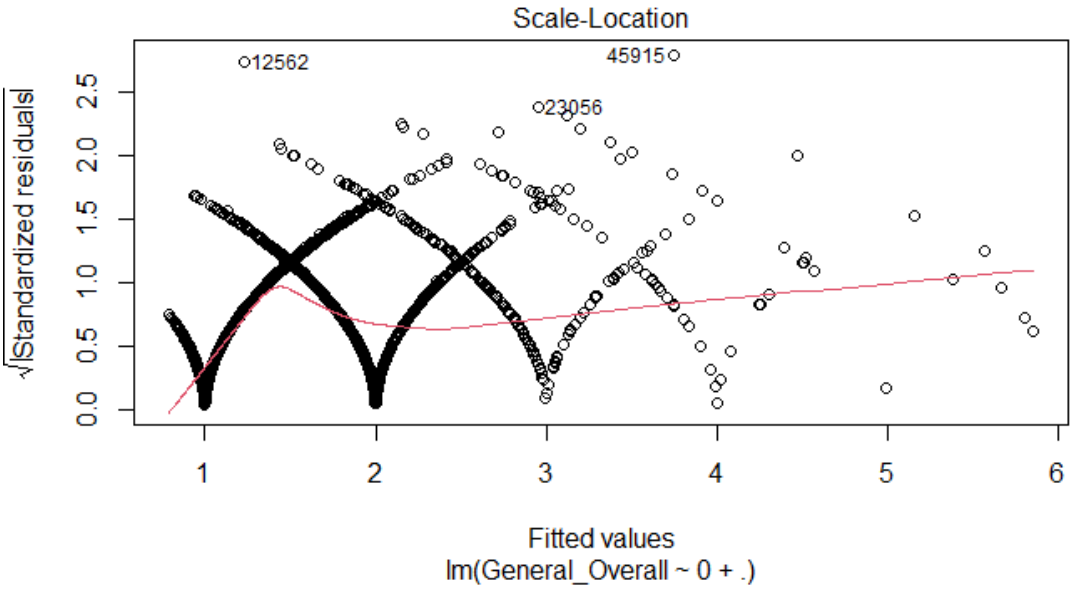


FIGURE 14: HOMOSCEDASTICITY.

According to Figure 15, the standardized residuals compared to the fitted regression line show lower values on the left side and higher values on the right side. In the optimal situation, this

graph should have the big round dots following the small black dotted line creating a perfectly straight line. In a way, it will be impossible to see that the black dotted line is there.

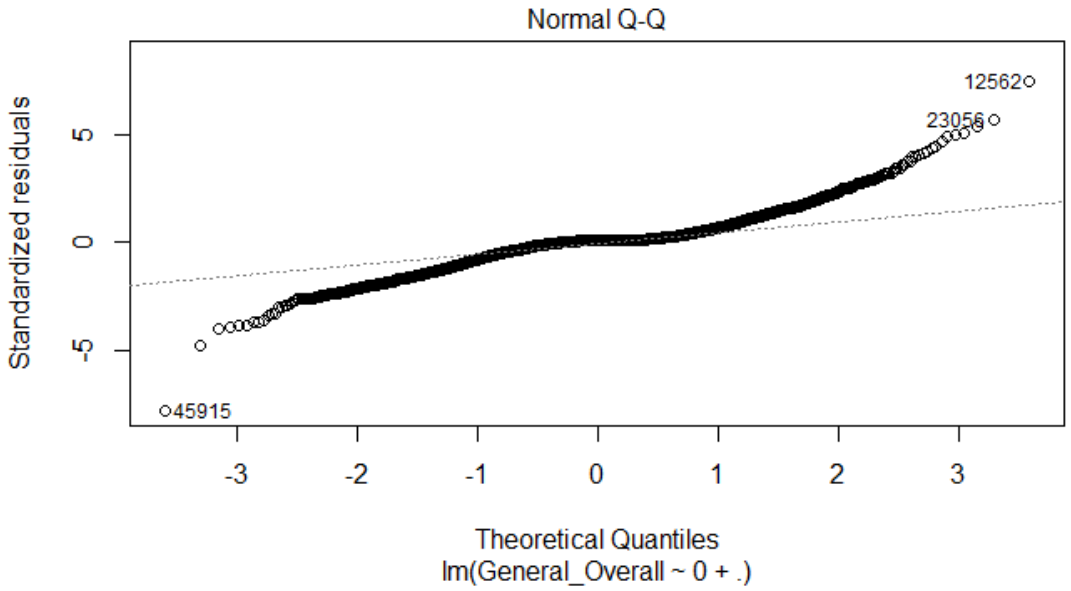


FIGURE 15: NORMALITY.

According to Figure 16, there are no outliers as the scale is bound to the range of 1 to 6. Moreover, the red dash-line, which represents the limit, is not even shown here due to the limited range. As the range is limited, this graph is not as useful as it could be, and in the optimal situation, it also shows the red dashed lines representing Cook's distance. The optimal situation will also bring the horizontal red line, and the dots should be equally distributed along the values.

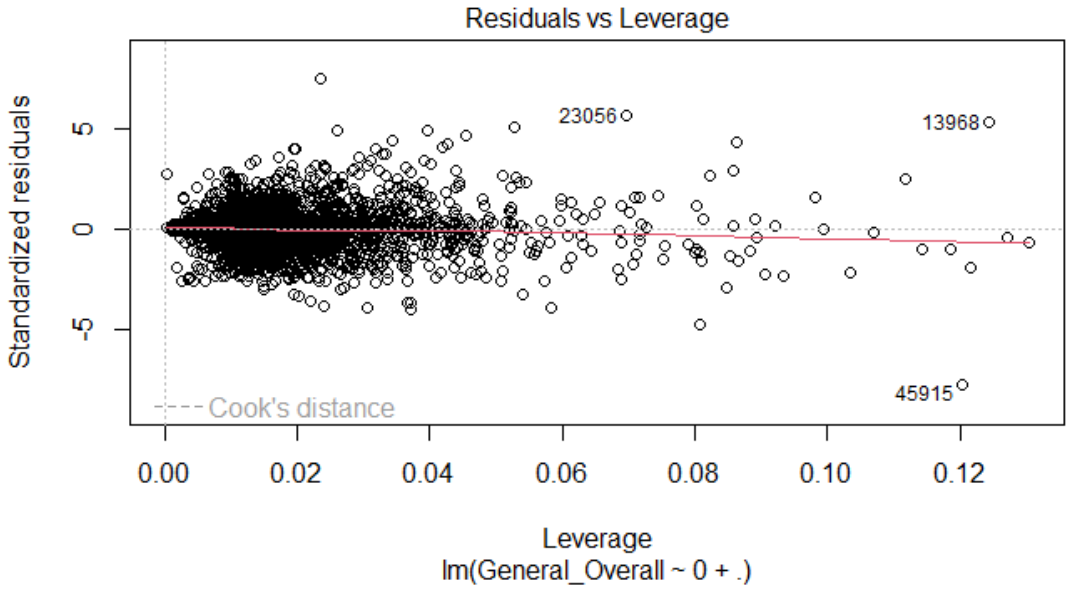


FIGURE 16: ABSENCE OF OUTLIER.

As one can see, multiple linear regression assumptions are not fully met. Therefore, the reader should be aware of possible distortions due to these violations.

After the assumptions are clarified and some preliminary data cleaning, the author runs the multiple linear regression with the following parameters: “General Overall” factor as a dependent, all the other 47 factors left as an independent. This process allows checking for the independent parameters' significant influence on the dependent one, thus testing the hypothesis mentioned in the literature review and hypothesis development chapter.

Hypothesis tested:

H0: No characteristic can predict the overall satisfaction of the customer of the skiing area in Austria.

H1: Overall satisfaction of the customer of the skiing area in Austria can be predicted by a specific characteristic of the skiing area.

At first, the model with intercept is used to see if the intercept p-value is lower than 5%. As it was not significant, the intercept value was excluded from the model. Furthermore, the final model summary is now showcased in Table 6.

Predictors	Estimates	p
General Ski Resort	0.11	<0.001
Gastronomy Atmosphere	0.06	<0.001
Gastronomy Waiting Time	-0.05	<0.001
General Offers	0.17	<0.001
General Atmosphere Winter Experience	0.14	<0.001
General Location	0.17	<0.001
General Accommodation	0.24	<0.001
Ski Resort Panorama	0.05	0.001
Apres Ski	0.04	0.001
Services Ski Schools	-0.03	0.004
Gastronomy Value for Money	0.03	0.007
General Hiking	-0.03	0.008
General Slopes	0.04	0.014
Lift Access	0.03	0.018
General Children Offers	0.03	0.019
Slopes Space	0.03	0.024
General Rodelbahn	0.02	0.052
Slopes Safety	0.02	0.077
	-0.02	0.136
	0.02	0.156
	0.02	0.157
	0.02	0.170
	-0.02	0.171
	-0.02	0.186
	-0.02	0.191
	-0.02	0.216
	-0.01	0.227
	-0.02	0.228
	-0.02	0.257
	-0.02	0.288
	0.01	0.326
	0.01	0.433
	-0.01	0.440
	0.01	0.444
	-0.01	0.524
	-0.01	0.525
	0.01	0.525
	0.01	0.584
	0.01	0.584
	-0.01	0.607

	-0.01	0.748
	0.00	0.750
	-0.00	0.812
	-0.00	0.815
	-0.00	0.866
	-0.00	0.888
	-0.00	0.893
Observations	3093	3093
R sq / R sq adjusted	0.952 / 0.951	0.952 / 0.951

TABLE 6: MULTIPLE LINEAR REGRESSION MODEL SUMMARY.

p<0.05; p<0.1;p>0.1

Before explaining the effects of the single factors, the overall outline of Table 6 is described. The variables in Table 6 are sorted by their p-value from the lowest to the highest. The first 16 are highlighted in light-green are having a p-value lower than 0.05. The following two highlighted light-yellow have a p-value lower than 0.1. While a p-value higher than 0.05 is not considered significant, the author added the values with a p-value lower than 0.1 and higher than 0.05 to expand the number of variables, as those are pretty close to being significant. As the lower part of Table 6 indicates, the test included 3,093 observations and led to the R squared and R squared adjusted of 0.952 and 0.951, respectively. The close to 1.00 result suggests that the independent variables are of good quality and can almost perfectly predict the dependent one.

At first, the variables with a positive effect on the overall satisfaction, meaning a positive estimated coefficient, are presented. Then the ones with negative coefficients and negative effects on overall satisfaction are given. The factors having p<0.001 are the ones having the most impactful estimated coefficients, and therefore they are discussed before the others in descending order.

The estimated coefficient for the variable "General Accommodation" is 0.24 (<0.001). This means if the evaluation of the "General Accommodation" increases by 1, the evaluation of the “General Overall” increases by 0.24. The estimated coefficient for the variable "General Offers" is 0.17 (<0.001). This means if the evaluation of the "General Offers" increases by 1, the evaluation of the “General Overall” increases by 0.17. The estimated coefficient for the variable "General Location" is 0.17 (<0.001). This means if the evaluation of the "General Location" increases by 1, the evaluation of the “General Overall” increases by 0.17. The estimated coefficient for the variable "General Atmosphere Winter Experience" is 0.14 (<0.001). This means if the evaluation of the "General Atmosphere Winter Experience" increases by 1, the evaluation of the “General Overall” increases by 0.14. The estimated coefficient for the variable "General Ski Resort" is 0.11

(<0.001). This means if the evaluation of the "General Ski Resort" increases by 1, the evaluation of the “General Overall” increases by 0.11. The estimated coefficient for the variable "Gastronomy Atmosphere" is 0.06 (<0.001). This means if the evaluation of the "Gastronomy Atmosphere" increases by 1, the evaluation of the “General Overall” increases by 0.06. The estimated coefficient for the variable "Ski Resort Panorama" is 0.05 (0.001). This means if the evaluation of the "Ski Resort Panorama" increases by 1, the evaluation of the “General Overall” increases by 0.05.

All the following variables have their estimated coefficient less or equal to 0.05. The estimated coefficient for the variable "Apres Ski" is 0.04 (0.001). This means if the evaluation of the "Apres Ski" increases by 1, the evaluation of the “General Overall” increases by 0.04. The estimated coefficient for the variable "Services Ski Schools" is -0.03 (0.004). This means if the evaluation of the "Services Ski Schools" increases by 1, the evaluation of the “General Overall” increases by -0.03. The estimated coefficient for the variable "Gastronomy Value for Money" is 0.03 (0.007). This means if the evaluation of the "Gastronomy Value for Money" increases by 1, the evaluation of the “General Overall” increases by 0.03. The estimated coefficient for the variable "General Slopes" is 0.04 (0.014). This means if the evaluation of the "General Slopes" increases by 1, the evaluation of the “General Overall” increases by 0.04. The estimated coefficient for the variable "Lift Access" is 0.03 (0.018). This means if the evaluation of the "Lift Access" increases by 1, the evaluation of the “General Overall” increases by 0.03. The estimated coefficient for the variable "General Children Offers" is 0.03 (0.019). This means if the evaluation of the "General Children Offers" increases by 1, the evaluation of the “General Overall” increases by 0.03. The estimated coefficient for the variable "Slopes Space" is 0.03 (0.024). This means if the evaluation of the "Slopes Space" increases by 1, the evaluation of the “General Overall” increases by 0.03. The estimated coefficient for the variable "General Rodelbahn" is 0.02 (0.052). This means if the evaluation of the "General Rodelbahn" increases by 1, the evaluation of the “General Overall” increases by 0.02. The estimated coefficient for the variable "Slopes Safety" is 0.02 (0.077). This means if the evaluation of the "Slopes Safety" increases by 1, the evaluation of the “General Overall” increases by 0.02.

While it is logical to expect that greater satisfaction with one parameter of the system should lead to greater overall satisfaction, when dealing with big data amounts, it can happen that there are some parameters in the model, the growth of which negatively affects the overall satisfaction. In this case, we have three of those:

The estimated coefficient for the variable "Gastronomy Waiting Time" is -0.05 (<0.001). This means if the evaluation of the "Gastronomy Waiting Time" increases by 1, the evaluation of the “General Overall” decreases by 0.05.The estimated coefficient for the variable "General Hiking" is -0.03 (0.008). This means if the evaluation of the "General Hiking" increases by 1, the evalua-

tion of the “General Overall” decreases by 0.03. The estimated coefficient for the variable "Services Ski Schools" is -0.03 (0.004). This means if the evaluation of the "Services Ski Schools" increases by 1, the evaluation of the “General Overall” decreases by 0.03.

The rest of the variables in Table 6 are greyed out. Their estimates parameter is in the range between -0.02 and 0.02, and the p-value is higher than 0.1.

Figure 17 and Figure 18 visualize the semantic differential scale of all significantly influential variables. The decision to divide the data into two figures is justified by the limitation of the software and the readability of the text inside of the figures on the printed version of the study.

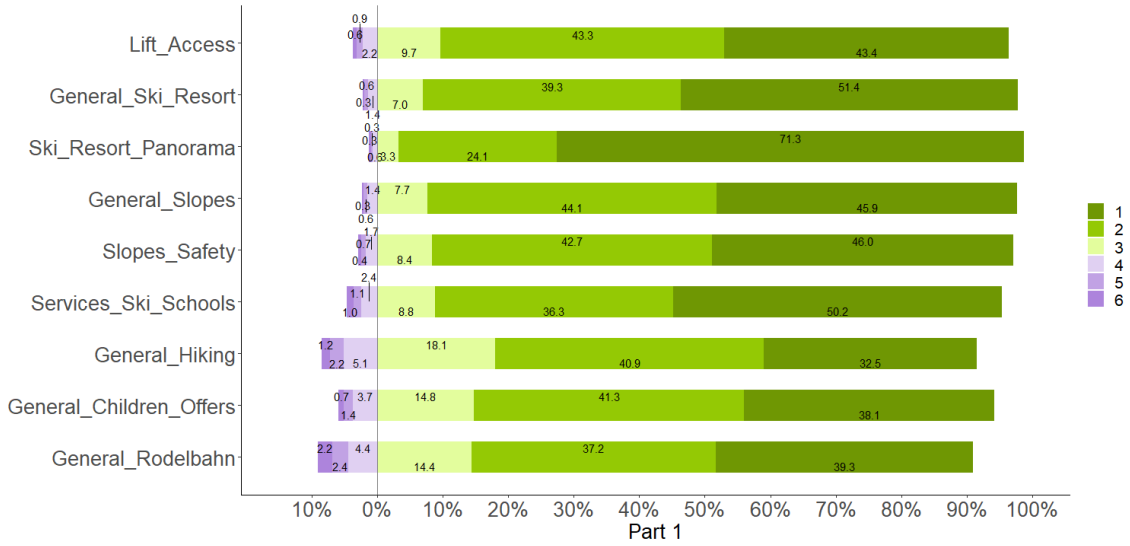


FIGURE 17: SEMANTIC DIFFERENTIAL SCALED FACTORS WITH SIGNIFICANT INFLUENCE, PART 1.

Figure 17 showcases “Lift Access,” “General Ski Resort,” “Ski Resort Panorama,” “General Slopes,” “Slopes Safety,” “Services Ski Schools,” “General Hiking,” “General Children Offers,” and “General Rodelbahn” factors.

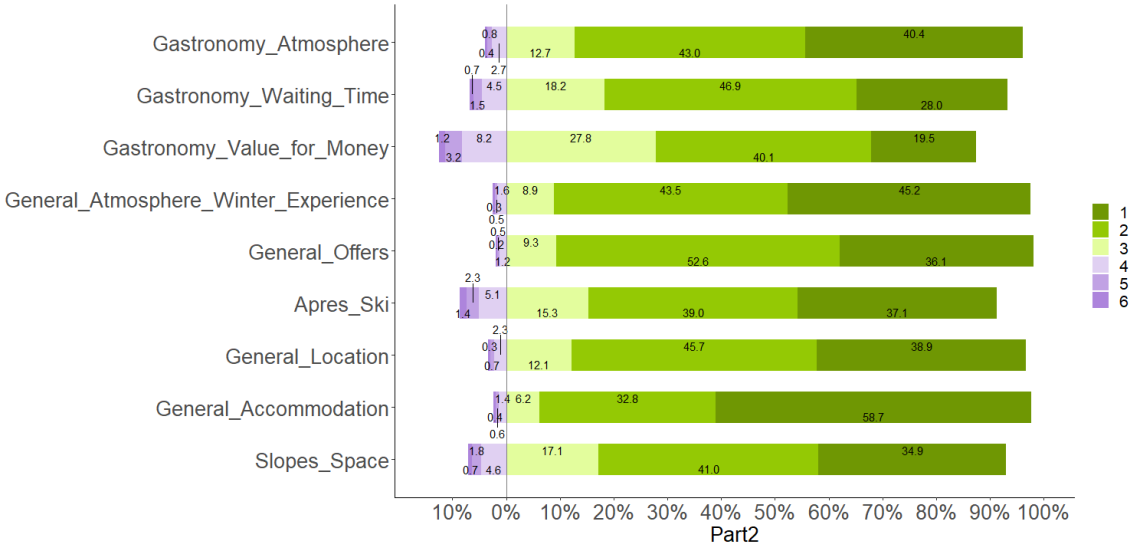


FIGURE 18: SEMANTIC DIFFERENTIAL SCALED FACTORS WITH SIGNIFICANT INFLUENCE, PART 2.

Figure 18 showcases "Gastronomy Atmosphere," "Gastronomy Waiting Time," "Gastronomy Value for Money," "General Atmosphere Winter Experience," "General Offers," "Après Ski," "General Location," "General Accommodation," and "Slopes Space" factors.

The "Ski Resort Panorama" variable has the highest share of the significantly satisfied response, with 71.3% of the respondents choosing 1 out of 6. The next factor is "General Accommodation," having 58.7% of the respondents evaluating the characteristic as exceptionally satisfactory. The characteristics "General Ski Resort" and "Services Ski Schools" can share the third place, as their share of highly satisfiable evaluations are 51.4% and 50.2%, respectively, which is very close to each other.

Figures 17 and 18 imply that most of the responses are located on the right side, where satisfactory positive answers are found. Leaving on the left side only less than ten percent of the unsatisfied responses, excluding one exception with "Gastronomy Value For Money" variable, that in total broke this threshold with 12.6% of respondents on the negative satisfaction side. The next closest would be the variables "General Hiking" and "General Rodelbahn," which still have less than 10% of dissatisfactory answers.

3.6.3 Cluster Analysis

This part of the research looks at the data output created after the multiple linear regression analysis from the two dimensions. Firstly, the customers are grouped, and then the variables are.

The dataset for this cluster analysis has 18 variables and more than 45 thousand respondents. The variables are rows, and respondents are columns for the clustering of the customers. Moreover, the transposed dataset with variables as columns and customers as rows is used for the later clustering of the variables.

In an attempt to group customers in an "a-posteriori" data-driven way, the test is done with the following parameters. Hierarchical clustering is chosen because there is no "a priori" knowledge of the number of groups there should be. The clustering procedure is done using Euclidean distance and Ward's linkage method. It uses the overall sum of the squared distance to the cluster means and groups together variables of clusters with the smallest increase in it. Also, important to mention as all the variables have the same scale, there is no standardization needed before the conduction of the hierarchical clustering procedure. The parameters of the test are outlined in Figure 19.

```
Call:
hclust(d = dist(dat1, method = "euclidean"), method = "ward.D2")

Cluster method   : ward.D2
Distance         : euclidean
Number of objects: 45161
```

FIGURE 19: RESPONDENTS' CLUSTERING PARAMETERS.

The number of respondents in this research is so high that the dendrogram of the clustering procedure is limited in its informativity. Nonetheless, after evaluating its height parameter and testing various other options, the tree was cut out into three groups. Meaning the respondents were grouped into three clusters. Figure 20 showcases the number of respondents in clusters after cutting the sample into three groups, dividing the total number into 15,064, 25,450, and 4,647 for groups one, two, and three, respectively.

```
hclustergroups
    1     2     3
15064 25450  4647
```

FIGURE 20: NUMBER OF RESPONDENTS IN GROUPS.

In order to find if there are segments with different priorities, the calculation of the group means for the variables is done and showcased in Table 7. In Table 7, it can be seen that the means increase step by step from group to group in all 18 variables. This means there is no multidimensional effect when clustering respondents. In other words, the grouping is done by the level of customer satisfaction overall and not by their identification as being satisfied by one group of variables while being dissatisfied with others, as it would be if there were some other data output other than gradually increasing means. For example, if there would be a variable with, e.g., the lowest value for group 3 and highest for group 1.

Group	Lift Access	General Ski Resort	Ski Resort Panorama
1	1.228385	1.112478	1.046428
2	1.858939	1.679820	1.380755
3	2.862660	2.810304	2.170790
Group	**General Hiking**	**General Children Offers**	**General Rodelbahn**
1	1.490549	1.362462	1.448937
2	2.295451	2.100086	2.183028
3	3.195668	3.001290	3.221799
Group	**General Atmosphere Winter Experience**	**General Offers**	**Apres Ski**
1	1.077729	1.151134	1.370313
2	1.844170	1.933458	2.192422
3	2.905786	2.966958	3.308291
Group	**General Slopes**	**Slopes Safety**	**Services Ski Schools**
1	1.176427	1.207216	1.263918
2	1.759511	1.797341	1.843347
3	2.814975	2.732726	2.669725
Group	**Gastronomy Atmos-phere**	**Gastronomy Waiting Time**	**Gastronomy Value for Money**
1	1.175460	1.445140	1.655883
2	1.986837	2.240111	2.587696
3	2.968874	3.155417	3.685797
Group	**General Location**	**General Accommodation**	**Slopes Space**
1	1.193087	1.138309	1.412126
2	1.998790	1.653516	2.089351
3	2.960769	2.314304	3.322398

TABLE 7: CLUSTER MEANS.

The results of the clustering procedure, in this case, groups the customers based on their level of satisfaction.

After inspecting the customers' segmentation possibilities in the previous phase, this part of the research is running the hierarchical cluster analysis to group variables in an “a posteriori” data-driven way. In this case, hierarchical clustering is used because there are not too many variables to group, and there is no “a priori” knowledge about how many groups there are. For this study, the hierarchical clustering procedure with Euclidean distance and Ward’s linkage method is chosen, as it is one of the most popular and reliable.

The following RStudio output (Figure 21) describes the clustering analysis done with the previously mentioned parameters:

```
Call:
hclust(d = dist(data, method = "euclidean"), method = "ward.D2")

Cluster method   : ward.D2
Distance         : euclidean
Number of objects: 18
```

FIGURE 21: VARIABLES' CLUSTERING PARAMETERS.

As an outcome of this cluster analysis, the cluster dendrogram is presented in Figure 22.

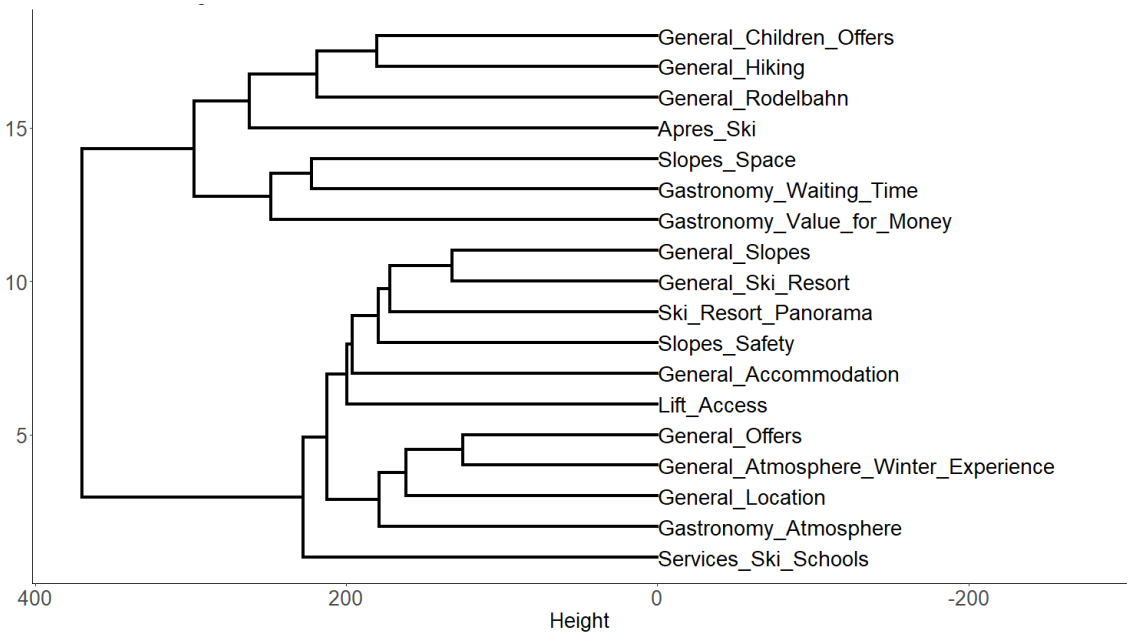


FIGURE 22: CLUSTER DENDROGRAM.

Based on the data provided in Figure 22, its height, and potential groupings, the variables are divided into three groups. This decision is visualized in Figure 23.

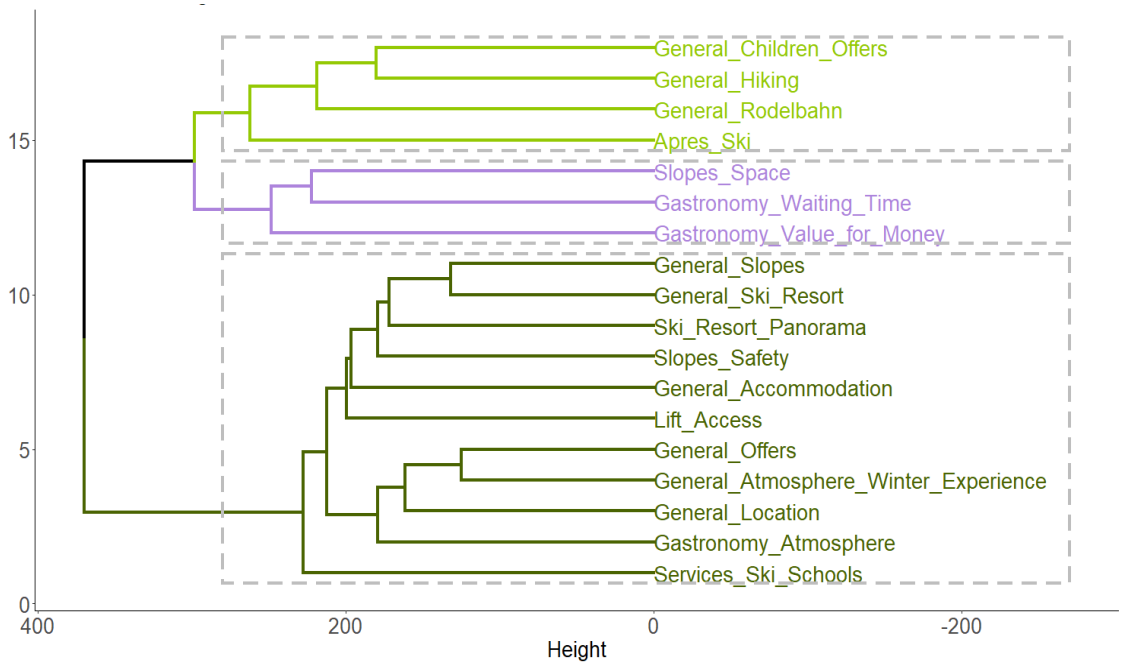


FIGURE 23: CLUSTER DENDROGRAM, WITH OUTLINED CLUSTERS.

As a result, we have three groups with 11, 4, and 3 variables in them. The decision to have three clusters is based on the variables grouped. When looking at the height of the dendrogram and considering 2 and 3 groups, both looked well, but after looking at the cluster plot in Figure 24 and seeing how concentrated group 1 is and how spread group 2 was, the decision was taken to cut the tree at 3 clusters and divide the group 2 into groups 2 and 3. This is further explained in the results chapter.

This part of the research represents the algorithm of the research from its basement and creation to its execution. The next part of the research elaborates more on the study's results, broadens the outputs provided in the data analysis subchapter, and explains their potential use in real-life scenarios.

4 RESULTS AND DISCUSSION

This chapter consists of three main parts. The first one gives an overview of the research, elaborating more about the reasons for this research. The last two discuss the results of the two main phases of the study, multiple linear regression and hierarchical cluster analysis.

4.1 Research Overview

Austrian skiing areas are essential for the economy (Pröbstl-Haider et al., 2021). Some of the regions of Austria rely on seasonal profits from winter tourism almost entirely (Pröbstl-Haider et al., 2021). In the situation where the alpine ski businesses suffer the consequences of various crises such as COVID-19 pandemics and climate change, it is wise to adjust business strategies according to the situation (Pröbstl-Haider, Mostegl, et al., 2021). However, while businesses adjust their strategies to be profitable, the customers should be satisfied (Vanat, 2020). In order to recognize the changes in customer behavior and to find out the satisfaction levels of the clients, the satisfaction surveys are conducted by the alpine ski resorts on a seasonal basis with the help of specialized market research companies. The more satisfied the customers are, the more the chances are that they will recommend the resort to their friends and family, and the higher the chances of a repeat visit (Hill, 2006). Therefore, such data gathering is essential. Also, while dealing with many characteristics, it is vital to identify which characteristics the management should be focused on in the first place and which are irrelevant or should be left for the second stage. This research is conducted to identify the most influential characteristics of the alpine ski resorts based on the data from the Austrian alpine ski resorts season 2019/2020 and then to segment the customers in the second stage to provide better service to the different customer groups. For the purpose of identifying the influential characteristics of the resort, 47 characteristics of the resort were taken as independent variables, and general satisfaction was taken as a dependent variable of the multiple linear regression.

4.2 Multiple Linear Regression Results Discussion

As a result of the multiple linear regression, 18 out of 47 variables are identified as the ones with significant influence on overall satisfaction. The result is based on the 3,093 observations as outlined in Table 6. Moreover, the R squared of the research and the R squared adjusted are 0.952 and 0.951, respectively. The R-squared results indicate the very high predictability of the dependent variable by the independent ones. The estimates of the significantly influential variables ranged from -0.05 to 0.24. The results of the rest of the factors ranged between -0.02 and 0.02.

The next stage of the research tried to segment customers in the “a-posteriori” data-driven way based on their satisfaction with the most influential characteristics.

4.3 Cluster Analysis Results and Discussion

As mentioned in the overview, after the attributes with significant influence on overall satisfaction are identified, the researcher runs a hierarchical cluster analysis. This procedure is done to specify if there is a need to assign customers to different groups and adjust the products and services provided by the ski resorts to suit the clients better.

The first part of the cluster analysis resulted in clustering the respondents into three groups without multidimensional differences. Meaning the respondents are grouped by their level of satisfaction. While it varies not significantly between variables, it can be concluded that group 1, based on all variables, is more satisfied than group 2, and group 2 is happier than group 3, according to Table 7.

The research idea included creating customers' personas for the groups created by clustering analysis if there would be multidimensional differences. In this case, creating personas for the three groups is unnecessary as there are not enough differences other than the general level of satisfaction to differentiate customers.

The second part of the cluster analysis is concerned with clustering the most influential variables. The results of this clustering procedure showed the three groups of variables with similar scores, as presented in Figure 23. Figure 24 showcases the cluster plot visualizing the distances between the variables.

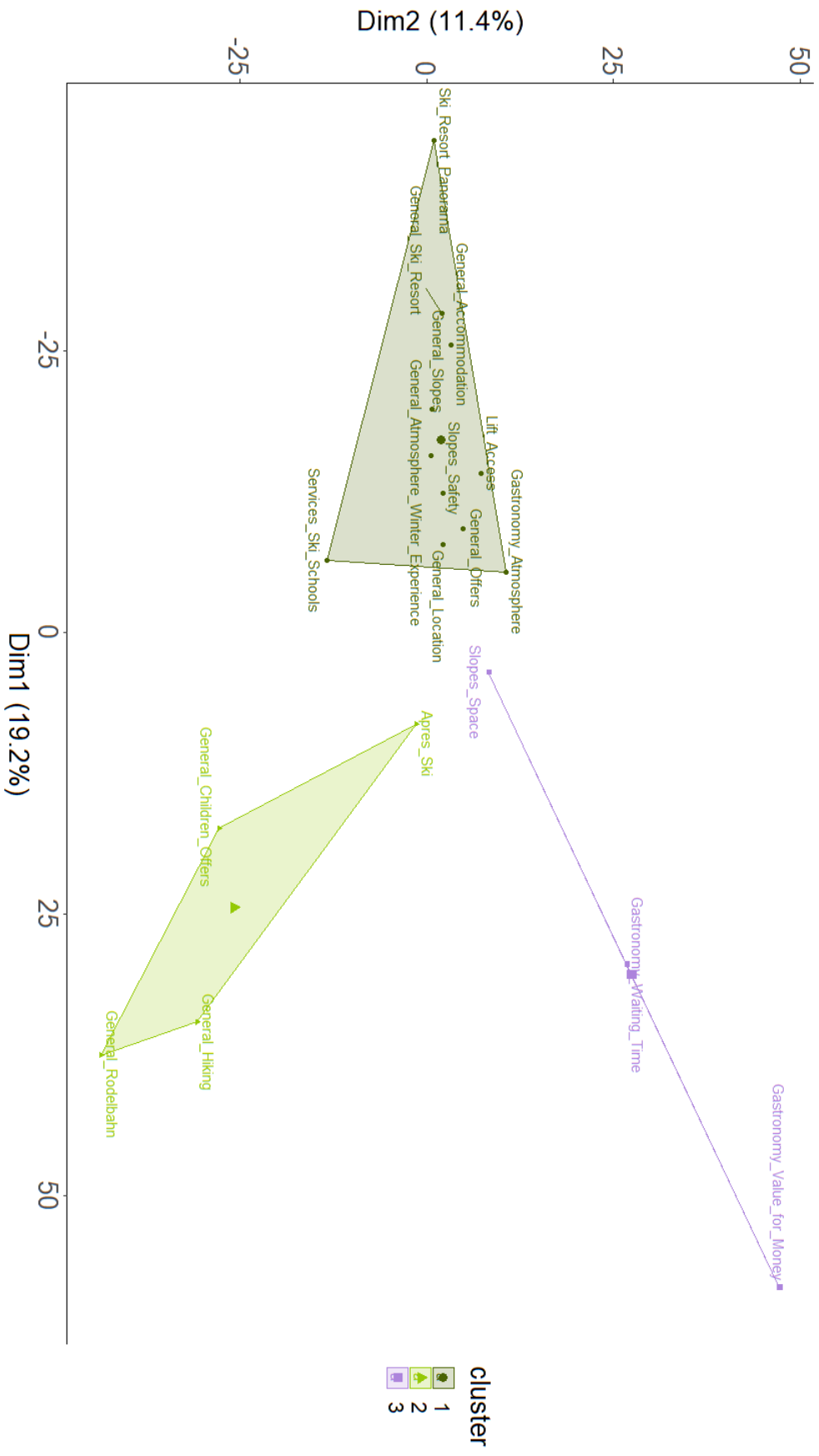


FIGURE 24: CLUSTER PLOT OF THE VARIABLES CLUSTER ANALYSIS.

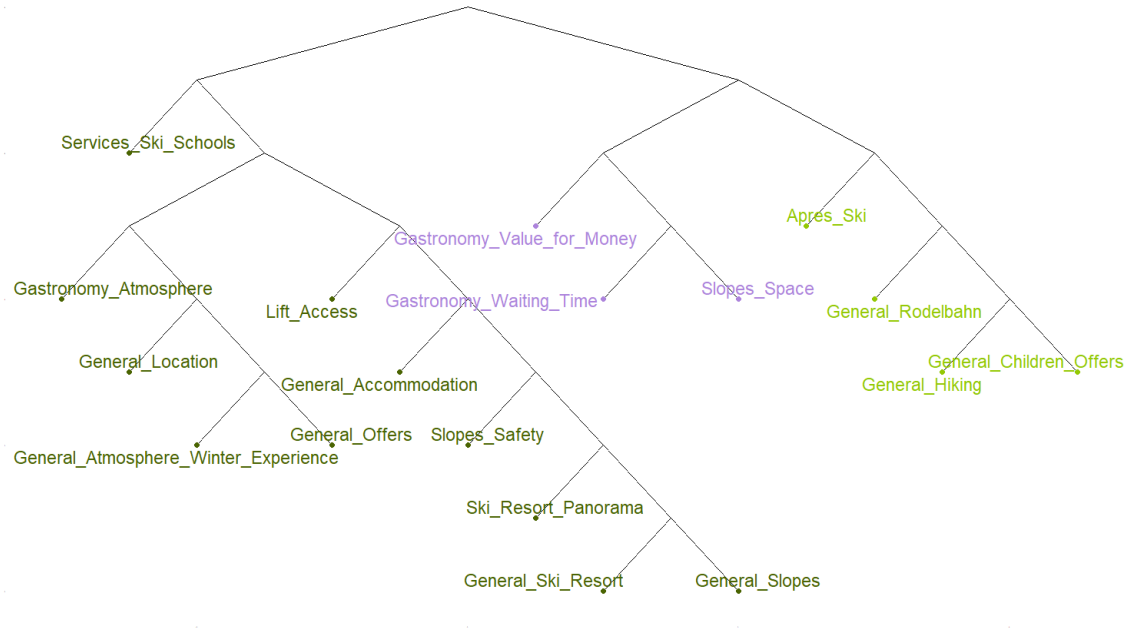


FIGURE 25: PHYLOGENIC DENDROGRAM OF THE VARIABLES CLUSTER ANALYSIS.

Figure 25 makes it easier to understand how the grouping happened to identify which variables have the closest score relationship and how the groups appeared and increased in size. The cluster analysis procedure does not test the hypothesis and assumes only a descriptive function. Therefore, it is not possible to compare clusters.

According to Figure 25 and Figure 23, the reader can see the development of the groups. Clusters 2 and 3 were created by adding the attribute one at a time. In contrast, cluster one was made from two sub-clusters and finally by adding the "Services Ski Schools" factor at the very end of the clustering analysis. The three clusters created represent the grouping of the most influential characteristics of the ski resorts. These attributes split can be used to identify the topics valuable for customers. For example, cluster 1 is all about skiing adventure, cluster 2 is about not skiing experiences, and cluster three represents usability and gastronomy-related issues.

After summarizing the results of the data analysis in this chapter, the following one concludes this research and outlines the potential for further use of the results and alternative ways to study the topic inspired by the results of this study.

5 CONCLUSION

As a result of the research, the most influential factors of the ski areas in Austria are identified. After that, based on the specified variables, the customers are grouped in the data-driven “a-posteriori” way. Also, the influential variables are clustered to see which characteristics are evaluated similarly.

The multiple linear regression identified "General Ski Resort," "Gastronomy Atmosphere," "Gastronomy Waiting Time," "General Offers," "General Atmosphere Winter Experience," "General Location," "General Accommodation," "Ski Resort Panorama," "Apres Ski," "Services Ski Schools," "Gastronomy Value for Money," "General Hiking," "General Slopes," "Lift Access," "General Children Offers," "Slopes Space," "General Rodelbahn" and "Slopes Safety" attributes of alpine ski resorts as significantly influential. Three primary characteristics greatly influence the customers' overall satisfaction and high estimation coefficients (Table 6). Those are “General Accommodation,” which increases the general satisfaction by 0.24 when it is increased by 1, and the “General Offers” parameter, which increases the general satisfaction by 0.17 when it is increased by 1. And finally “General Location” parameter increases the overall satisfaction by 0.17 when it is increased by 1. Therefore, accommodation offers and location are the most important attributes to focus the attention of the alpine ski resorts. While the location factor can be considered an uncontrolled factor, not all the respondents may see the parameter as a solely geographical one, which gives some ways for further improvement. Especially when the winter experience-related factor is so close in the evaluation and is the one having the next closest estimation coefficient of 0.14 (Table 6). The “General Location” factor includes various other factors about the location other than its physical geographical location. The other two parameters, accommodation and offers, are the parameters under the control of the ski resorts.

Comparing the results of the multiple linear regression output to the three most influential aspects identified by Mazanec (gastronomy, slopes, location), the presence of all three aspects is visible in the outcomes (Mazanec, 2007). However, instead of gastronomy and slopes, general accommodation and offers factors round out the top three aspects of this study. In the work of Mazanec, other parameters of the ski resort were the evaluations of access, ticket office, prices, lift, and service(Mazanec, 2007). Therefore the accommodation parameter was either not considered or was a part of one of the mentioned groups. Also, while the level and variety of prices can be seen as a factor comparative to general offers, it can be assumed that those talk about different things, one narrowing more into prices and the other into a variety of offers. Therefore, comparing them can be untrustworthy.

The following cluster analysis of the participants returned the result, which divides the customers into groups based on their level of satisfaction without multidimensional differences. Meaning there are no groups with different satisfaction profiles except for the level of satisfaction.

After that, the cluster analysis with the transposed dataset groups the variables into three groups. While two smaller groups have factors that are logically connected with each other, the biggest group has various aspects in it. "Apres Ski," "General Rodelbahn," "General Children Offers," and "General Hiking" are the factors of group 2, which all represent activities outside of the skiing scope. In other words, those are activities not necessarily for the skiers and snowboarders. The plot in Figure 24 provides good visualization of the variables in different segments. Factors of group 3, "Gastronomy Value for Money," "Gastronomy Waiting Time," and "Slopes Space," all contribute either to the amount of space or to gastronomy.

Factors of group 1 are "General Ski Resort", "Gastronomy Atmosphere", "General Offers", "General Atmosphere Winter Experience", "General Location", "General Accommodation", "Ski Resort Panorama", "Services Ski Schools", "General Slopes", "Lift Access", and "Slopes Safety". This group is concentrated around the “General Slope” factor in the middle of cluster 1, as seen in Figure 24. Cluster one represents a triangle and the three corners of which are “Ski Resort Panorama,” “Gastronomy Atmosphere,” and “Services Ski Schools.” Those three extremes are all allocated to the three different subdivisions of cluster 1, as also can be visually identified in Figure 25.

5.1 Contribution to Knowledge

The contribution to the knowledge of the multiple linear regression analysis conducted is much easier seen through this research than the contribution of cluster analysis, as multiple linear regression, in this case, provides the data output that is much easier to interpret. The data provided in Table 6 suggests how the factors affect the customer's overall satisfaction. Therefore, improvements in the parameters with lower p-values and positive estimated coefficients are encouraged. As those parameters affect overall satisfaction, their positive presence in ski area advertisements is also advised. It is also important to mention that improvements in the variables with negative estimation coefficients are worth considering. While the multiple regression outcomes suggest that their improvement will lover the overall satisfaction, it is not logical. As well as lowering the satisfaction with those will not result in an increase in satisfaction. This problem commonly appears in models with many independent predictors with coefficients close to zero.

The cluster analysis procedure for customer segmentation identifies that there are no multidimensional differences between clusters, and they differ only by the level of satisfaction. As cluster analysis did not identify the groups that are satisfied with one variable while being dissatisfied with others, there is not enough variability in the data to differentiate respondents using the customer persona tool. The customers seem somewhat satisfied with all the variables with the approximately same gradual score across groups, as seen in Table 7.

The cluster analysis procedure for the variables provides interesting outputs showing the similarity of scoring patterns between the most influential factors and outlines the distance between the customer's feedback (Figure 25). While clustering of the questions cannot help to group customers, the data of this clustering procedure identifies which factors are evaluated similarly and which are not and creates three groups of characteristics that all significantly influence the overall satisfaction but differ in their purpose. The first cluster is all about adventure-related characteristics. The second cluster is about alternative leisure activities in the ski resort, and the third is about gastronomy and usability. Those three topics can be used as three main pillars on which to base the ski resorts' advertisements.

5.2 Implications for Relevant Stakeholders

The core stakeholders in this research are the ski area management, the data collector firm, and the ski area customers. As a result of this research, ski area management receives valuable data, potentially influencing ski resorts' further development and decisions. After the implementation period, the data collector firm gets an instrument to use, and customers are expected to profit by getting better service. But there are indirect stakeholders, such as the country and its economy. Austria heavily relies on the winter tourism areas such as alpine ski resorts, and what benefits the businesses, in the long run, benefits the economy.

As of 2022, prices for skiing tourism are likely to go up because ski operators are facing an increase in their expenditures (Maguire, 2022). This issue is also mitigated with the help of energy-saving strategies (Maguire, 2022). While increasing prices may be the option for all ski resorts, this decision can decrease customer satisfaction with offers. The better option would be to mix the price increase with better competitive advantages and implement different advertisement and investment strategies, by focusing on the topics important to the customers and increasing their overall satisfaction with their holidays. Therefore, making it visible to customers that prices did not just increase but they are getting a better product or service.

The second cluster analysis procedure results can help managers of the ski areas advertise the alpine ski resorts according to the three topics identified: ski adventure, alternative leisure activities, and gastronomy and usability.

5.3 Future Research

This research is done in a way that it can be relatively quickly repeated with other datasets, including other season data, location data, or both. This can mean other countries, a particular resort, a group of resorts, or a particular ski region. Therefore, for the future research author can advise extending the time and location frame. It would be interesting to compare the results of this research to similar research from other countries, such as Finland, the United States of America, and especially nearby countries such as Italy, Switzerland, Czech Republic, Slovakia,

France, and Poland. Also, the results of the different seasons inside Austria can be interesting. The timeseries comparison of the seasons before, during, and after the COVID-19 pandemic would be interesting to observe the shifts in customer satisfaction, if there are any.

The next idea would be to filter the data by some parameters like “size of the ski resort” and conduct similar research for small, medium-sized, and big resorts and afterward compare if there are differences and how big they are.

When conducting this research author had difficulties with data being measured on an ordinal scale. Therefore, it is wise to offer an option for further research to conduct ordinal regression and compare the results.

Another idea would be to offer a shorter survey to half of the respondents, which will include not all the satisfaction aspects but only the ones found influential in this research. Further comparison of the results may be interesting, as the shorter length of the survey may introduce the opinions of the people who don’t have time to participate in the more extended version.

As an alternative approach, future research can focus on adjusting the data gathering to the newly introduced parameters, such as health safety precautions at the ski resort. Those may be, from now on, always required by the law and should be beneficial for the customers. But the whole experience can be ruined if those are implemented in the wrong way. Therefore a satisfaction-based study of those aspects is recommended.

BIBLIOGRAPHY

Ahmed, S. A., Barber, M., & D'astous, A. (1997). Segmentation of the Nordic Winter Sun Seekers Market. Journal of Travel & Tourism Marketing, 7(1), 39–63. https://doi.org/10.1300/j073v07n01_04

Anekwe, L. (2021). Inductive reasoning. New Scientist. https://www.newscientist.com/definition/inductive-reasoning/

Bearden, W. O., & Teel, J. E. (1983). Selected Determinants of Consumer Satisfaction and Complaint Reports. Journal of Marketing Research, 20(1), 21. https://doi.org/10.2307/3151408

Berends, M. S., Luz, C. F., Friedrich, A. W., Sinha, B. N. M., Albers, C. J., & Glasner, C. (2022). AMR: An R Package for Working with Antimicrobial Resistance Data. Journal of Statistical Software, 104(3). https://doi.org/10.18637/jss.v104.i03

Berger, P. L., & Luckmann, T. (1967). The Social Construction of Reality: A Treatise in the Sociology of Knowledge. Sociological Analysis, 28(1), 55. https://doi.org/10.2307/3710424

Bhattacherjee, A. (2012). Social Science Research: Principles, Methods, and Practices. Textbooks Collection. https://scholarcommons.usf.edu/oa_textbooks/3

Bradley, C., Oliveira, L., Birrell, S., & Cain, R. (2021). A New Perspective on Personas and Customer Journey Maps: Proposing Systemic UX. International Journal of Human-Computer Studies, 148, 102583. https://doi.org/10.1016/j.ijhcs.2021.102583

Broberg, O., Andersen, V., & Seim, R. (2011). Participatory Ergonomics in Design Processes: The Role of Boundary Objects. Applied Ergonomics, 42(3), 464–472. https://doi.org/10.1016/j.apergo.2010.09.006

Cadotte, E. R., Woodruff, R. B., & Jenkins, R. L. (1987). Expectations and Norms in Models of Consumer Satisfaction. Journal of Marketing Research, 24(3), 305–314. https://doi.org/10.1177/002224378702400307

Campos Rodrigues, L., Freire-González, J., González Puig, A., & Puig-Ventosa, I. (2018). Climate Change Adaptation of Alpine Ski Tourism in Spain. Climate, 6(2), 29. https://doi.org/10.3390/cli6020029

Cassidy-Smith, T. N., Baumann, B. M., & Boudreaux, E. D. (2007). The Disconfirmation Paradigm: Throughput Times and Emergency Department Patient Satisfaction. The Journal of Emergency Medicine, 32(1), 7–13. https://doi.org/10.1016/j.jemermed.2006.05.028

Churchill, G. A., & Surprenant, C. (1982). An Investigation into the Determinants of Customer Satisfaction. Journal of Marketing Research, 19(4), 491. https://doi.org/10.2307/3151722

Cleave, P. (2019, November 14). 6 Ways to Measure Customer Satisfaction. SmartSurvey. https://www.smartsurvey.co.uk/blog/6-ways-to-measure-customer-satisfaction

Cooper, A. (1999). The Inmates are Running the Asylum. Berichte Des German Chapter of the ACM, 17–17. https://doi.org/10.1007/978-3-322-99786-9_1

Creswell, J. W., & Creswell, D. J. (2018). Research Resign : Qualitative, Quantitative, and Mixed Methods Approaches. Sage Publications, Inc.

Danaher, P. J., & Haddrell, V. (1996). A Comparison of Question Scales Used for Measuring Customer Satisfaction. International Journal of Service Industry Management, 7(4), 4–26. https://doi.org/10.1108/09564239610129922

Danaher, P. J., & Mattsson, J. (1994). Customer Satisfaction during the Service Delivery Process. European Journal of Marketing, 28(5), 5–16. https://doi.org/10.1108/03090569410062005

Devlin, S. J., Dong, H. K., & Brown, M. (1993). Selecting a Scale for measuring Quality. Marketing Research: A Magazine of Management and Applications, 5(3), 12–17.

Dolnicar, S., Grabler, K., & Mazanec, J. A. (1999). Analyzing Destination Images: A Perceptual Charting Approach. Journal of Travel & Tourism Marketing, 8(4), 43–57. https://doi.org/10.1300/j073v08n04_03

Dolnicar, S., & Leisch, F. (2003a). Winter Tourist Segments in Austria: Identifying Stable Vacation Styles Using Bagged Clustering Techniques. Journal of Travel Research, 41(3), 281–292. https://doi.org/10.1177/0047287502239037

Dolnicar, S., & Leisch, F. (2003b). Data-Driven Market Segmentation - A Structure-Based Conceptual Framework for Management Decision Support. Faculty of Commerce - Papers (Archive). https://ro.uow.edu.au/commpapers/319/

Elsasser, H., & Messerli, P. (2001). The Vulnerability of the Snow Industry in the Swiss Alps. Mountain Research and Development, 21(4), 335–339. https://doi.org/10.1659/0276-4741(2001)021[0335:tvotsi]2.0.co;2

Everitt, B. S., Landau, S., Leese, M., & Stahl, D. (2011). Cluster Analysis. Wiley.

Flores-Sosa, M., Avilés-Ochoa, E., Merigó, J. M., & Kacprzyk, J. (2022). The OWA Operator in Multiple Linear Regression. Applied Soft Computing, 124, 108985. https://doi.org/10.1016/j.asoc.2022.108985

Formica, S., & Uysal, M. (2001). Segmentation of Travelers Based on Environmental Attitudes. Journal of Hospitality & Leisure Marketing, 9(3-4), 35–49. https://doi.org/10.1300/j150v09n03_04

Fornell, C. (1992). A National Customer Satisfaction Barometer: the Swedish Experience. Journal of Marketing, 56(1), 6–21. https://doi.org/10.1177/002224299205600103

Frank, B., Herbas Torrico, B., Enkawa, T., & Schvaneveldt, S. J. (2014). Affect versus Cognition in the Chain from Perceived Quality to Customer Loyalty: The Roles of Product Beliefs and Experience. Journal of Retailing, 90(4), 567–586. https://doi.org/10.1016/j.jretai.2014.08.001

Füller, J., & Matzler, K. (2008). Customer Delight and Market Segmentation: An Application of the Three-factor Theory of Customer Satisfaction on Lifestyle Groups. Tourism Management, 29(1), 116–126. https://doi.org/10.1016/j.tourman.2007.03.021

Galili, T., Benjamini, Y., Simpson, G., Jefferis, G., Gallotta, M., Renaudie, J., The R Core Team, Hornik, K., Ligges, U., Spiess, A.-N., Horvath, S., Langfelder, P., Van Der Loo, M., de Vries, A., Gu, Z., Ma, J., G, K., Hummel, M., Clark, C., & Graybuck, L. (2021, October 28). dendextend: Extending "dendrogram" Functionality in R. R-Packages. https://cran.r-project.org/web/packages/dendextend/index.html

Gawel, J. (1996). Herzberg's Theory of Motivation and Maslow's Hierarchy of Needs. Practical Assessment, Research, and Evaluation, 5, 11. https://doi.org/10.7275/31qy-ea53

Gerson, R. F. (1993). Measuring Customer Satisfaction: A Guide to Managing Quality Service. Crisp Publications.

Giese, J. L., & Cote, J. A. (2000). Defining Consumer Satisfaction. Academy of Marketing Science Review.

Goodwin, K. (2009). Designing for the Digital Age [electronic resource] : How to Create Human-Centered Products and Services. John Wiley & Sons.

Gössling, S., Scott, D., Hall, C. M., Ceron, J.-P., & Dubois, G. (2012). Consumer Behaviour and Demand Response of Tourists to Climate Change. Annals of Tourism Research, 39(1), 36–58. https://doi.org/10.1016/j.annals.2011.11.002

Grigoroudis, E., & Siskos, Y. (2010). Customer Satisfaction Evaluation. In International Series in Operations Research & Management Science. Springer US. https://doi.org/10.1007/978-1-4419-1640-2

Hallowell, R. (1996). The Relationships of Customer Satisfaction, Customer loyalty, and Profitability: An Empirical Study. International Journal of Service Industry Management, 7(4), 27–42. https://doi.org/10.1108/09564239610129931

Haugom, E., & Malasevska, I. (2019). The Relative Importance of Ski Resort- and Weather-related Characteristics when going Alpine Skiing. Cogent Social Sciences, 5(1), 1681246. https://doi.org/10.1080/23311886.2019.1681246

Heinzen, E., Sinnwell, J., Atkinson, E., Gunderson, T., Dougherty, G., Votruba, P., Lennon, R., Hanson, A., Goergen, K., Lundt, E., Broderick, B., & McCullough, M. (2021, June 4). arsenal: An Arsenal of "R" Functions for Large-Scale Statistical

Summaries. R-Packages. https://cran.r-project.org/package=arsenal

Hill, N., & Alexander, J. (2006). The Handbook of Customer Satisfaction and Loyalty Measurement. Gower Publishing Ltd.

Hudson, S. (2000). The Segmentation of Potential Tourists: Constraint Differences between Men and Women. Journal of Travel Research, 38(4), 363–368. https://doi.org/10.1177/004728750003800404

Hunt, H. K. (1977). Conceptualization and Measurement of Consumer Satisfaction and Dissatisfaction. Marketing Science Institute.

IBM. (2021a). K-Means Cluster Analysis. Www.ibm.com. https://www.ibm.com/docs/en/spss-statistics/29.0.0?topic=edition-k-means-cluster-analysis

IBM. (2021b, September 11). Cluster Analysis. IBM. https://www.ibm.com/docs/en/spss-statistics/28.0.0?topic=edition-cluster-analysis

IBM. (2022). About Linear Regression | IBM. IBM. https://www.ibm.com/topics/linear-regression

Jang, S. C., Morrison, A. M., & O'Leary, J. T. (2002). Benefit Segmentation of Japanese Pleasure Travelers to the USA and Canada: selecting Target Markets based on the Profitability and Risk of Individual Market Segments. Tourism Management, 23(4), 367–378. https://doi.org/10.1016/S0261-5177(01)00096-6

Jean Philippe Deschamps, & P Ranganath Nayak. (1995). Product Juggernauts: How Companies Mobilize to Generate a Stream of Market Winners. Harvard Business School Press.

Kandampully, J., & Suhartanto, D. (2003). The Role of Customer Satisfaction and Image in Gaining Customer Loyalty in the Hotel Industry. Journal of Hospitality & Leisure Marketing, 10(1-2), 3–25. https://doi.org/10.1300/j150v10n01_02

Kassambara, A. (2017). Multivariate Analysis I Practical Guide To Cluster Analysis in R Unsupervised Machine Learning. STHDA. http://www.sthda.com/

Kassambara, A., & Mundt, F. (2020, April 1). factoextra: Extract and Visualize the Results of Multivariate Data Analyses. R-Packages. https://cran.r-project.org/web/packages/factoextra/index.html

Konu, H., Laukkanen, T., & Komppula, R. (2011). Using Ski Destination Choice Criteria to Segment Finnish Ski Resort Customers. Tourism Management, 32(5), 1096–1105. https://doi.org/10.1016/j.tourman.2010.09.010

Kotler, P. (1994). Analysis, Planning, Implementation and Control. Prentice Hall International.

Kumar, V., & Shah, D. (2004). Building and Sustaining Profitable Customer Loyalty for the 21st Century. Journal of Retailing, 80(4), 317–329. https://doi.org/10.1016/j.jretai.2004.10.007

Larmarange, J., Ludecke, D., Wickham, H., Bojanowski, M., & Briatte, F. (2022, September 14). labelled: Manipulating Labelled Data. R-Packages. https://cran.r-project.org/web/packages/labelled/index.html

Lentz, M., Berezan, O., & Raab, C. (2021). Uncovering the Relationship Between Revenue Management and Hotel Loyalty Programs. Journal of Revenue and Pricing Management. https://doi.org/10.1057/s41272-021-00331-0

Likert, R. (1932). A Technique for the Measurement of Attitudes. Archives of Psychology, 22(140), 5–55.

Lincoln, Y. S., & Guba, E. G. (1985). Naturalistic Inquiry. Sage Publications.

Lüdecke, D. (2021, July 10). Data Visualization for Statistics in Social Science [R package sjPlot version 2.8.9]. Cran.r-Project.org. https://CRAN.R-project.org/package=sjPlot

Lundberg, C., Gudmundson, A., & Andersson, T. D. (2009). Herzberg's Two-Factor Theory of Work Motivation Tested Empirically on Seasonal Workers in Hospitality and Tourism. Tourism Management, 30(6), 890–899. https://doi.org/10.1016/j.tourman.2008.12.003

Mackrill, J., Marshall, P., Payne, S. R., Dimitrokali, E., & Cain, R. (2017). Using a Bespoke Situated Digital Kiosk to Encourage User Participation in Healthcare Environment Design. Applied Ergonomics, 59, 342–356. https://doi.org/10.1016/j.apergo.2016.08.005

Maechler, M., original), P. R. (Fortran, original), A. S. (S, original), M. H. (S, Hornik [trl, K., maintenance(1999-2000)), ctb] (port to R., Studer, M., Roudier, P., Gonzalez, J., Kozlowski, K., pam()), E. S. (fastpam options for, & Murphy (volume.ellipsoid({d >= 3})), K. (2019, June 19). cluster: "Finding Groups in Data": Cluster Analysis Extended Rousseeuw et

al. R-Packages. https://cran.r-project.org/web/packages/cluster/index.html

Maguire, H. (2022, October 11). What to Expect from the Ski Season in Austria this Winter. The Local Austria. https://www.thelocal.at/20221011/what-to-expect-from-the-ski-season-in-austria-this-winter/

Maslow, A. H. (1954). Motivation and Personality. Harper & Row.

Massnick, F. (1997). The Customer Is CEO: How to Measure What Your Customers Want — and Make Sure They Get It. Amacom, American Management Assoc.

Matsuoka, K. (2022). Effects of Revenue Management on Perceived Value, Customer Satisfaction, and Customer Loyalty. Journal of Business Research, 148, 131–148. https://doi.org/10.1016/j.jbusres.2022.04.052

Matzler, K., Fuchs, M., & Schubert, A. (2004). Employee Satisfaction: Does Kano's Model Apply? Total Quality Management & Business Excellence, 15(9-10), 1179–1198. https://doi.org/10.1080/1478336042000255569

Mazanec, J. A. (2007). Exploring Tourist Satisfaction with Nonlinear Structural Equation Modeling and Inferred Causation Analysis. Journal of Travel & Tourism Marketing, 21(4), 73–90. https://doi.org/10.1300/j073v21n04_06

Mertens, D. M. (2010). Research and Evaluation in Education and Psychology : Integrating Diversity with Quantitative, Qualitative, and Mixed Methods. Sage.

Miaskiewicz, T., & Kozar, K. A. (2011). Personas and User-centered Design: How can Personas Benefit Product Design Processes? Design Studies, 32(5), 417–430. https://doi.org/10.1016/j.destud.2011.03.003

Middleton, V. T. C., & Clarke, J. R. (2012). Marketing in Travel and Tourism. Taylor And Francis.

Mills, A. S., Couturier, H., & Snepenger, D. J. (1986). Segmenting Texas Snow Skiers. Journal of Travel Research, 25(2), 19–23. https://doi.org/10.1177/004728758602500204

Miro. (2022). About Miro | Meet the team | Our mission. Https://Miro.com/. https://miro.com/about/

Mkpojiogu, E. O. C., & Hashim, N. L. (2016). Understanding the Relationship Between Kano Model's Customer Satisfaction Scores and Self-stated Requirements Importance. SpringerPlus, 5(1). https://doi.org/10.1186/s40064-016-1860-y

Murphy, A. B. (2022). The Modernist Territorial Order and the Ukraine crisis. Political Geography, 102667. https://doi.org/10.1016/j.polgeo.2022.102667

Musyoka, S. M., Mutyauvyu, S. M., Kiema, J. B. K., Karanja, F. N., & Siriba, D. N. (2007). Market Segmentation using Geographic Information Systems (GIS). Marketing Intelligence & Planning, 25(6), 632–642. https://doi.org/10.1108/02634500710819987

Oliver, R. L. (1977). Effect of Expectation and Disconfirmation on Postexposure Product Evaluations: An Alternative Interpretation. Journal of Applied Psychology, 62(4), 480–486. https://doi.org/10.1037/0021-9010.62.4.480

Oliver, R. L. (1980). A Cognitive Model of the Antecedents and Consequences of Satisfaction Decisions. Journal of Marketing Research, 17(4), 460. https://doi.org/10.2307/3150499

Oliver, R. L. (1992). An Investigation of the Attribute Basis of Emotion and Related Affects in Consumption: Suggestions For a Stage-Specific Satisfaction Framework. NA - Advances in Consumer Research, 19, 237–244. https://www.acrwebsite.org/volumes/7302/volumes/v19/NA-19

Oliver, R. L. (1997). Satisfaction: A Behavioral Perspective on the Consumer. Mcgraw Hill.

Øvad, T. (2020, December 1). Likert Scale & Semantic Differential Scale. Preely. https://preely.com/likert-scale-semantic-differential-scale/#:~:text=A%20Likert%20scale%20will%20provide

Palmatier, R. W., Dant, R. P., Grewal, D., & Evans, K. R. (2006). Factors Influencing the Effectiveness of Relationship Marketing: A Meta-Analysis. Journal of Marketing, 70(4), 136–153. https://doi.org/10.1509/jmkg.70.4.136

Parasuraman, A., & Zeithaml, V. (1988). SERVQUAL: A Multiple-Item Scale for Measuring Consumer Perceptions of Service Quality. Journal of Retailing, 64(1).

Park, D.-B., & Yoon, Y.-S. (2009). Segmentation by Motivation in Rural Tourism: A Korean Case Study. Tourism Management, 30(1), 99–108. https://doi.org/10.1016/j.tourman.2008.03.011

Perdue, R. R. (2004). Sustainable Tourism and Stakeholder Groups: A Case Study of Colorado Ski Resort Communities. Consumer Psychology of Tourism, Hospitality and Leisure, 253–264. https://doi.org/10.1079/9780851997490.0253

Peterson, R. A., & Wilson, W. R. (1992). Measuring Customer Satisfaction: Fact and Artifact. Journal of the Academy of Marketing Science, 20(1), 61–71. https://doi.org/10.1007/bf02723476

Prettenthaler, F., & Formayer, H. (2011). Tourismus im Klimawandel: Zur regionalwirtschaftlichen Bedeutung des Klimawandels für die österreichischen Tourismusgemeinden. Studien zum Klimawandel in Österreich (Vol. 6). Verlag der Österreichischen Akademie der Wissenschaften.

Price, M. F., Borowski, D., & Macleod, C. (2011). From Rio 1992 to 2012 and beyond: 20 years of Sustainable Mountain Development. In Weadapt. https://www.weadapt.org/sites/weadapt.org/files/2017/may/smd_alps_final_20120228_rio_alps.pdf

Pröbstl-Haider, U., Mostegl, N., & Damm, A. (2021). Tourism and Climate Change – A Discussion of Suitable Strategies for Austria. Journal of Outdoor Recreation and Tourism, 34, 100394. https://doi.org/10.1016/j.jort.2021.100394

Pröbstl-Haider, U., Wanner, A., Feilhammer, M., & Damm, A. (2021). Tourism and Climate Change – An Integrated Look at the Austrian Case. Journal of Outdoor Recreation and Tourism, 34. https://doi.org/10.1016/j.jort.2020.100361

QuestionPro Survey Software. (2021). Semantic Differential Scale, Example, and Question Types | QuestionPro. Www.questionpro.com. https://www.questionpro.com/semantic-differential-scale.html#:~:text=A%20semantic%20differential%20scale%20is

R Core Team. (2019). R: The R Project for Statistical Computing. R-Project.org. https://www.r-project.org/

RStudio. (2019, February 7). Download RStudio - RStudio. RStudio. https://www.rstudio.com/products/rstudio/download/

RStudio. (2021). RStudio Benefit Corporation. In RStudio. https://www.rstudio.com/assets/img/RStudio-BenefitCorporation-2021-Annual-Report.pdf

Rust, R. T., Zahorik, A. J., & Keiningham, T. L. (1994). Return on Quality : Measuring the Financial Impact of your Company's Quest for Quality. Probus.

Sauro, J. (2019, October 2). Five Scales to Measure Customer Satisfaction. MeasuringU. https://measuringu.com/csat-scales/

Singh Minhas, R., & Jacobs, E. M. (1996). Benefit Segmentation by Factor Analysis: An Improved Method of Targeting Customers for Financial Services. International Journal of Bank Marketing, 14(3), 3–13. https://doi.org/10.1108/02652329610113126

Slife, B. D., & Williams, R. N. (1995). What's Behind the Research? : Discovering Hidden Assumptions in the Behavioral Sciences. Sage Publications.

Statistics Austria. (2019). Tourismus in ¨Osterreich 2018. Ergebnisse der Beherbergungsstatistik. Bundesanstalt Statistik Österreich. https://www.statistik.at/web_de/services/publikationen/13/index.html?includePage=detailedView&sectionName=Tourismus&pubId=569

Steiger, R., & Scott, D. (2020). Ski Tourism in a Warmer World: Increased Adaptation and Regional Economic Impacts in Austria. Tourism Management, 77, 104032. https://doi.org/10.1016/j.tourman.2019.104032

Taormina, R. J., & Gao, J. H. (2013). Maslow and the Motivation Hierarchy: Measuring Satisfaction of the Needs. The American Journal of Psychology, 126(2), 155. https://doi.org/10.5406/amerjpsyc.126.2.0155

The R Foundation. (2019). R: What is R? R-Project.org. https://www.r-project.org/about.html

Tse, D. K., & Wilton, P. C. (1988). Models of Consumer Satisfaction Formation: An Extension. Journal of Marketing Research, 25(2), 204–212. https://doi.org/10.2307/3172652

Tsiotsou, R. (2006). Using Visit Frequency to Segment Ski Resorts Customers. Journal of Vacation Marketing, 12(1), 15–26. https://doi.org/10.1177/1356766706059029

UNIX Full Form. (2020, April 16). GeeksforGeeks. https://www.geeksforgeeks.org/unix-full-form/#:~:text=UNIX%20was%20earlier%20known%20to

Vanat, L. (2020). 2020 International Report on Snow & Mountain Tourism. Overview of the key industry figures for ski resorts. In vanat.ch. https://www.vanat.ch/RM-world-report-2020.pdf

Wedel, M., & Kamakura, W. (2002). Introduction to the Special Issue on Market Segmentation. International Journal of Research in Marketing, 19(3), 181–183. https://doi.org/10.1016/s0167-8116(02)00075-7

Westbrook, R. A., & Reilly, M. D. (1983). Value-Percept Disparity: An Alternative to the Disconfirmation of Expectations Theory of Consumer Satisfaction. ACR North American Advances., 10(1).

Wickham, H., Chang, W., Henry, L., Pedersen, T. L., Takahashi, K., Wilke, C., Woo, K., Yutani, H., Dunnington, D., & RStudio. (2019). Create Elegant Data Visualisations Using the Grammar of Graphics [R package ggplot2 version 3.2.1]. R-Project.org. https://doi.org/https://CRAN.R-project.org/package=ggplot2

Wickham, H., François, R., Henry, L., Müller, K., & RStudio. (2020, August 18). dplyr: A Grammar of Data Manipulation. R-Packages. https://cran.r-project.org/web/packages/dplyr/index.html

Wickham, H., Miller, E., Smith, D., & RStudio. (2019). Package “haven” Title Import and Export “SPSS”, “Stata” and “SAS” Files. https://cran.r-project.org/web/packages/haven/haven.pdf

Wild, R. (1980). Operations management: A policy framework. Pergamon Press.

Wirtschaftskammer Österreich. (2022, August 8). Zahlen/Daten/Fakten. Www.wko.at. https://www.wko.at/branchen/transport-verkehr/seilbahnen/ZahlenDatenFakten.html

Yan, W., Chen, C.-H., & Pheng Khoo, L. (2007). Identification of Different Demographical Customer Preferences for Product Conceptualization. Journal of Engineering Design, 18(1), 39–54. https://doi.org/10.1080/09544820600664952

Yang, J., Huang, B., Yang, Q., & Zhou, Y. (2022). Impact of the US–China Trade War on Resource Allocation: Evidence From China's Land Supply. China Economic Review, 101854. https://doi.org/10.1016/j.chieco.2022.101854

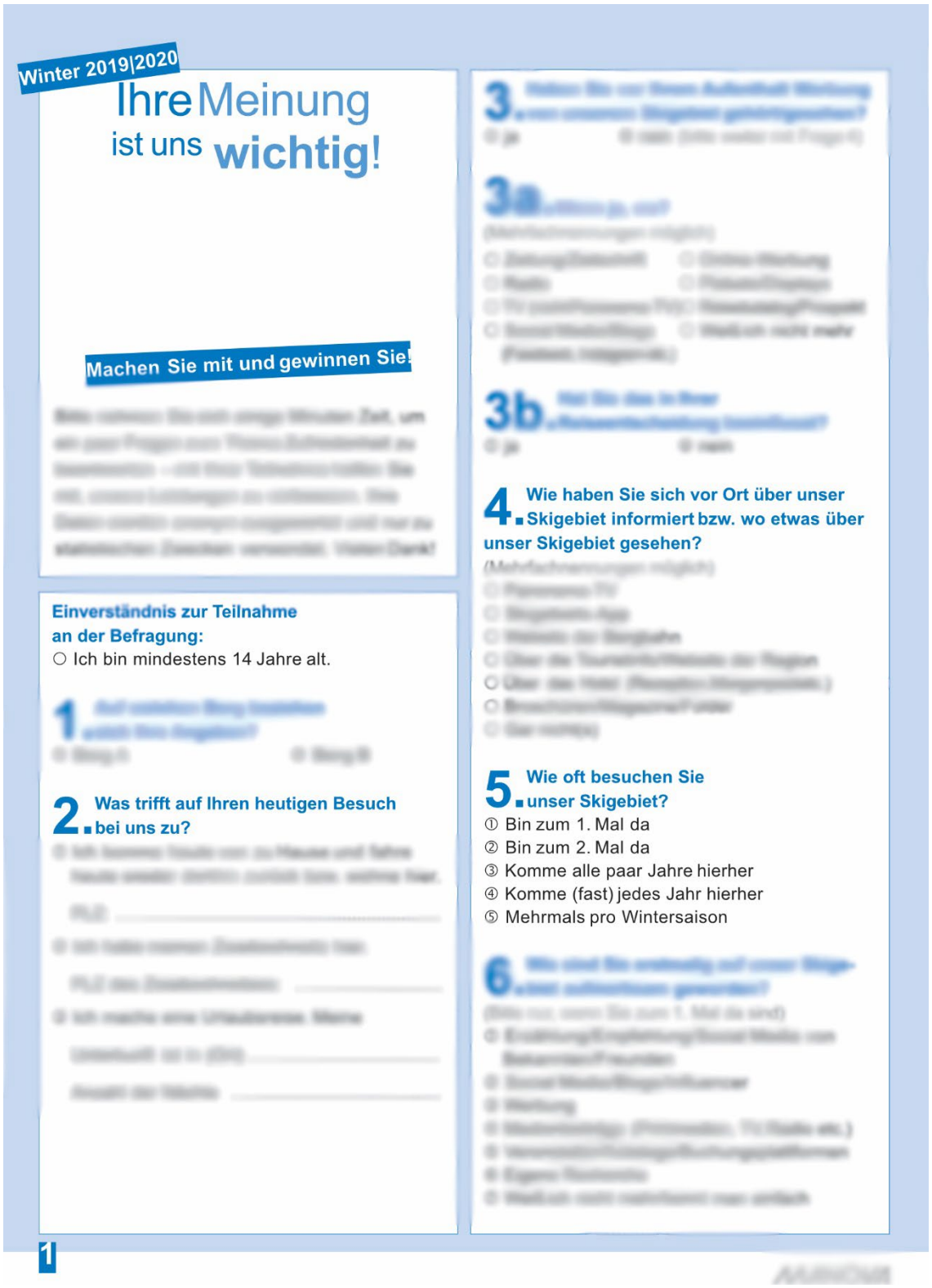
Zeithaml, V. A., Berry, L. L., & Parasuraman, A. (1996). The Behavioral Consequences of Service Quality. Journal of Marketing, 60(2), 31. https://doi.org/10.2307/1251929

Zografos, C., & Allcroft, D. (2007). The Environmental Values of Potential Ecotourists: A Segmentation Study. Journal of Sustainable Tourism, 15(1), 44–66. https://doi.org/10.2167/jost572.0

APPENDICES

Appendix 1: The DEMO survey

The original survey example consisted of 29 questions, some of which had multiple choice answers (e.g., 11), some Likert scaled (e.g., 12), some semantic differential scaled (e.g., 13-22), and some open-ended. To be able to participate, you were required to be at least 14 years old. This survey example is partly provided in the following figures (Figure 1-Figure 6). Almost all the questions and answers that were not used in this research are blurred. An exception is given for general questions and a few specific questions and answers to provide an overview of how many aspects were considered during the creation of this survey. While the main interest of this research was looking at semantic differential scaled questions, and those were translated and discussed in the methodology chapter, therefore, here they are also partly blurred.



Winter 2019|2020

Ihre Meinung ist uns wichtig!

Machen Sie mit und gewinnen Sie!

Bitte nehmen Sie sich einige Minuten Zeit, um ein paar Fragen zum Thema Schneesicherheit zu beantworten – und Ihre Teilnahme hatten Sie evtl. unseren Lotsinnen zu verdanken. Ihre Daten werden anonym ausgewertet und nur zu statistischen Zwecken verwendet. Vielen Dank!

Einverständnis zur Teilnahme an der Befragung:

○ Ich bin mindestens 14 Jahre alt.

1 Auf welchem Berg befinden wir Sie gerade?

○ Berg A ○ Berg B

2 Was trifft auf Ihren heutigen Besuch bei uns zu?

① Ich komme heute von zu Hause und fahre heute wieder dorthin zurück bzw. wohne hier.

PLZ _____

② Ich habe meinen Zweitwohnsitz hier.

PLZ des Zweitwohnsitzes _____

③ Ich mache eine Urlaubsreise. Meine Unterkunft ist in (Ort) _____

Anzahl der Nächte _____

3 Haben Sie vor Ihrem Aufenthalt Werbung von unserem Skigebiet gesehen?

○ ja ○ nein (bitte weiter mit Frage 4)

3a Wenn ja, wo?

(Mehrfachnennungen möglich)

○ Zeitung/Zeitschrift ○ Online-Werbung
○ Radio ○ Plakate/Displays
○ TV (zum/Panorama TV) ○ Reisekatalog/Prospekt
○ Social Media/Blogs ○ Weiß ich nicht mehr (Facebook, Instagram etc.)

3b Hat Sie das in Ihrer Reiseentscheidung beeinflusst?

○ ja ○ nein

4 Wie haben Sie sich vor Ort über unser Skigebiet informiert bzw. wo etwas über unser Skigebiet gesehen?

(Mehrfachnennungen möglich)

○ Panorama TV
○ Skigebiets-App
○ Website der Bergbahn
○ Über die Tourismus/Website der Region
○ Über das Hotel (Rezeption/Angestellte)
○ Broschüren/Magazine/Folder
○ Gar nicht(s)

5 Wie oft besuchen Sie unser Skigebiet?

① Bin zum 1. Mal da
② Bin zum 2. Mal da
③ Komme alle paar Jahre hierher
④ Komme (fast) jedes Jahr hierher
⑤ Mehrmals pro Wintersaison

6 Wie sind Sie erstmalig auf unser Skigebiet aufmerksam geworden?

(Bitte nur, wenn Sie zum 1. Mal da sind)

① Erzählung/Empfehlung/Social Media von Bekannten/Freunden
② Social Media/Blogs/Influencer
③ Werbung
④ Medienberichte (Printmedien, TV, Radio etc.)
⑤ Veranstalter/Katalog/Buchungsplattformen
⑥ Eigene Recherche
⑦ Weiß ich nicht mehr/kommt man einfach

1

7. An wie vielen Tagen nutzen Sie die Bergbahn während Ihres Aufenthalts?

(Bitte nur ausfüllen, wenn Sie eine Urlaubsreise machen oder Ihren Zweitwohnsitz hier haben!)

gesamt: _____ Tage)

davon Skitage: _____ Tage)

davon Skitage bei uns: _____ Tage)

8. Welche Art von Skipass verwenden Sie HAUPTSÄCHLICH in unserem Skigebiet?

- Stunden- bzw. Halbtagesskipass
- Tagesskipass
- Mehrtagesskipass für _____ Tage
- Saisonskipass (Jahreskarte
- Fußgängerskipass
- Sonstiges

9. Wer begleitet Sie?

(Mehrfachnennungen möglich)

- Ich absolviere Begleitung
- Partner/in
- Kinder unter 6 Jahren
- Kinder 6-13 Jahre
- Kinder 14-18 Jahre
- Kinder im Erwachsenenalter
- Eltern
- Großeltern, Tante, Onkel etc.
- Freunde, Geschwister etc.
- Organisierte Gruppe

10. Warum haben Sie sich für unser Skigebiet entschieden?

(Bitte nur entscheidende Punkte bei Ihrer Wahl des Skigebietes ankreuzen!)

- Größe des Skigebietes
- Modernität/Komfort der Liftanlagen
- Geringe Wartezeiten
- Schnee-/Pistenverhältnisse
- Schneesicherheit auf den Pisten
- Skipisten
- Unterkunft
- Après Ski
- Talabfahrten)
- Schnee-/Winterlandschaft
- Panorama
- Snowpark
- Freundlichkeit der Mitarbeiter
- Erreichbarkeit
- Ort (Angebot/Image)
- Skipasspreis
- Kinderangebot
- Hütten/Restaurants
- Freizeitmöglichkeiten
- Überschaubarkeit des Skigebiets

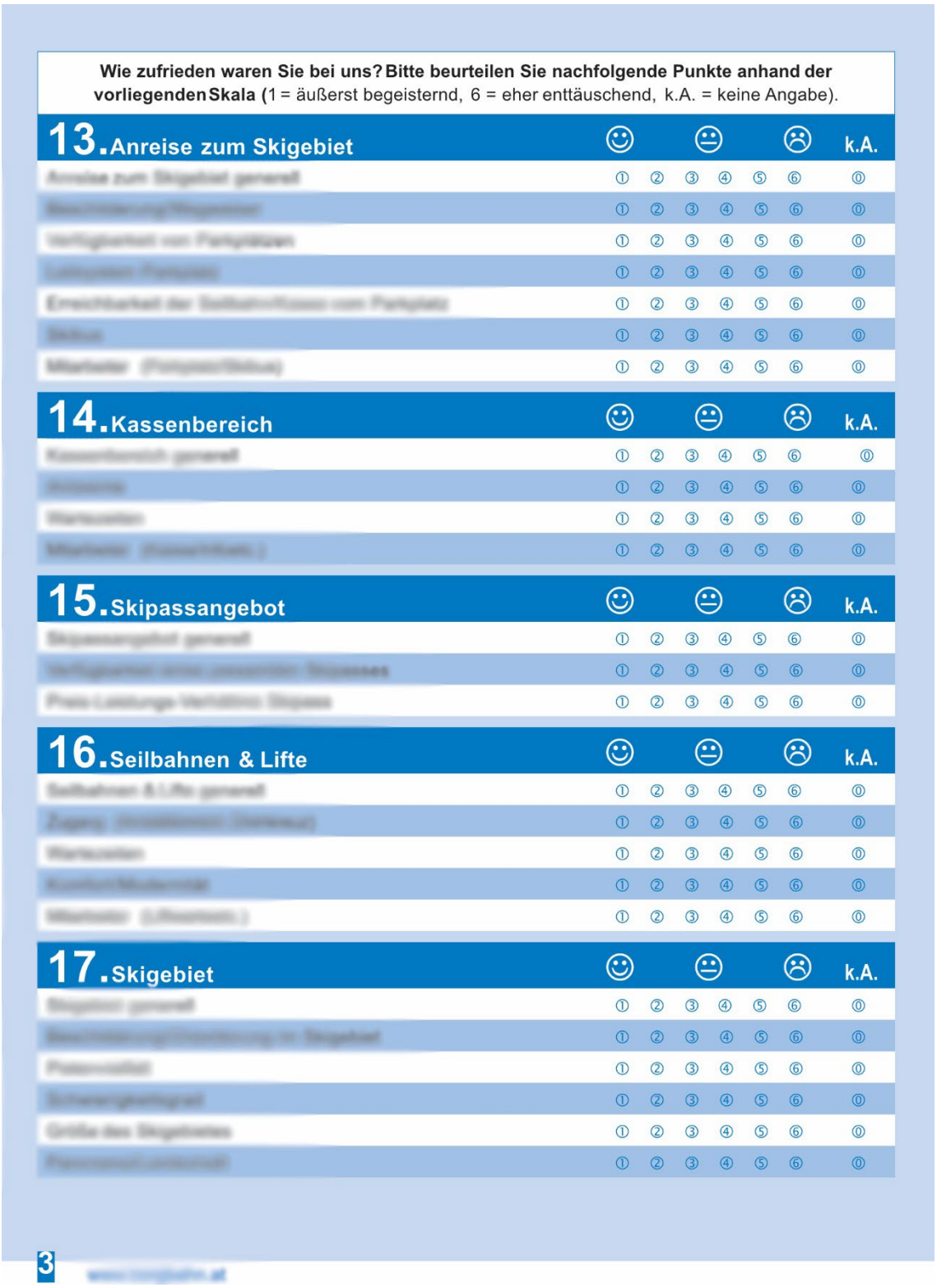
11. Welche Eigenschaften passen auf unser Skigebiet?

(Mehrfachnennungen möglich)

- unterklassig
- modern
- sportlich anspruchsvoll
- erholsam
- exklusiv
- genussvoll
- authentisch und echt
- abwechslungsreich
- überlaufen
- innovativ
- nachhaltig
- familienfreundlich
- teuer
- unspektakulär
- sympathisch
- gastfreundlich
- erlebnisreich
- gemütlich
- trendig
- romantisch
- cool

12. Wie sehr stimmen Sie folgenden Aussagen zu?	stimme voll zu					stimme überhaupt nicht zu
Die Zeit hier im Skigebiet "DEMO-Skigebiet" hat eine große persönliche Bedeutung für mich.	①	②	③	④	⑤	⑥
Ich fühle mich mit dem Skigebiet "DEMO-Skigebiet" emotional verbunden.	①	②	③	④	⑤	⑥
Wenn ich nicht in das Skigebiet "DEMO-Skigebiet" kommen könnte, würde mir das sehr viel ausmachen.	①	②	③	④	⑤	⑥
Das Skigebiet "DEMO-Skigebiet" ist mein Lieblingsskigebiet.	①	②	③	④	⑤	⑥

2 www.[redacted].at



Wie zufrieden waren Sie bei uns? Bitte beurteilen Sie nachfolgende Punkte anhand der vorliegenden Skala (1 = äußerst begeisternd, 6 = eher enttäuschend, k.A. = keine Angabe).

13. Anreise zum Skigebiet	☺			☹	k.A.
[illegible]	① ② ③	④	⑤ ⑥		⓪
[illegible]	① ② ③	④	⑤ ⑥		⓪
[illegible]	① ② ③	④	⑤ ⑥		⓪
[illegible]	① ② ③	④	⑤ ⑥		⓪
[illegible]	① ② ③	④	⑤ ⑥		⓪
[illegible]	① ② ③	④	⑤ ⑥		⓪
[illegible]	① ② ③	④	⑤ ⑥		⓪

14. Kassenbereich	☺			☹	k.A.
[illegible]	① ② ③	④	⑤ ⑥		⓪
[illegible]	① ② ③	④	⑤ ⑥		⓪
[illegible]	① ② ③	④	⑤ ⑥		⓪
[illegible]	① ② ③	④	⑤ ⑥		⓪

15. Skipassangebot	☺			☹	k.A.
[illegible]	① ② ③	④	⑤ ⑥		⓪
[illegible]	① ② ③	④	⑤ ⑥		⓪
[illegible]	① ② ③	④	⑤ ⑥		⓪

16. Seilbahnen & Lifte	☺			☹	k.A.
[illegible]	① ② ③	④	⑤ ⑥		⓪
[illegible]	① ② ③	④	⑤ ⑥		⓪
[illegible]	① ② ③	④	⑤ ⑥		⓪
[illegible]	① ② ③	④	⑤ ⑥		⓪
[illegible]	① ② ③	④	⑤ ⑥		⓪

17. Skigebiet	☺			☹	k.A.
[illegible]	① ② ③	④	⑤ ⑥		⓪
[illegible]	① ② ③	④	⑤ ⑥		⓪
[illegible]	① ② ③	④	⑤ ⑥		⓪
[illegible]	① ② ③	④	⑤ ⑥		⓪
[illegible]	① ② ③	④	⑤ ⑥		⓪
[illegible]	① ② ③	④	⑤ ⑥		⓪

3 [illegible]

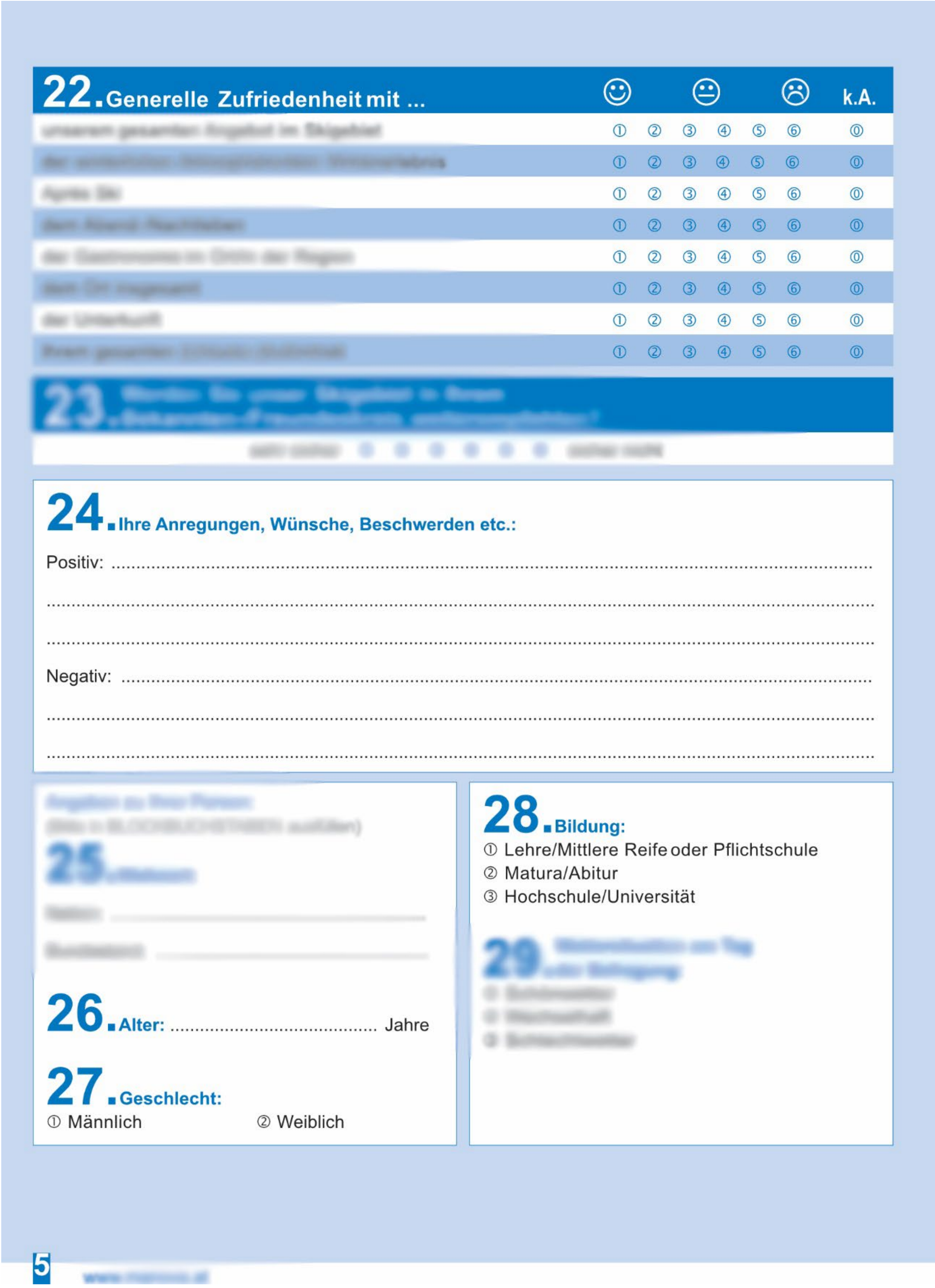
18. Pisten	☺			😐		☹	k.A.
[illegible]	①	②	③	④	⑤	⑥	⓪
[illegible]	①	②	③	④	⑤	⑥	⓪
[illegible]	①	②	③	④	⑤	⑥	⓪
[illegible]	①	②	③	④	⑤	⑥	⓪
[illegible]	①	②	③	④	⑤	⑥	⓪
[illegible]	①	②	③	④	⑤	⑥	⓪
[illegible]	①	②	③	④	⑤	⑥	⓪

19. Infrastruktur - Services	☺			😐		☹	k.A.
[illegible]	①	②	③	④	⑤	⑥	⓪
[illegible]	①	②	③	④	⑤	⑥	⓪
[illegible]	①	②	③	④	⑤	⑥	⓪

20. Erlebnisangebote am Berg	☺			😐		☹	k.A.
[illegible]	①	②	③	④	⑤	⑥	⓪
[illegible]	①	②	③	④	⑤	⑥	⓪
Angebote für Kinder/Kinderland	①	②	③	④	⑤	⑥	⓪
[illegible]	①	②	③	④	⑤	⑥	⓪
[illegible]	①	②	③	④	⑤	⑥	⓪
[illegible]	①	②	③	④	⑤	⑥	⓪
[illegible]	①	②	③	④	⑤	⑥	⓪
[illegible]	①	②	③	④	⑤	⑥	⓪
[illegible]	①	②	③	④	⑤	⑥	⓪
[illegible] am Berg (Unterhaltung,Aussichtsl.)	①	②	③	④	⑤	⑥	⓪

21. Gastronomie im Skigebiet	☺			😐		☹	k.A.
[illegible]	①	②	③	④	⑤	⑥	⓪
[illegible]	①	②	③	④	⑤	⑥	⓪
Angebot an Speisen und Getränken	①	②	③	④	⑤	⑥	⓪
[illegible]	①	②	③	④	⑤	⑥	⓪
Atmosphäre	①	②	③	④	⑤	⑥	⓪
[illegible]	①	②	③	④	⑤	⑥	⓪
[illegible]	①	②	③	④	⑤	⑥	⓪
[illegible]	①	②	③	④	⑤	⑥	⓪
Mitarbeiter	①	②	③	④	⑤	⑥	⓪
[illegible]	①	②	③	④	⑤	⑥	⓪

4



22. Generelle Zufriedenheit mit ...

	☺		😐		☹		k.A.
[illegible]	①	②	③	④	⑤	⑥	⓪
[illegible]	①	②	③	④	⑤	⑥	⓪
[illegible]	①	②	③	④	⑤	⑥	⓪
[illegible]	①	②	③	④	⑤	⑥	⓪
[illegible]	①	②	③	④	⑤	⑥	⓪
[illegible]	①	②	③	④	⑤	⑥	⓪
[illegible]	①	②	③	④	⑤	⑥	⓪
[illegible]	①	②	③	④	⑤	⑥	⓪

24. Ihre Anregungen, Wünsche, Beschwerden etc.:

Positiv: ....

....

....

Negativ: ....

....

....

26. Alter: .... Jahre

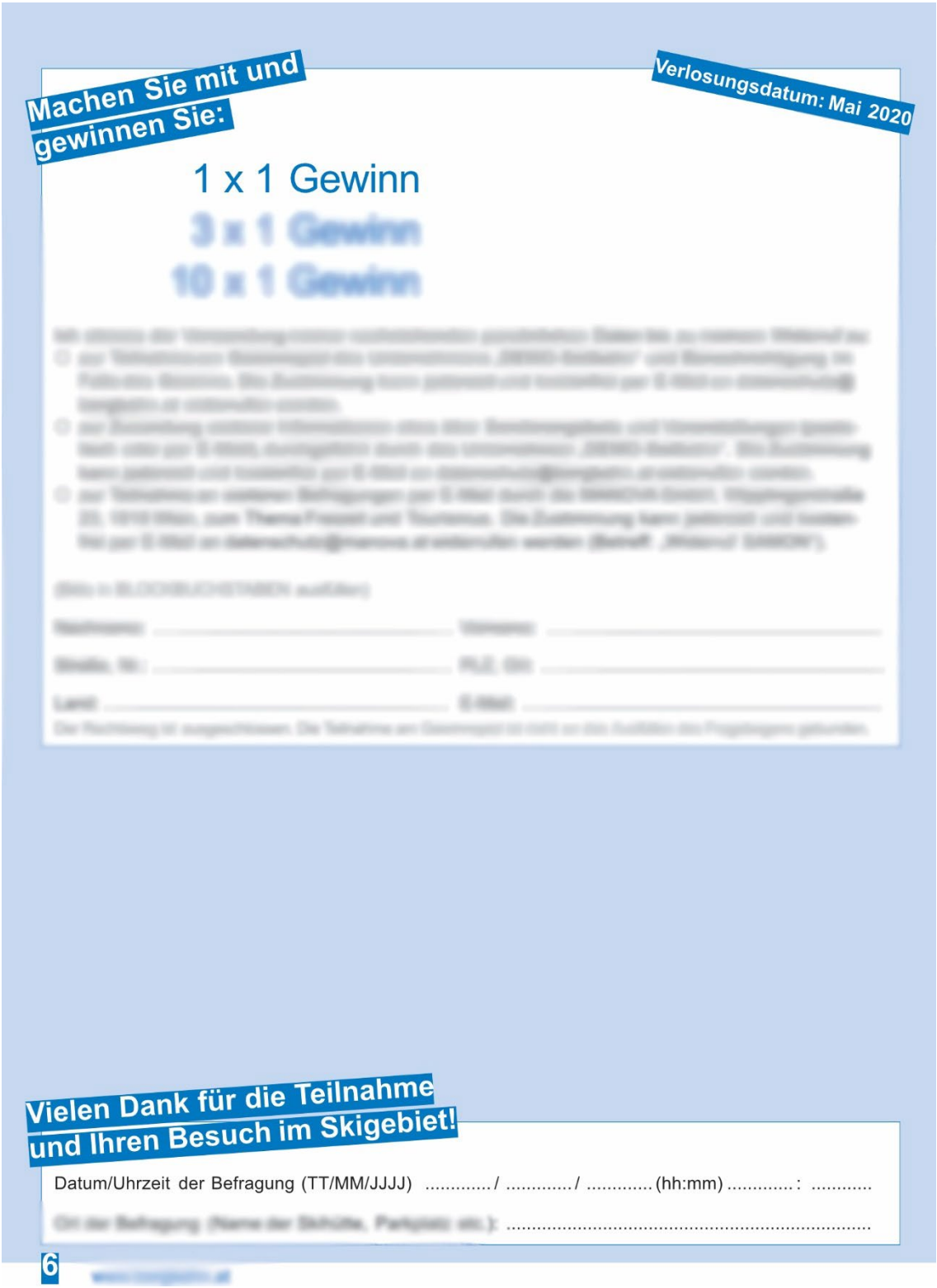
27. Geschlecht:

① Männlich ② Weiblich

28. Bildung:

① Lehre/Mittlere Reife oder Pflichtschule
② Matura/Abitur
③ Hochschule/Universität

5



Appendix 2: The Code

This appendix chapter provides the full RStudio code used in this research.

The comments section, due to the RStudio limitations, starts with a # symbol.

```
#Clean environment
rm(list = ls())
#Clean console: Ctrl+L
#Set path
setwd("C:/Rstd")
library(haven)
#get data:
dataset1 <- read_sav("mt.sav")

#Descriptive Statistics General#

###Frequency Gender###
dataset1$sex<-replace(dataset1$sex, dataset1$sex == 1, "Male")
dataset1$sex<-replace(dataset1$sex, dataset1$sex == 2, "Female")
dataset1$sex<-replace(dataset1$sex, dataset1$sex == 3, "Other")
dataset1$sex<-replace(dataset1$sex, dataset1$sex == 0, NA)

freq<-table(dataset1$sex)
print ("Frequency count of Gender")
print (freq)

library(dplyr)
library(ggplot2)

###Bar CHart

ggplot(data = dataset1, aes(x = factor(sex),
                            y = prop.table(stat(count)),
                            label = scales::percent(prop.table(stat(count))))) +
 geom_bar(position = "dodge", fill = "#e3fd9d") +
 geom_text(stat = 'count',
```

```
        position = position_dodge(.9),
        vjust = -0.5,
        size = 5) +
 scale_y_continuous(labels = scales::percent) +
 labs(x = 'Gender', y = '') + theme_classic() +
 theme(axis.text=element_text(size=20),text=element_text(size=20))

###Frequency Age###

library(AMR)

dataset1$age_groups<-age_groups(dataset1$alter, split_at = c(14, 26, 46, 66), na.rm = FALSE)

freq2<-table(dataset1$age_groups)
print ("Frequency count of Age Groups")
print (freq2)

###Bar Chart

ggplot(data = dataset1, aes(x = factor(age_groups),
                    y = prop.table(stat(count)),
                    label = scales::percent(prop.table(stat(count))))) +
 geom_bar(position = "dodge", fill = "#e3fd9d") +
 geom_text(stat = 'count',
        position = position_dodge(.9),
        vjust = -0.5,
        size = 5) +
 scale_y_continuous(labels = scales::percent) +
 labs(x = 'Age', y = '') + theme_classic() +
 theme(axis.text=element_text(size=20),text=element_text(size=20))

###Frequency Education###
```

```
dataset1$bildung<-replace(dataset1$bildung, dataset1$bildung == 1, "Lower secondary education")
dataset1$bildung<-replace(dataset1$bildung, dataset1$bildung == 2, "Upper secondary education")
dataset1$bildung<-replace(dataset1$bildung, dataset1$bildung == 3, "Tertiary education")

dataset1$bildung<-replace(dataset1$bildung, dataset1$bildung == 4, NA)

freq3<-table(dataset1$bildung)
print ("Frequency count of Gender")
print (freq3)

library(dplyr)
library(ggplot2)

###Bar Chart

ggplot(data = dataset1, aes(x = factor(bildung),
                    y = prop.table(stat(count)),
                    label = scales::percent(prop.table(stat(count))))) +
 geom_bar(position = "dodge", fill = "#e3fd9d") +
 geom_text(stat = 'count',
         position = position_dodge(.9),
         vjust = -0.5,
         size = 5) +
 scale_y_continuous(labels = scales::percent) +
 labs(x = 'Education', y = '') + theme_classic() +
 theme(axis.text=element_text(size=20),text=element_text(size=20))


#Descriptive statistics Satisfaction Variables later

###Cleaning the data###
```

```
#clean the columns with 100% NA:

data_noNAcolumns <- dataset1[,colSums(is.na(dataset1))<nrow(dataset1)]


#create data frame with only variables interesting for the research:

df_variables <- data_noNAcolumns[,47:103]


#additional cleaning of the rows and columns (columns and rows with more than 50%NA deleted)

df_variables_clean2 <- df_variables[which(rowMeans(!is.na(df_variables)) > 0.5), which(colMeans(!is.na(df_variables)) > 0.5)]

# save csv to create long command in excel

#write.csv(df_variables_clean, "df_variables_clean.csv")


###Get rif of 0s in the 1-6 range

df_variables_clean<-replace(df_variables_clean2, df_variables_clean2 == 0, NA)


#Rename variables to English versions of the Labels (initial labels and variables in German)

colnames(df_variables_clean) <- c("General_Cable_Cars","Lift_Access","Lift_Waiting_Times","Lift_Comfort","Lift_Employees","General_Ski_Resort","Slopes_Orientation","Area_Slopes_Variety","Ski_Resort_Difficulty","Ski_Resort_Size","Ski_Resort_Panorama","General_Slopes","Slopes_Readiness","Slopes_Length","Slopes_Snow","Slopes_Safety","Slopes_Width","Services_Ski_Schools","Services_Rental_Service","Services_Ski_Depot","General_Experiences","General_Hiking","General_Children_Offers","General_Freeride","General_Rodelbahn","General_Snowpark","General_Funslope","General_Additional_Offers","General_Attractions","General_Gastronomy","Gastronomy_Variety","Gastronomy_Offer_Food_and_Drinks","Gastronomy_Quality","Gastronomy_Atmosphere","Gastronomy_Place_Availability","Gastronomy_Waiting_Time","Gastronomy_Cleanliness","Gastronomy_Value_for_Money","Gastronomy_Employees","General_Offers","General_Atmosphere_Winter_Experience","Apres_Ski","Gastronomy_Resort","General_Location","General_Accommodation","General_Overall","General_Evening_Nightlife","Slopes_Space")


#Remove German Labels

library(labelled)

df_variables_clean <-remove_labels(df_variables_clean)


#library(tidyverse)

library(arsenal)
```

```
#Creating descriptive statistics data for all satisfaction variables (Mean,(SD))
my_controls<-tableby.control(
 test = T,
 total = T,
 numeric.test = "kwt", cat.test = "chisq",
 numeric.stats = c("meansd", "range"),
 cat.stats = c("countpct", "Nmiss2"),
 stats.labels = list(
  meansd = "Mean (SD)",
  medianq1q3 = "Median (Q1, Q3)",
  range = "Min - Max",
  Nmiss2 = "Missing"
 )
)

table_one <- tableby(~.,
                data = df_variables_clean,
                control = my_controls
)

tab1<-summary(table_one, title = "Satisfaction Variables")

###'*The HTML is saved to the C:/Rstd/ please check your destination*
write2html(tab1, "C:/Rstd/Descriptive_Statistics.html")
###'*Look inside the HTML*

#Model fit with cleaned data

summary(lm(df_variables_clean$General_Overall~0+df_variables_clean$General_Cable_Cars+df_variables_clean$Lift_Access+df_variables_clean$Lift_Waiting_Times+df_variables_clean$Lift_Com-
```

```
fort+df_variables_clean$Lift_Employees+df_variables_clean$General_Ski_Resort+df_varia-
bles_clean$Slopes_Orientation+df_variables_clean$Area_Slopes_Variety+df_variables_clean$Ski_Re-
sort_Difficulty+df_variables_clean$Ski_Resort_Size+df_variables_clean$Ski_Resort_Panorama+df_varia-
bles_clean$General_Slopes+df_variables_clean$Slopes_Readiness+df_varia-
bles_clean$Slopes_Length+df_variables_clean$Slopes_Snow+df_variables_clean$Slopes_Safety+df_var-
iables_clean$Slopes_Width+df_variables_clean$Services_Ski_Schools+df_variables_clean$Ser-
vices_Rental_Service+df_variables_clean$Services_Ski_Depot+df_variables_clean$General_Experi-
ences+df_variables_clean$General_Hiking+df_variables_clean$General_Children_Offers+df_varia-
bles_clean$General_Freeride+df_variables_clean$General_Rodelbahn+df_variables_clean$Gen-
eral_Snowpark+df_variables_clean$General_Funslope+df_variables_clean$General_Additional_Of-
fers+df_variables_clean$General_Attractions+df_variables_clean$General_Gastronomy+df_varia-
bles_clean$Gastronomy_Variety+df_variables_clean$Gastronomy_Offer_Food_and_Drinks+df_varia-
bles_clean$Gastronomy_Quality+df_variables_clean$Gastronomy_Atmosphere+df_varia-
bles_clean$Gastronomy_Place_Availability+df_variables_clean$Gastronomy_Waiting_Time+df_varia-
bles_clean$Gastronomy_Cleanliness+df_variables_clean$Gastronomy_Value_for_Money+df_varia-
bles_clean$Gastronomy_Employees+df_variables_clean$General_Offers+df_variables_clean$Gen-
eral_Atmosphere_Winter_Experience+df_variables_clean$Apres_Ski+df_variables_clean$Gastron-
omy_Resort+df_variables_clean$General_Location+df_variables_clean$General_Accommoda-
tion+df_variables_clean$General_Evening_Nightlife+df_variables_clean$Slopes_Space))


library(sjPlot)

#library(sjmisc)

#library(sjlabelled)


#m2 <-
lm(zuf_ges~0+zuf_lift+lif_zuga+lif_wart+lif_komf+lif_mita+zuf_gebi+pis_mark+geb_viel+geb_grad+geb_
groe+geb_pan+zuf_pis+pis_prae+pis_laenge+pis_snow+pis_sich+pis_breit+ser_schule+ser_serv+ser_de-
pot+zuf_alternativ_generell+zuf_alternativ_winterwandern+zuf_alternativ_kinder+zuf_alterna-
tiv_freeride+zuf_alternativ_rodeln+zuf_alternativ_snowpark+zuf_alternativ_funslope2+zuf_alterna-
tiv_pistenzusatz+zuf_alternativ_attraktio-
nen+zuf_gast+gas_viel+gas_ang+gas_qua+gas_atmo+gas_sitz+gas_warte+gas_saub+gas_pl+gas_mita+z
uf_sb+zuf_erl+zuf_apr+zuf_go+zuf_ort+zuf_unterk+pis_platz+zuf_nachtleben, data = df_varia-
bles_clean)

m2<-lm(General_Overall~0+., data = df_variables_clean)

#create HTML output in viewer

tab_model(m2)

#plot click: Linearity/Normality/Homoscedasticity(homogeneity of variance)/No Outliers

plot(m2) + theme_classic() + theme(axis.text=element_text(size=50),text=element_text(size=50))
```

```
#'*create dataframe for cluster analysis:*

#p<0.1

df_cluster <- df_variables_clean[, c("Lift_Access","General_Ski_Resort","Ski_Resort_Panorama","General_Slopes","Slopes_Safety","Services_Ski_Schools","General_Hiking","General_Children_Offers","General_Rodelbahn","Gastronomy_Atmosphere","Gastronomy_Waiting_Time","Gastronomy_Value_for_Money","General_Atmosphere_Winter_Experience","General_Offers","Apres_Ski","General_Location","General_Accommodation","Slopes_Space"

)]


#####Scale visuals #Will be divided in two/three parts to fit the A4 page#

#install.packages("sjPlot")

library(sjPlot)

#divide the data

df_cluster_part1<-df_cluster[,c(1:9)]

df_cluster_part2<-df_cluster[,c(10:18)]

# Plot the scales

sjplot(df_cluster_part1, fun = c( "likert"), grid.range=c(.15,1), grid.breaks = 0.1, geom.colors = c("#6f9703","#94c904","#e3fd9d","#e0d0f2","#c1a2e4","#ad84dc"), values = "show", show.n = FALSE, show.legend = TRUE, show.prc.sign = FALSE, expand.grid = TRUE) +

labs(x = '', y = 'Part 1') + theme_classic() +

  theme(axis.text=element_text(size=20),text=element_text(size=20))

sjplot(df_cluster_part2, fun = c( "likert"), grid.range=c(.15,1), grid.breaks = 0.1, geom.colors = c("#6f9703","#94c904","#e3fd9d","#e0d0f2","#c1a2e4","#ad84dc"), values = "show", show.n = FALSE, show.legend = TRUE, show.prc.sign = FALSE, expand.grid = TRUE) +

labs(x = '', y = 'Part2') + theme_classic() +

  theme(axis.text=element_text(size=20),text=element_text(size=20))

#Data transpose

data<-as.data.frame(t(df_cluster))
```

```
####Final move to Cluster Analysis:


hcluster<-hclust(d=dist(data,method="euclidean"),method = "ward.D2")
hcluster


library(cluster)
library(dendextend)
library(factoextra)


grp <- cutree(hcluster, k = 3)
head(grp, n = 5)
table(grp)


rownames(data)[grp == 1]


#Creating cluster dendrogram



#nothing outlined
fviz_dend(hcluster, # Cut in three groups
        cex = 1.5, # label size
        k_colors = c("#4a6402","#ad84dc","#94c904"),
        repel = FALSE, lwd = 1.5,
        type = c("rectangle"),
        horiz = TRUE,
        color_labels_by_k = TRUE, # color labels by groups
        rect = TRUE ,# Add rectangle around groups
        labels_track_height = 270
)  + theme_classic() + theme(axis.text=element_text(size=20),text=element_text(size=20))
#Clusters outlined
fviz_dend(hcluster, k = 3, # Cut in three groups
        cex = 1.5, # label size
```

```
        k_colors = c("#4a6402","#ad84dc","#94c904"),
        repel = FALSE, lwd = 1.5,
        type = c("rectangle"),
        horiz = TRUE,
        color_labels_by_k = TRUE, # color labels by groups
        rect = TRUE ,# Add rectangle around groups
        labels_track_height = 270
)  + theme_classic() + theme(axis.text=element_text(size=20),text=element_text(size=20))



#Creating a Cluster Plot
#looking for zero variance columns
which(apply(data, 2, var)==0)
#getting rid of zero variance columns in data2 to be able to produce cluster plot
data3<-data[ , which(apply(data, 2, var) != 0)]


fviz_cluster(list(data = data3, cluster = grp),
          palette = c("#4a6402","#94c904","#ad84dc"),
          ellipse.type = "convex", # Concentration ellipse
          repel = TRUE, # Avoid label overplotting (slow)
          show.clust.cent = TRUE) + theme_classic() + theme(axis.text=element_text(size=20),text=element_text(size=20))


#Philogenic Plot with layout as a tree
fviz_dend(hcluster, k = 3, # Cut in three groups
        cex = 1.5, # label size
        k_colors = c("#4a6402","#ad84dc","#94c904"),
        repel = TRUE, lwd = 1.5,
        type = c("phylogenic"),
        color_labels_by_k = TRUE, # color labels by groups
        rect = TRUE ,# Add rectangle around groups
        phylo_layout = "layout_as_tree"
```

```
) + theme(axis.text=element_text(size=0),text=element_text(size=20))



dat1 <- df_cluster[which(rowMeans(!is.na(df_cluster)) > 0.5), ]
hcluster1<-hclust(d=dist(dat1,method="euclidean"),method = "ward.D2")
library(cluster)
library(dendextend)
library(factoextra)

grp <- cutree(hcluster, k = 3)
head(grp, n = 5)
table(grp)

rownames(dat1)[grp == 1]

#plot(hcluster) ! careful us with enough RAM may lead to fatal error

#fviz_dend(hcluster1, # Cut in three groups
#         cex = 1.5, # label size
#         k_colors = c("#4a6402","#ad84dc","#94c904"),
#         repel = FALSE, lwd = 1.5,
#         type = c("rectangle"),
#         horiz = TRUE,
#         color_labels_by_k = TRUE, # color labels by groups
#         rect = TRUE ,# Add rectangle around groups
#         labels_track_height = 270
#)



hclustergroups<-cutree(hcluster1, k=3)
hclustergroups
```

```
table(hclustergroups)



aggregate(dat1,by=list(hclustergroups),FUN=mean, na.rm=T)

table22<-aggregate(dat1,by=list(hclustergroups),FUN=mean, na.rm=T)

write2html(table22, "C:/Rstd/aggregate.html")
```