ROUGH SET KANSEI ENGINEERING: MULTIPLE USERS, MULTIPLE KANSEIS

A Dissertation by

Ali Ahmady

Master's Degree, Isfahan University of Technology, 1998

Bachelor's Degree, Iran University of Science and Technology, 1994

Submitted to the Department of Industrial and Manufacturing Engineering
and the faculty of Graduate School of
Wichita State University
in partial fulfillment of
the requirements of the degree of
Doctor of Philosophy

December 2010

ROUGH SET KANSEI ENGINEERING: MULTIPLE USERS, MULTIPLE KANSEIS

The following faculty members have examined the final copy of this dissertation for form and content, and recommend that it be accepted in partial fulfillment of the requirement for the degree of Doctor of Philosophy with a major in Industrial Engineering.

Don Malzahn, Committee Chair

Seyed Hossein Cheraghi, Committee Member

Abu Masud, Committee Member

Barbara S. Chaparro, Committee Member

Hamid M. Lankarani, Committee Member

Accepted for the College of Engineering

Zulma Toro-Ramos, Dean

Accepted for the Graduate School

J. David McDonald, Dean

To my beloved parents, my wife Sahar, and my children

In completing one discovery we never fail to get an imperfect knowledge of others of which we could have no idea before, so that we cannot solve one doubt without creating several new ones

Joseph Priestley

ACKNOWLEDGMENTS

First and foremost, I offer my sincerest gratitude to my supervisor, Dr. Don Malzahn, who supported me throughout my dissertation with his patience and knowledge while allowing me the room to work in my own way. I definitely benefitted from his scientific intuition, which made him a constant source of ideas and passion throughout this research.

I express my appreciation to Dr. Seyed Hossein Cheraghi for his supervision, advice, and guidance in helping me initiate this research as well as giving me support and encouragement. I am thankful to both Dr. Malzahn and Dr. Cheraghi for the extraordinary experiences they gave me throughout this work by opening my mind to new ideas and ways to approach the topic, develop the method, and conduct the research, all of which inspired and enriched my growth as a student and researcher.

Special thanks to Dr. Barbara Chaparro, faculty member in the Department of Psychology and a member of my supervisory committee, for her critical contribution in providing the data necessary to implement this research and her extremely helpful comments. She played a major role in the success of my research.

Deepest gratitude is also due to other members of the supervisory committee, Dr. Abu Masud and Dr. Hamid M. Lankarani.

Some of my inspiration for this research began at the very early stages in Dr. Janet Twomey’s class. I really appreciate her encouragement and kindness to answer my questions throughout this research.

I wish to express my great love and gratitude to my parents who planted the seeds for the enjoyment of learning in my early years. Thank you for your support, guidance, encouragement, prayers, and love.

Words fail me to express appreciation to my wife whose dedication, love, and persistent confidence in me has taken the load off my shoulders. I am indebted to her for her unselfish devotion and commitment to helping me achieve my goals.

Also, I am grateful to my brothers and sisters and my parents-in-law, for their understanding, prayers, support, and endless love throughout the duration of my studies.

It is a pleasure to thank a few professors from other universities who helped during the process of this study:

- Dr. Wojciech Ziarko, Professor in the Department of Computer Science at University of Regina in Regina, Canada, who provided access to his class notes and corrected my path to understanding and applying rough set theory to develop the approach.
- Dr. Aijun An, Associate Professor in the Department of Computer Science and Engineering at York University in Toronto, Ontario, Canada, who assisted me in finding the appropriate software for this study.
- Professor Roman Słowiński, Professor and Founding Head of the Laboratory of Intelligent Decision Support Systems within the Institute of Computing Science, Poznań University of Technology, Poland, for his very valuable guidance and answers to my questions regarding rough set theory.

And last, but not least, I would like to thank everyone who had a contribution in conducting this study.

ABSTRACT

The method proposed in this dissertation addresses the need to relate product features to customer expectations. This is particularly difficult given the variety of consumer perspectives and the uncertainty in their assessments. Current statistical methods may not relate all of the market research information available to customer-oriented product-development approaches.

Rough set-based Kansei Engineering (RSBKE) is an approach for reasoning under uncertainty and deals with imperfect information originating from the imprecision of human assessment. This mathematically powerful approach extracts knowledge from customer survey data and develops product design rules based upon single or multiple subjective impressions (Kansei) from single or multiple users.

A two-stage user-oriented product development approach generates market segmentation rules and product design rules for either a single or multiple Kansei(s). RSBKE provides an enhanced means of defining primary customer groupings and automatically generating design rules. Several extensions to target marketing, lead-user identification, and Kano model applications are presented. RSBKE can be extended to the decision attributes of functional customer requirements. The approach presented here is compared to statistical methods.

A case study involving a website design was used to illustrate this approach. The results identified distinctive classes of users who had the same perception of a set of websites. The system generated a set of strict design rules for each class.

TABLE OF CONTENTS

TABLE OF CONTENTS (continued)

TABLE OF CONTENTS (continued)

LIST OF TABLES

Table Page

LIST OF TABLES (continued)

LIST OF TABLES (continued)

Table Page

LIST OF TABLES (continued)

Table Page

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

To remain competitive, companies and organizations continually strive to enhance the quality of their products and services. Because of the diversity of users, companies can either expand their designs and produce a product for all users (generalization) or produce a specific product for each individual user (individualization). Either approach would be costly and/or wasteful for users and companies. Balancing the trade-off between these two methods is a major responsibility of designers. The question is how to accomplish this balance.

How can a designer fashion a product that meets users' needs and maximizes producers' profit? Techniques such as market segmentation and customer analysis (Juran, 1992) have been developed to help companies find their position between generalization and individualization (customization). The main aim of these techniques is to define the most effective set of customer classes so that companies can concentrate on their requirements to better serve them. But even in the most optimistic situation, customers in the same group will not have the same requirements even if they are categorized precisely. Also, not all requirements may have the same level of importance. With these considerations in mind, how can designers design products to best incorporate customers' desires? A purely customized design is costly and wasteful while too much generalization may not satisfy some of the customers' needs. Therefore, what should companies do? The answer to this question is the subject of this study.

1.1 Customer-Oriented Product Development

Adam Smith introduced the concept of "sovereignty of the consumer" two hundred years ago (Brown, 1993). Today, there is no doubt about this concept. This simple notion has led many organizations to develop their products according to customers' needs. They compete with each other to obtain and retain more customers. This competition will never end. Since companies are using more advanced technologies, they can offer more efficient products. This raises customers' expectations, and at the same time, customers are more mature, cognizant, and sophisticated. To respond quickly to change, many product development techniques have been developed. These methods consider the customer as part of the product development process. Today, product development is a strategic process for many organizations (Margaret & Biemans, 1995).

Product design is at the core of the product development process and is intended to satisfy customer needs. Many quality scholars such as Crosby (1928–2001), Deming (1900–1993), and Juran (1904–2008) considered customer requirements as the key to quality improvement of any product. Today, companies should not only consider functional customer needs but also map customer feelings and emotions into product design to fully satisfy them. Giving extensive attention to the end-user's feelings and emotions is called Human Centered Design or emotional design and is advocated by Norman (2004). In Japan, Nagamachi (1995) developed a systematic approach to translate customer feelings into product design parameters and called it Kansei Engineering (KE). KE has been successfully applied in many industries. This study mainly focuses on this approach.

Quality, which any organization is eager to achieve, is nothing but "the degree to which a set of inherent characteristics fulfills requirements" (International Organization for Standardization, 2005). But who are the customers?

Juran (1992) defined the customer as "anyone who is impacted by the product or process." Based on his definition, a customer could be internal or external. Internal customers are impacted by the product and are members of the company. The parties who are impacted by the product but are not members of the company are external customers. Therefore, clients (either purchasers or end-users) who buy the product, government regulatory bodies, labor unions, non-governmental organizations and other advocacy groups, local and national communities, and the public, all of whom might be influenced by the consequences of unsafe products or damage to the environment, can be considered as external customers. In this research, external customers are the main concern.

1.2 Product Value

One of the main strategic goals of a company is to increase the value of its consumer products. If a company insists on competing on price alone, then it probably has chosen the weakest position upon which to compete, since a motivated customer using the Internet can find the cheapest product very quickly.

The value of a product is the difference between what a customer receives—including product features, quality, and service—and what a customer forfeits—including the amount paid for the product, plus the time and effort spent acquiring the product and learning how to use it. To increase value to the customer, a company should maximize what a customer receives from a product and minimize what a customer gives in order to obtain it. For this reason, a company should not compete based solely on price.

1.3 Customer-Attribute Hierarchy

It is obvious that customers define product value. They determine value based on what they perceive about products, usually using different layers of product concepts to assess product value and hence their preference (Margaret & Biemans, 1995). Based on the means-end chain model (Gutman, 1982), three different layers can be recognized:

- Perception of a product (product attributes).
- Perceived relationship between a product and the person (consequences of using the product).

- The person (human values).

1.3.1 Product Attributes

Each product has certain characteristics. When these characteristics are perceived by consumers, they are usually considered to be attributes. Two kinds of attributes through which a product is perceived by consumers are concrete and abstract. Concrete attributes are tangible product characteristics that are physical and perceived objectively, and they are usually perceived by all customers in the same way, such as the color of a car. On the other hand, abstract attributes are subjective and intangible characteristics of a product. The quality of a product could be considered an abstract attribute and, as mentioned before, is part of the product value. Different customers might perceive these attributes differently.

1.3.2 Relationship Between Product and Person

A product not only has characteristics that are perceived by customers but also has two kinds of consequences: functional and psychosocial. Functional consequences are the direct practical outcomes of the product for the customers and are more or less objective. For example, a call-making capability could be considered a functional consequence of a cell phone. In addition, a product may have a psychosocial consequence, that is, the psychological and social effects of using or owing a product. Psychosocial consequences are, to some extent, emotional in nature. They are more subjective than functional consequences. Also, since users have different emotions and belong to different social classes, psychosocial consequences can be different for different users.

1.3.3 Human Values

Each person has a personal set of values, which are fundamental convictions believed to be good (Rokeach, 1973). Usually these kinds of values are assumed to guide human behavior, but in this study, it is assumed that they have less direct impact on customer perceptions.

One of the main points in the means-end chain model is that the above three layers of product meaning are interconnected. Concrete attributes can affect abstract attributes. For instance, the weight of a car (concrete attribute) may affect the customer's perception of the car's reliability (abstract attribute). In addition, recognition of concrete attributes and perception of abstract product attributes can influence customer perception of functional and psychosocial product consequences.

1.4 Heterogeneous Population

A product can have many users who are not necessarily homogenous. These different users constitute a heterogeneous population or market. If one wants to classify heterogeneous customers in terms of their similarities and differences, each customer should be considered a class, since individuals are in fact different from one another (Kano, 1996). "In this age of diversification and individualization it is safe to say that each individual is a customer." But the question is, is it possible or economically feasible to customize one specific product for a user?

1.5 Generalization Versus Customization

Kano (1996) states that the ideal is to design one specific product for each user. However, this is usually very costly and sometimes practically impossible. In fact, a company could generalize a product, meaning that it designs and produces one specific product for all users, regardless of whether or not the product's market is heterogeneous or homogenous. Generalization could be efficient, but there is no guarantee of satisfying all customers. For example, consider a cell phone that has many functions in order to cover a wide range of customer requirements. Obviously, some of those functions will not be used by certain customers; therefore, these functions are wasteful for both the company and those customers.

On the other hand, customization can be costly for both sides—companies and customers. In other words, if the company wants to design and produce a particular product for a specific user, it is very costly for the company, not to mention the technological limitations. In some cases, it would be difficult for certain customers to afford such an expensive product. Therefore, subject to organizational interest and the difficulty of designing a product that can please everyone, designers should try to find an optimal point between individualization (customization) and generalization. To do so, companies may design for all customers (extreme), customize for each customer (extreme), or customize for each group of customers. Usually companies consider cost and their technological abilities in deciding the most beneficial solution. Each point between these two extreme points represents the level of generalization or individualization.

1.6 Uncertainty and Inconsistency in Heterogeneous Markets

As mentioned earlier, the market for a product is usually heterogeneous rather than homogenous. Heterogeneity can cause some uncertainty and inconsistency.

In general, inconsistency is the quality of not being in agreement and lacking a harmonious uniformity among things or parts (http://dictionary.die.net/inconsistency). In fact, inconsistency is a kind of ignorance that could be recognized because of inaccuracy, conflict, contradiction, or confusion. "Inconsistency can result from assignments and substitutions that are wrong, conflicting, or biased, producing confusion, conflict, or inaccuracy, respectively" (Ayyub & Klir, 2006).

Heterogeneity in the market could result in inconsistencies. Since different users may have different perceptions with respect to the same product, lack of harmony could occur in their thoughts and/or feelings. Therefore, designers will assign different users to different distinguishable customer groups using market segmentation techniques to decrease the level of inconsistency among customer perceptions. Confusion for designers and conflict in user perceptions can occur from inconsistent assignments and substitutions, whereas inaccuracy results from a level bias or error in these assignments and substitutions.

Besides inconsistency, heterogeneity in a market can cause uncertainty in different ways for designers who want to design a product that is matched to user requirements and feelings. The sources of uncertainty can be ambiguity, approximation, and likelihood. All these sources of uncertainty underlie the current research problem. Since the input data of the problem are taken by sample, there is an inherent likelihood of uncertainty in the problem. On the other hand, the nature of the problem can have ambiguous uncertainty for designers, since it is quite possible to have multiple outcomes for the product evaluation process in a heterogeneous market. Perhaps the most important source of uncertainty is approximation, which can cause uncertainty for designers. Approximation includes vagueness, coarseness, and simplification. “Vagueness results from the imprecise nature of belonging and non-belonging of elements to a set or a notion of interest, whereas coarseness results from approximating a set by subsets of an underlying partition of the set’s universe that would bound the crisp set of interest” (Ayyub & Klir, 2006). When users consider some products as “reliable” and some as “unreliable,” this is an uncertain situation for designers who want to determine the most influential product features based on user-reliability perceptions. Therefore, designers have to perform approximation through reduction and generalization.

According to Zimmerman’s uncertainty definition, “uncertainty applies to the particular situation that a person does not dispose about information which quantitatively and qualitatively is appropriate to describe, prescribe or predict deterministically and numerically a system, its behavior or other characteristics” (Zimmermann, 2000). He categorizes the causes of uncertainty in different classes. Either one, two, or all of these sources of uncertainty could be considered as causes of uncertainty in a heterogeneous market and are explained below.

1.6.1 Lack of Information

Lack of information is the most frequent cause of uncertainty in a heterogeneous market. Since a product could have many users, designers will not or do not want to consider all user preferences to design a perfect product due to time and cost limitations. Therefore, by tolerating a reasonable error, transition from a situation of uncertainty caused by lack of information to a situation of greater certainty could be achieved using approximation methods, which are discussed in probability theory and statistics as well as survey sampling methods (Scheaffer, 1996) and (Soler, 2005).

1.6.2 Abundance of Information

Sometimes the amount of data is too large to be perceived and processed at the same time by people (here, designers) because of the limitation of human beings’ ability (Zimmermann, 2000). In a heterogeneous market, a large number of subjective attributes of a product could be perceived by multiple users. Resolving this uncertainty would be a challenge for designers who want to use all of this data. In this kind of situation, people usually “transform the available data into perceivable information by using a coarser grid or a rougher granularity or by focusing their attention on those features which seem to them most important and neglecting all other information or data” (Zimmermann, 2000). This cause of uncertainty is the type of uncertainty that is considered in this study.

1.6.3 Conflicting Evidence

In a heterogeneous market, if some users perceive a product as high quality and others consider it as low quality, this will definitely be an uncertain situation for designers. As Zimmerman states, "if two classes of available information are conflicting, then an increase of information might not reduce uncertainty at all, but rather increase the conflict" (Zimmermann, 2000). This kind of uncertainty could be resolved by checking the correctness of data, deleting some pieces of information that could be outliers, or putting the information on a rougher grid. This cause of uncertainty exists in a heterogeneous market and is discussed in Chapter 3.

1.6.4 Ambiguity

In all languages, some words have different meanings in different contexts. Therefore, a particular word can mean different things for multiple individuals, if they are not in the same condition and/or they do not know the context of the word. This kind of uncertainty can happen in customer surveys in a heterogeneous market. In this study, this cause of uncertainty will not be discussed.

1.6.5 Belief

One of the main causes of uncertainly is people's differences in their beliefs, opinions, and convictions. As a cause of uncertainty, all information available to people is subjective (Zimmermann, 2000).

Existence of this cause of uncertainty is obvious in a heterogeneous market, which contains different users with different minds. Abstract attributes perceived by different users is a good example of this kind of cause for uncertainty, which is included in this study.

1.7 Kansei Engineering and Multiple Users

Although Kansei Engineering is one of the product development techniques described in detail in the ‘background” section in Chapter 2, some points regarding this technique and multiple users are mentioned here. KE is a concept whereby the main aim is to determine product characteristics that satisfy customer attributes (specifically, customer feelings) using a variety of different qualitative and quantitative methods. Customer feelings could come from one user or a population of users who are not necessarily the same (homogenous). Therefore, in this context, inconsistency among multiple users’ different points of view will arise. This inconsistency needs to be resolved by an appropriate method, which is the subject of this study.

1.8 Rough Set Theory and Multiple Users

Rough set theory provides a framework of approximation and reasoning about data (Kudo & Murai, 2007). Using this method, various decision rules can be generated from Kansei data. In this study, the possibility of using rough set theory to determine the most important product characteristics and generate “if-then” design rules from customer evaluation data are investigated. This method is detailed in Chapter 2.

1.9 Conclusion

In summary, companies are looking for ways to increase product value to their customers. Customers define this value, but it is the company’s (including designer’s) goal to maximize it.

When designers create a product, they should translate subjective attributes, including abstract attributes, and psychosocial and functional requirements, into objective characteristics. During this process, some uncertainty and inconsistency will arise and must be resolved.

Markets are not usually homogenous; there are multiple users for a product. On the other hand, when different users perceive product characteristics (which are objective) and their corresponding consequences, these become subjective attributes. These subjective attributes again should be converted to objective characteristics so they can be used by designers. Many methods have been developed for this conversion. Among them, Kansei Engineering maps user feelings to product characteristics. The core of KE is to find the most influential product features based on customer perceptions. Usually this is accomplished by discovering the relationships between customer feelings and product features. Many techniques have been used to find this relationship, such as statistical multiple regression, rough set theory, and artificial neural network (ANN). Among these, rough set theory is the one that has been used most recently. Rough set is a powerful mathematical approach to data mining, cognitive science, and decision analysis. Also, in this study, rough set is preferred over multiple regression since it can handle non-linear and non-normal data, which are very common with human evaluation data and also generate decision rules. Furthermore, due to the data size reduction property of rough set theory, which is based upon set theory, this method is preferred over neural networks for the purpose of this study.

In this dissertation, rough set theory is utilized to translate efficiently the subjective perceptions (Kansei) of multiple users in a heterogeneous market into product characteristics. In Chapter 2, background and state-of-the-art literature on this subject are reviewed. In Chapter 3, the problem is defined in detail. The proposed approach is discussed in Chapter 3. In Chapter 4, the proposed approach is implemented in a real-world context of website design. The results of implementation are demonstrated and discussed in Chapter 5. Chapter 6 includes the conclusions and future work.

CHAPTER 2

BACKGROUND AND LITERATURE REVIEW

This chapter includes two subsections—background and literature review. Since Kansei Engineering, rough set theory, and traditional product development approaches are used in this study, these methods are briefly introduced in the background section; then state-of-the-art literature on the research subject is reviewed in the literature review section.

2.1 Background

2.1.1 Kansei Engineering

Kansei Engineering is a customer-oriented technique, which originated in Hiroshima University, Japan, around 1970. This technique, along with other total quality management (TQM) tools, has been used successfully in many Japanese firms to improve the quality of their products. KE is one of the techniques that attempt to find the relationship between customer's feelings and a product's features. Other parallel techniques such as quality function deployment (QFD) use functional customer requirements as dependent variables (Akao, 2004).

There is no accurate synonym in English for the Japanese word "Kansei," which is comprised of two words—Kan and Sei—both of which can be interpreted as sensitivity, sensibility, feeling, or aesthetics (Schütte, 2005). However, Nagamachi (1995) defined Kansei as "customer's feeling and includes the customer's feeling about the product design, size, color, mechanical function, feasibility of operation, and price as well." In addition, Kansei Engineering can be defined as an efficient method for rendering customer feelings into product design elements (Matsubara & Nagamamachi, 1997).

2.1.1.1 How Kansei Engineering Works

Since the feelings and preferences of customers usually are vague and unclear, a method is needed to analyze, interpret, and translate them into technical designer language. Kansei Engineering assists designers in making decisions and concentrating on associated product design elements that match human feelings. The following three main points of this method are addressed (Nagamachi M. , 2002):

- Recognizing the user’s Kanseis.
- Clarifying the relationship between the customer’s Kansei and the product features.
- Constructing a computerized-assisted database that will help designers create and improve a product.

Although there are six different types of Kansei Engineering, which are explained later in more detail later, their structure is almost the same. A framework for KE methodology is depicted in Figure 2.1 (Schutte, Eklund, Axelsson, & Nagamachi, 2004).

The first step of this methodology is to define the Kansei domain, which is the selection of a customer group and a market segment, and the specification of the product. The product can be an existing product or concept, or an unknown design solution. The second step is to span the semantic space by collecting and measuring Kansei words. Kansei words can be obtained from magazines, relevant literature, manuals, experts, experienced users, and related Kansei studies.

Some methods have been developed for measuring the gathered Kansei words. Physiological response (e.g., heart rate, EMG, EEG), people’s behavior, facial and body expressions, and spoken words are just some of these methods that are sorted based on the complexity of behavioral patterns (Schütte et al., 2004).

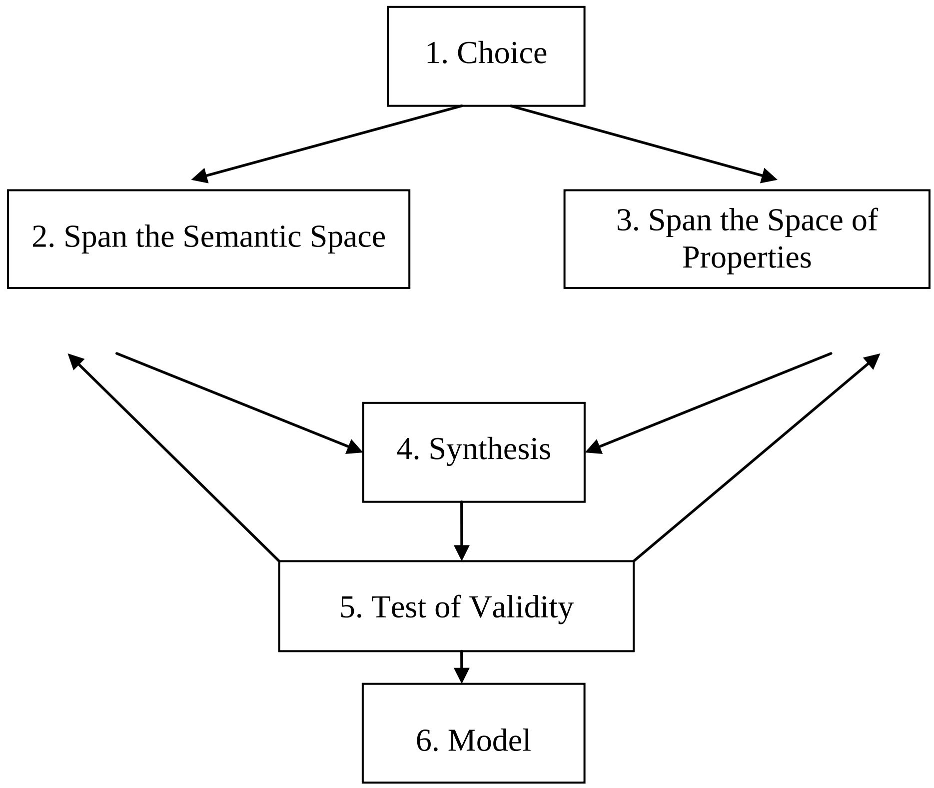


Figure 2.1. Model of Kansei Engineering concept.

Sometimes in practice, the number of Kansei words is decreased to a more practical number. This can be done by a pilot study using semantic differentials, factor analysis, focus groups, and/or expert groups and affinity diagrams.

After the Kansei words are gathered, measured, and refined, the third step is to collect the product features in which the space of properties is spanned. These features should be chosen based on the largest impact on Kansei words. The following sources can be helpful in providing product characteristics (Schütte et al., 2004): technical documents, comparisons of competing products, magazines, relevant literature, manuals, experts, experienced users, and related Kansei studies.

Since the aim of Kansei Engineering is to make customers feel better about products, in the next step, those product attributes that have the most relationship with Kansei words are defined. This step is called “synthesis.” Multiple regression analysis (or Hayashi’s quantification theory type I and Hayashi’s quantification theory type II, both of which are a kind of multi-variate analysis for qualitative data), are just two of the mathematical methods that can be used to find the relationship between Kansei words and design category items. Other tools that have been applied to Kansei Engineering are the general linear model, neural network, genetic algorithm (GA), or rough set analysis (Schütte et al., 2004).

Sometimes, when the relationships between Kansei words and product features are determined, some of the product characteristics do not address Kansei words. This can occur as the result of an inaccurate reduction of Kansei words in the first step or because of outdated product attributes. This kind of reviewing of results is referred to by Schütte et al. (2004) as a test of validity. After the validity test is completed, a model can be made based on the data gathered from the synthesis. In this kind of model, the independent variables are the product properties, and the model can project the Kansei score for certain words. Based on the tool applied, the function can be qualitative, linear, or non-linear.

Kansei Engineering, like other product development systems based on customer demands, has specific steps. The first step is to collect the customer demands or Kansei words. In this step, some market research tools such as focus groups or depth interviews can be used (Hill & Alexander, 2000). Second, the product’s technical specifications that are selected to be developed must be characterized. In the case of a complex product that has many components, several Kansei words must be defined for each component. The third step is to find and analyze the relationship between Kansei words and design elements. Here, some mathematical tools like multiple regression analysis can be applied. In some types of KE, the qualitative evaluation by a cross-functional team is used. In more advanced types of KE, newer and more complicated evaluation methods, like a genetic algorithm or artificial neural network, are used.

Once the product characteristics are distinguished, designers use their creativity to set the value of these characteristics to gratify the customer's Kansei. Sometimes the current product with its existing features does not have the potential to cover all Kansei words. In this case, designers create new features so that all Kansei words have counterpart characteristics in the product.

2.1.1.2 Types of Kansei Engineering

Kansei Engineering can be categorized into five types (Nagamachi M. , 1999):

- Type One: Category Classification
- Type Two: Kansei Engineering Computer System
- Type Three: Kansei Engineering Modeling
- Type Four: Hybrid Kansei Engineering
- Type Five: Virtual Kansei Engineering

A sixth category can also be added—Collaborative Kansei Engineering (Schütte et al., 2004).

2.1.1.2.1 Type One: Category Classification

Mazda (a Japanese automotive maker) applied type one, category classification, of Kansei Engineering to develop a new type of sports car called the "Miata" (Japanese name, Unos Roadster). It was simply a conceptual tool that did not use mathematical methods. Category classification is a method by which Kansei words are decomposed into a tree structure to define the design element details. In this method, the basic concepts of customers' desires, called the zero-level category, are obtained via market survey. These concepts are in the customers' language and do not address design characteristics. In KE category classification, these basic concepts must be broken down into subconcepts, as the design features of a product are obtained. In Miata's case (Nagamachi M. , 1999), "human-machine unity" was one of the basic concepts obtained via market research. This concept emphasized that most sports car drivers are young and want the feeling of oneness between themselves and the car during their driving. The developing team members subdivided this zero-level category into four levels: "tight feeling," "direct feeling," "speedy feeling," and "communications." For example "tight feeling" can be translated into greater detail such as "fitting closely to the machine" and "neither large nor small." To satisfy these feelings, they found that some technical characteristics, like chassis length, were related to them.

2.1.1.2.2 Type Two: Kansei Engineering Computer System

A Kansei Engineering computer system is an expert system that helps to transfer Kansei words and images into physical design elements. This system usually has five databases and an inference engine, as shown in Table 2.1 (Nagamachi M. , 1995).

TABLE 2.1 KANSEI ENGINEERING COMPUTER SYSTEM PROCESS

Database	Operations
Kansei Words	Gather Kansei words
	Analyze by multivariate analysis, such as factor analysis and cluster analysis
	Evaluate by semantic differential method
	Analyze by Hayashi's quantification theory type II
Image	Construct relationship between Kansei words and design elements
Knowledge	Construct control database, color conditioning principle, and round design guideline
Shape	Construct relationships between Kansei words and shapes
Color	Construct relationships between Kansei words and colors

Kansei words are analyzed by some kind of multivariate analysis, such as factor analysis or cluster analysis, and are stored in a Kansei words database. Then the data that are evaluated by the semantic differential (SD) method are analyzed using the Hayashi quantification theory type II (a kind of multiple regression analysis for qualitative data). In an image database, the statistical relationship between Kansei words and design elements are saved so that a customer or designer can link contributory items in the design parameters to particular Kansei words, and vice versa. The knowledge base includes some rules (if-then rules) to control the database and design guidelines. In addition, the shape design database includes the relationship between shape aspects of design and Kansei words, and the color database consists of the relationship between colors and Kansei words. The entire design, including shape and color, can be extracted by the particular inference system based on if-then rules and can be shown in graphics on a screen. Table 2.1 summarizes the Kansei Engineering computer system process. As can be seen, the whole process includes three major functions: gathering and analyzing Kansei words, identifying the most influential design elements on Kansei words, and using these elements and incorporating them with suitable shapes and colors to obtain the design best-matched to customer feelings.

2.1.1.2.3 Type Three: Kansei Engineering Modeling

Usually there is an expert system, including if-then rules, inside the Kansei Engineering system. In KE modeling, the expert system is substituted with mathematical models, such as an artificial neural network, genetic algorithm, rough set theory, and so forth. In other words, KE modeling establishes relationships between Kansei words and product specifications using mathematical models so that designers can determine how and how much each specification influences Kansei words, how to satisfy each Kansei word, and how much each design element has to be changed.

2.1.1.2.4 Type Four: Hybrid Kansei Engineering

Hybrid Kansei Engineering includes two directions: forward and backward. In forward Kansei Engineering, a designer can input Kansei words that have been gathered from customers into the system, and then the system provides the most relative design elements with Kansei words as output. On the other hand, sometimes designers want to see how well the initial design matches customer feelings before going further. In this case, inputting the initial sketch into the system will evaluate it based on Kansei words and provide a score that shows to what extent the initial design satisfies customer feelings. The KE computer system that includes these two directions is called hybrid Kansei Engineering (Nagamachi M. , 2002).

2.1.1.2.5 Type Five: Virtual Kansei Engineering

In this type of Kansei Engineering, a virtual product is created by the designers and shown to the customers virtually. Customers can walk through the virtual Kansei environment using a head mounted display (HMD) and data gloves, and evaluate the virtual product with respect to their Kansei. Using this customer feedback, designers can improve their designs and invite customers to share their feelings regarding the redesigned products.

2.1.1.2.6 Type Six: Collaborative Kansei Engineering Designing

In collaborative Kansei Engineering designing, all databases, including the Kansei words database, are accessible via the Internet. Utilizing intelligent software, this system can bring the viewpoints of designers and customers together rapidly.

2.1.1.3 Kansei Engineering and Its Applications

2.1.1.3.1 Type One: Category Classification

Due to the simplicity and similarity to other traditional product development approaches, Kansei Engineering type one is the most popular application. Since Nagamachi helped Mazda develop the "Miata" in 1986 (Nagamachi M. , 1997), this type of KE has been applied to many products. Schütte et al. (2005) applied KE type one to the design of rocker switches for work vehicles. Three major Swedish vehicle manufacturers—BT Industries AB, Saab Automobile AB, and Scania AB—in cooperation with Linkoping University, started a project, which had four main goals. First, they identified the list of Kansei words for rocker switches in order to cover the complete product domain with as few words as possible. Second, they identified how rocker switches were perceived by their users. Third, they determined which switch's properties generated these perceptions. Finally, they suggested some recommendations to improve the design of these switches. The method they used consisted of six steps:

- Step 1: Choose the product domain. In this step, user groups, market segments, and type and class of products were defined. The domain chosen was "rocker switches for use in work-utility vehicles and similar environments." This definition identified the products, the market segments, and potential target groups.
- Step 2: Span the semantic space. This step had three substeps: (a) collecting Kansei words from different sources, such as the Internet, literature, newspapers, and manuals; (b) reducing the number of words without losing any crucial information by using some method, such as an affinity diagram; and (c) selecting the representative of each group as the final Kansei words. After taking these steps, 118 Kansei words were gathered, of which 32 were identical and therefore eliminated. Ten Kansei words were added at the suggestion of the industries. Using an affinity diagram, they categorized 94 Kansei words into 14 groups, from which 29 words were drawn as representative Kansei words.
- Step 3: Span the space of properties. Like the previous step, this step has three substeps: (a) collecting product properties; (b) identifying important properties, which was done by company experts; and (c) making a final selection of properties. After spanning the space of rocker switch properties, 12 different properties were identified and put into three main groups: mechanical design, electrical design, and form design. These properties were identified by designers, expert users, and advanced users.

- Step 4: Synthesize. In this step, the interaction between the scale space (Kansei words) and concept space (product samples) were analyzed. Participants in this study rated each product with respect to each Kansei word using a semantic differential scale with five main points and two extremes points.
- Step 5: Test the validity and update. In this step, some tests were done to ensure that the data were valid. In this study, the Kolmogorow-Smirnov test was carried out to see if the data followed a normal distribution. It was found that when participants were confused about a particular Kansei word, they tended to give middle values of the scale, so a one-sample t-test was conducted to see whether the mean value was different from the middle value of the scale. In cases where the data were not normally distributed or the middle value for one Kansei word was equal to four, that Kansei word was subjected to further examination or excluded from the list of Kansei words. Here, the data consisted of the ratings given to each product for each Kansei word by 71 persons who participated in the study. In this project, all Kansei words showed normal distributions. In addition, there was no indication of deviation from the expected value, so no Kansei word was excluded from the list.
- Step 6: Build a model. In this step, a mathematical model was built to facilitate understanding, present results, and obtain a future prediction. Here, a three-dimensional matrix can be considered, where concepts (product sample), scales (Kansei words), and subjects represent these three dimensions. Since most statistical approaches need two-dimensional data, taking the mean over subjects reduces one of the dimensions. However, it may lose some information. Hence, the data were analyzed by using factor analysis. Doing this analysis provided information about how the Kansei words correlated with

each other and which Kansei words were of highest importance. Then the most important Kansei words were connected to product properties using quantification theory type I, which is a kind of regression method.

In this study, using factor analysis, the vehicle manufacturers defined the most important factors and, using them, calculated the weighted values, which represent the satisfaction level for each concept. Out of 14 switch concepts, five received higher Kansei values than the rest of the switches. Using quantification theory type I, they calculated the correlation between product properties and important factors (Kansei words). At this step, they were able to say which product properties had the most influence on the most important Kansei words. Based on these findings, they developed three different switch concepts.

2.1.1.3.2 Type Two: Kansei Engineering Computer System

The Kansei Engineering computer system is an expert system that helps to transfer Kansei words into design elements. Shimizu, Sadoyama, Kamijo, Hasoya, et al. (2004) created a KE computer system with three main steps: collecting Kansei words, synthesizing data, and compiling data (Kansei words and product features) to obtain products matched with customer requirements, feelings, and validation. Their objective was to design a system that would allow customers to input their specific requirements. With this system, the intermediates were eliminated. To show how this theory can be put into practice, the authors applied this system to the textile industry as interactive production system apparel, and they developed some clothing such as shirts, socks, and jackets. Their approach was as follows:

1. Create a database that contains many kinds of clothing in terms of colors, forms, and materials.
2. Create a retrieval system in which customers can evaluate and select the products according to their self concept, desire, and physiological body characteristics.

3. Create an individualized clothing pattern-making system. After a customer selected his/her desired clothing, a three-dimensional digitization of the shape determined the clothing measurements on the body. Then an interactive body model provided accurate information for individual pattern design. Following that, an interactive 3D-CAD created a human body model with control points, which could be modified interactively. This information was used to make clothing patterns and to simulate clothing pattern fitting for individual body shapes.
4. Validate the designed products. Customers were allowed to evaluate the comfort and feel of the clothes. For example, they used an electroencephalogram (EEG) to measure the waist belt pressure, which affects comfort and blood circulation.

In summary, the authors tried to demonstrate the benefits of the interactive production system. This method allows customers to enter their requirements into the system, manufacturers to produce the appropriate product based on the requirements, and manufacturers to obtain feedback in order to improve their products. Also, this system eliminates intermediate layers and excess production that results in overstock of goods. In addition, the authors suggested individuality using Kansei, which is the opposite of mass production.

2.1.1.3.3 Type Three: Kansei Engineering Modeling

In type 3, Kansei Engineering modeling, the “if-then rule” is inside the Kansei system. This kind of expert system can be an artificial neural network, genetic algorithm, rough set theory, or others. Kosaka, Nishitani, and Watanabe (2005) used a neural network to estimate the reaction force of a keyboard switch based on Kansei information. This research used a neural network for selecting and designing a switch so that the operators felt most comfortable. The aim of this study was to develop a product design system where the input is Kansei data and the output is parameters for designing keyboard switches. These parameters are four kinds of reaction forces: Fl: initial reaction force, F2: peak reaction force, F3: drop reaction force, and F4: ending reaction force. The authors found seven Kansei words representing the feelings of switch users. Changes in reaction forces affected these Kansei words and user feelings. Some surveys were done to generate data relative to the relationship between reaction forces and Kansei. It was discovered that some relationships between reaction forces and Kansei words were nearly linear and some were nonlinear. Therefore, they used an artificial neural network as an appropriate tool to estimate these relationships. This ANN has the capability to define the reaction forces based on particular Kansei words. Knowing the set of reaction forces, designers can create a product matching a particular customer's feelings.

In addition, Shimohara and Shimazaki (2005) used rough set theory and Kansei Engineering to improve the design of a walking space. They used the rough set model to find the most important combination of elements of walking space that affect pedestrians' comfort preferences. The factors affecting people's preferences are considered a combination of various elements. For example, people prefer a walking space with roadside trees but not if the walk is narrow. In the case of walking space without trees, a wider walking space is preferred. To find these combinations of elements, first a sample of people was asked to look at 28 photographs of different walking spaces and express their opinions in terms of them being pleasant or unpleasant. Each photograph included some elements of a walking space. Using the reduct method on a rough set model, the attributes that did not affect preference were defined. In this case, roadside trees, poles, and width were contained in all combinations, so they were essential. Also, using rough set theory, the authors defined the decision rules that show combinations of composition elements that affect preferences. For instance, some decision rules considering walking space as pleasant are as follows: "pavement type is interlocking and roadside trees are on both sides or randomly located," or "no poles, textured paving block color is yellow."

In summary, using the rough set theory, the most important combination of elements of walking space that affected pedestrians' preferences could be found by simply asking the people which walking space was pleasant or unpleasant for them; the rough set then defined their decision rules and essential attributes.

2.1.1.3.4 Type Four: Hybrid Kansei Engineering

Hybrid Kansei Engineering has two directions: backward and forward. In backward hybrid Kansei Engineering, designers input their designs to the system and obtain the evaluations of their designs in terms of customer Kansei words.

Bouchafra and Tan (2003) tried to introduce a new methodology that maps designs to human perceptions using a Markov model and Kansei Engineering. In other words, their paper answers the question of how designers can evaluate their designs in terms of customer perceptions before the object is produced. First, the authors created a database that computed semantic relationships among customers' perceptions. Then, they divided the set of perceptions into k clusters that could be used to classify them further. Finally, they developed a new classifier called the structural hidden Markov model (SHMM), which can anticipate and learn user perceptions given an object's design. They applied this approach to Kansei Engineering in order to map external shapes of cars to customers' senses. They collected 114 images of regular cars with three views and presented these images to 100 young female students. These students were asked to express their feelings about the images, and the responses were used to create sets of perceptions. The authors put a new design into the system and obtained the Kansei evaluation. Since the optimal prediction of user perceptions is fed to the design engineer before the object is produced, from an economical standpoint, their model can save industrial companies a considerable amount of money.

2.1.1.3.5 Type Five: Virtual Kansei Engineering

Virtual reality can be used to obtain the feelings of customers regarding a specific product, even if the product is simply a design and not a real item. A virtual reality system creates a virtual environment so that the customer can use and feel the product virtually, and designers use this feedback to improve the product.

Oyama (1997) used virtual Kansei Engineering to enhance human communication between a doctor and a patient, and between a nurse and a patient. Using virtual reality, they analyzed and measured cancer patients' Kansei in a virtual reality. Then, based on feedback, the quality of communication was improved.

2.1.1.3.6 Type Six: Collaborative Kansei Engineering Designing

Collaborative Kansei Engineering designing or the Internet Kansei Designing System (IKDS) is an Internet-supported KE system that helps to collect the viewpoints of customers and designers via the Internet. Using this method, the early phases of product development can be reduced and shortened.

2.1.1.4 Another Kansei Engineering Classification

Kansei Engineering, like other customer-oriented product development methods, clarifies the relationship between the customer's feelings (Kansei) and the product's features. Each of the six aforementioned types of KE derives this relationship using its own approach (Anitawati, Nor Laila, & Nagamachi (2007); Jindo & Hirasago, (1995); Nagamachi, (1999); Tanoue, Ishizaka, & Nagamachi, (1997)). Information availability, complexity, and required performance are those factors upon which the current classification of KE methods is based. However, these methods could be classified according to other criteria, such as the method that is used to derive the relationship. Generally, there are two methods that can be used to find relationships: mathematical (quantitative or computational) and nonmathematical (qualitative or noncomputational). Kansei Engineering types two, three, and four could be considered computational KE, and the others can be classified noncomputational KE.

Figure 2.2 illustrates this classification, and the two categories of KE (computational and noncomputational) are explained below (Ahmady A. , 2008).

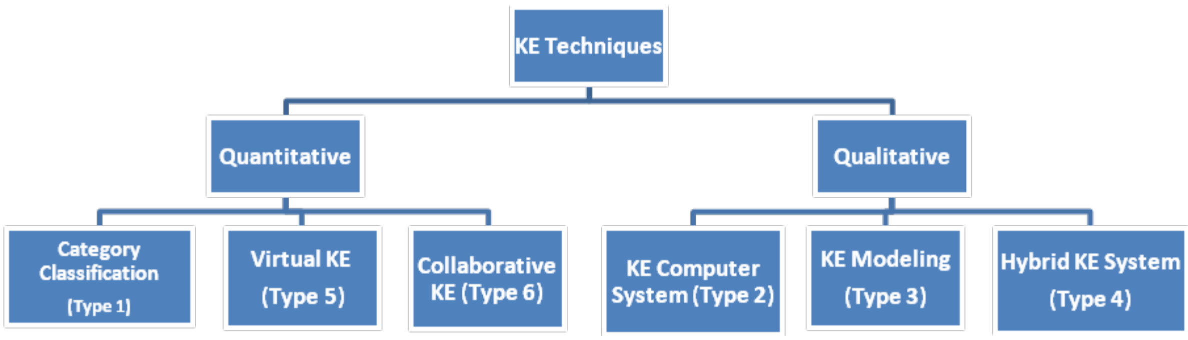


Figure 2.2. Kansei Engineering classification techniques.

2.1.1.4.1 Computational Kansei Engineering

In this category of Kansei Engineering methods, feelings (Kansei) words are considered dependent variables, and product features are independent variables. The purpose here is to find the most contributory product characteristics to stimulate or satisfy customer's feelings. In the KE computer system (type two) and KE modeling (type three), product features are identified from Kansei words (forward method); however, in the hybrid KE (type four), designers have additional features so that their designs can be evaluated with respect to Kansei words (backward method). The most common methods used to find relationships between Kansei and product feature sets are multiple linear regression analysis, neural network, genetic algorithm, and rough set analysis.

2.1.1.4.2 Noncomputational Kansei Engineering

Category classification (type one), virtual KE (type five), and collaborative KE (type six) are grouped in this class. In KE types one and five, no specific mathematical or computational methods are used to define the relationships between variable sets. For example, in category classification, the developing team used qualitative approaches such as the Delphi method, while in type five, virtual reality technology is used. Although in the collaborative method, designers used the Kansei database and the intelligent system, which is the KE computer system (type two), the difference between this method and the KE computer system is that in the collaborative method, computer and network capabilities are used extensively as a communication base for designers. So this method could be considered a noncomputational KE method.

2.1.1.5 Kansei Engineering: Advantages and Disadvantages

Kansei Engineering is just one of many methods that have been developed to consider customer needs and requirements. KE has some advantages and disadvantages compared to other product-development techniques.

Relative to advantages, first, Kansei Engineering can be supported by many other techniques to reach its goal, which is to design a product that addresses customer emotions. While other product development techniques, such as quality function deployment or customer-oriented product concept (COPC), use only a few qualitative techniques to find the relationship between product design characteristics and customer demands, Kansei Engineering has the capability of using both qualitative and quantitative methods to find this relationship (Shutte & Eklund, 2005). Artificial neural network, rough set theory, genetic algorithm, fuzzy logic, factor analysis, multiple regression analysis, multivariate analysis, and so forth are just some of the techniques that have been used in KE so far and have been mentioned in this study.

Second, over the years, different types of Kansei Engineering have been generated, depending on concepts and tools that have been applied in this method (Takagi, Watada, & Yubazaki, 2004), (Nagasawa, 2004). For example, once KE used virtual reality to obtain feedback from customers and the many benefits were discovered, then others began to use virtual reality in KE. Thus, different types of Kansei Engineering became popular. In addition, as mentioned previously, the newest type of KE is type six, collaborative Kansei Engineering designing, which is only a concept now, and its application has not yet been published. It is anticipated that other types of KE will be developed in the future.

Third, as previously mentioned, Kansei Engineering is one method that considers customer feelings in product development. In fact, this method enhances the quality of a product in terms of customer feelings and can include functional requirements, while most other techniques can only improve quality in terms of a customer’s functional requirements.

Finally, since Kansei Engineering uses certain mathematical models and computer abilities, its process is more repeatable. Thus, the results are more reproducible (since the quantitative data can be manipulated in a certain way) than traditional product development methods, which do not have the capability of using a quantitative model.

Kansei Engineering, like every other method, has some disadvantages. Although Oyama (1997) tried to analyze a cancer patient’s Kansei using virtual reality technology, thus far, few KE applications in service industries have been seen. Nevertheless, in the future, KE might be applied in service industries. In addition, using KE in most companies is costly and time consuming, and requires experts in the field. Perhaps that is why most of the applications of this method have been done in large organizations.

2.1.2 Rough Set Theory

2.1.2.1 Introduction

The rough set model was advocated by Pawlak in 1982 (Pawlak, 1991). This model was used to determine the causal relationship between illness and patient symptoms from medical data. This section provides a brief introduction to rough set theory.

Vagueness and uncertainty are the nature of many real-life situations. Medicine (i.e., the notion of a healthy [or ill] person), economy, politics, engineering, and many other fields are examples in which situations may not be clear. For example, in economics and politics, vague concepts are essentially the basis of thinking and debate.

To be clearer, the concept of odd and even numbers is precise or crisp because a number is either even or odd. However, the concept of beauty is vague because one cannot, without bias, conclude whether a particular thing is beautiful or not. “Vague concepts form the basis for common sense reasoning in many fields connected with real-life situations” (Pawlak, 1998).

The concept of set, introduced by George Cantor in 1883, is a basic notion of modern mathematical thinking and is needed to provide exactness. However, this concept leads to antinomies and is not strong enough to handle vagueness and uncertainty (Peters & Skowron, 2004). To overcome this defect, many improvements in set theory have been proposed. Among them, the one proposed by Lesniewski (Peters & Skowron, 2004), which replaces the “membership relation between elements and set in classical set theory” with “being a part,” has attracted more attention than others. Rough set theory and fuzzy set theory, which were introduced by Pawlak (1991) and Zadeh and Kacprzyk (1992), respectively, are two mathematical approaches to vagueness. While vagueness is defined by the degree of membership in fuzzy set theory, the boundary region of a set is a vagueness representation in rough set theory.

The basic idea of fuzzy set theory is based on the fuzzy membership function that defines the degree of membership of an element to a set (which is a sort of vagueness). In rough set theory, vagueness is due to lack of information about some elements. For example, two patients who are middle-aged women are suspected of having the same disease, but the results of testing for a specific disease shows that one of them has that disease (positive), whereas the other one does not have it (negative). In this case, before the test is done, based on available information, these two patients (two elements) are indiscernible. Here, one cannot say that a person in the same class of age, status of disease, and gender belongs to the set of positive test results or negative test results. This kind of indiscernibility leads to a concept which, in philosophical literature, is known as a boundary line. A boundary line or boundary region is the set of elements that cannot be linked to the specific concept (or set) for certain (in the above case, the concepts are positive and negative results sets). However, the sets that include people with negative (or positive) results (including uncertain cases, such as patients with disease symptoms and negative results, and vice-versa) are called rough sets, since the elements or objects are assigned to sets roughly, not exactly, based on available information.

Today, rough set theory, which is based on imperfect information, has been used successfully in many areas, such as artificial intelligence and cognitive sciences, including machine learning, knowledge acquisition, decision analysis, expert systems, and pattern recognition, as well as decision support systems and data mining (Inuiguchi, Hirano, & Tsumoto, 2003). In addition, it is applied to solve many real-life problems in medicine, pharmacology, engineering, finance, banking, and market analysis.

Generally, the main advantages of rough set theory over similar approaches, such as fuzzy set theory, neural networks, or multiple regression analysis, are as follows:

- Does not need any preliminary information about data, such as the grade of membership function, which is used in fuzzy set theory (Suraj, 2004).
- Can reduce the knowledge that is required to find the minimal set of conditions so that the objects are discernible (Pawlak, 1991).
- Offers efficient methods, algorithms, and tools for finding a hidden pattern in data (Pawlak & Slowinski, 1994).
- Allows evaluating the significance of data (Pawlak & Slowinski, 1994).
- Can generate automatically the set of decision rules (policies) from data (Pawlak & Slowinski, (1994); An, Shan, Chan, Cercone, & Ziarko, (1996)).
- Is robust, since decision rules in rough set theory are obtained without additional assumptions (like probability in statistics or grade of membership in fuzzy set theory) (Suraj, 2004).
- Is able to propose reduced sets of criteria so that by using them decision makers will have the same ability to approximate the decision as the whole set.

Particularly in Kansei Engineering applications, when other techniques are used to establish the relationship between Kansei words and product features, there are some shortcomings compared to rough set theory. For example, a genetic algorithm, in which product features are coded as chromosomes (Hideyoshi & Fukuda, 2005), can calculate only a few optimum solutions (Shimohara & Shimazaki, 2005). Moreover, to use multiple regression analysis, quantification theory, or even ANN, a large amount of data is required to determine the relationship between the set of attributes and the set of customer feelings (Hideyoshi & Fukuda, 2005). Handling this amount of input is costly, time consuming for the design team, and also difficult for customers to answer. Furthermore, calculation in rough set theory is simpler, and the results are more understandable compared to other nonlinear discrimination methods such as ANN (Hideyoshi & Shuichi, 2007). In addition, Nagamachi, Okazaki, & Ishikawa (2006)mentioned that since some Kansei have nonlinear features, it is not correct to apply statistical linear analysis, such as linear regression analysis or factor analysis. However, the rough set approach can handle rough and ambiguous data, regardless of linear or nonlinear data characteristics. For example, the rough set model can treat nonlinear problems such as the aesthetic image of a product. Finally, since Kanseis usually are not normal and since most common statistics techniques are based upon “normality in population distribution” assumptions (Kaplan, 2004), rough set theory can be used more appropriately than statistical linear analysis.

2.1.2.2 Basic Concept of Rough Set Theory

Rough set theory has seven basics characteristics (Pawlak & Slowinski, 1994):

1. Information system
2. Indiscernibility relation
3. Set approximation
4. Membership function
5. Dependency of attributes
6. Reduction of attributes
7. Decision rule synthesis

Each of these is explained briefly below.

2.1.2.2.1 Information System

One of the basic components of rough set theory is an information system, which can be a table. The rows of this table represent the objects (actions, alternatives, candidate, patients, etc.), and the columns represent attributes. The entries of the table are attribute values or descriptors (Inuiguchi, Hirano, & Tsumoto, 2003).

The information system is called a decision table if the set of attributes is divided into two subsets: condition attributes (criteria, tests, symptom, etc.) and decision attributes (decision, classification, taxonomies, etc.). If the decision table uniquely describes the decisions that have to be made when some conditions are satisfied, then it can be said that the decision table is deterministic (iff C → D); otherwise, it is nondeterministic.

2.1.2.2.2 Indiscernibility Relation

Having knowledge about objects allows them to be classified. Objects in the same class are assumed to have nonsignificant differences. Each class is called an indiscernible class, meaning that there is no difference between elements of that class with respect to specific criteria. These indiscernible classes are called concepts or basic building blocks. For example, some cell phones could be classified based on their weight and size; then, the indiscernible classes are the sets that contain phones which are heavy small, heavy large, light small, or light large.

Indiscernible objects in terms of condition attributes (criteria) usually prevent their exact assignment to a set, which is generated by the decision attribute. “In this case, the only sets which can be characterized precisely in terms of the classes of indiscernible objects are lower and upper approximations of the given set” (Pawlak & Slowinski, Decision Analysis Using Rough Sets, 1994). In other words, let $S = \langle U, Q, V, \rho \rangle$ be an information system; U be called a universe set; Q be a finite set of attributes; $V = \bigcup_{q \in Q} Vq$, where Vq is a domain of the attribute q; and $\rho: U \times Q \rightarrow V$ be an information function, so that $\rho(x, q) \epsilon Vq$ for every $q \in Q \subseteq U$. For example $\rho(x_1, c_1) \ = 3$ means that the information (i.e., it could be coming from a decision-maker’s evaluation) of object X_1 with respect to attribute C is 3.

Also, let $P \subseteq Q$ and x, $y \in U$, where x and y are indiscernible by the set of attributes P in S iff $\rho(x, q) = \rho(y, q)$ for every $q \in P$. For example, if $Q = \{c_1, c_2, c_3, c_4, c_5\}$ and $P = \{c_1, c_3, c_5\}. x_1$, then x_2 are indiscernible by the set of attributes P if $\rho(x_1, c_1) = \rho(x_2, c_1)$, $\rho(x_1, c_3) = \rho(x_2, c_3)$, and $\rho(x_1, c_5) = \rho(x_2, c_5)$. Thus, every $P \subseteq Q$ generates a binary relation on U, which is called an indiscernibility relation, denoted by $IND\ (P)$. A binary relation is a constraint on ordered pairs such as (x, y), where $x \in U, y \in U$. Therefore, "a formalization of a relation is done via a choice of a set of pairs which satisfy this constraint" (Polkowski, 2002). It is obvious that $IND\ (P)$ is an equivalence relation for any P. A relation that satisfies reflexivity, symmetry, and transitivity is called an equivalence relation (Polkowski, 2002). Equivalence classes of $IND\ (P)$ are called p-elementary sets in S. In summary, for any $P \subseteq Q$, there is an associated equivalence relation $IND\ (P)$, which is

$$IND\ (P) = \{(x, y) \in U^2 | \forall a \in P, a(x) = a(y)\}.$$

If $(x, y) \in IND\ (P)$, then x, y are indiscernible by attributes from P. Therefore, for any selected subset of attributes P, there is a set of objects that are indiscernible based on those attributes.

2.1.2.2.3 Set Approximation

Let a finite set of objects U and a binary relation $R \subseteq U * U$ be given. U set is called a universe, and R is called an indiscernibility relation. Also, R is meant to be an equivalence relation. An elementary portion of knowledge that we are able to perceive due to R is called the equivalence classes of the relation R or granules. Knowing the above definitions, the concepts of lower and upper approximation can be introduced. The union of all granules that are entirely included in the set is called the lower approximation of the set. Also, the union of all granules that have a nonempty intersection with the set is called the upper approximation of the set. The boundary region is the difference between the upper and lower approximations of the set. Figure 2.3 shows the graphical representation of this definition.

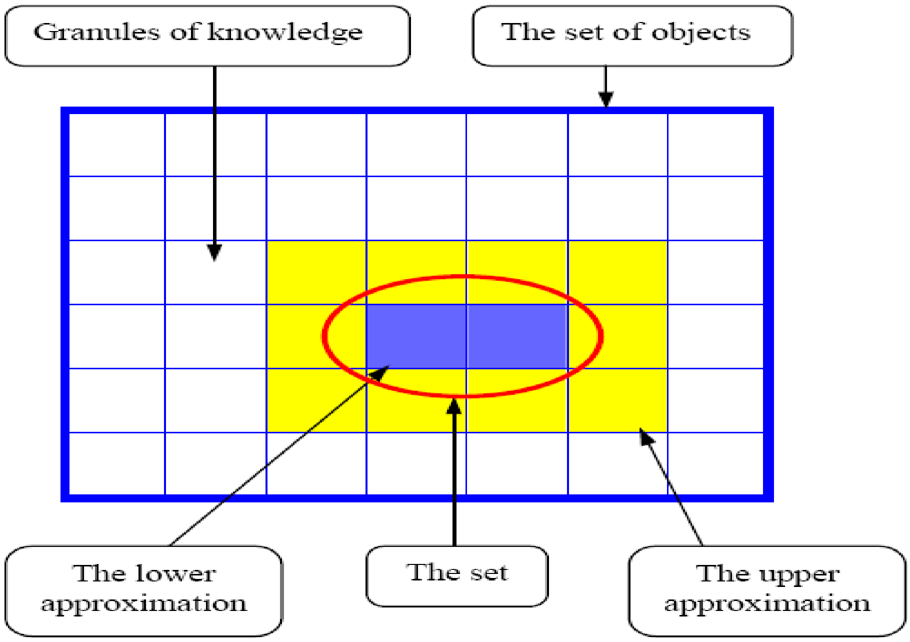


Figure 2.3. Graphical illustration of set approximation (Suraj, 2004).

In other words, let $P \subseteq Q$ and $Y \subseteq U$. Then the P(subset of condition attributes Q, or equivalence relation P, or knowledge P) lower approximation of Y, denoted by $\underline{P}Y$, and the P-upper approximation of Y, denoted by $\overline{P}Y$, are defined as

$$\underline{P}Y = \bigcup \{X \epsilon U|P : X \subseteq Y\} \tag{2.1}$$

$$\overline{P}Y = \bigcup \{X \epsilon U|P : X \cap Y \neq \emptyset\} \tag{2.2}$$

The P-boundary (doubtful region) of set Y is defined as

$$Bnp\ (Y) = \overline{P}Y - \underline{P}Y \tag{2.3}$$

With every set $Y \subseteq U$, an accuracy of approximation of set Y by P in S could be associated as

$$\alpha_p(Y) = \frac{Card(\underline{P}Y)}{Card(\overline{P}Y)} \tag{2.4}$$

Also, if $У = \{Y1, Y2, \dots, Yn\}$ is a partition of U, then

$$\gamma p(У) = \frac{\sum_{i=1}^{n} card(\underline{P}Y_i)}{card(U)} \tag{2.5}$$

is called the quality of approximation or quality of sorting of partition У by the set of attributes P. A rough set is a set defined by its lower and upper approximations.

Furthermore, $POS_P(У) = \bigcup_{Y_i \in У} \underline{P}\, Y_i$ is called the P-positive region of У and implies the set of all objects of the universe U that can be certainly classified as a member of each class of $У = \{Y_1, Y_2, \dots, Y_n\}$ using knowledge P.

2.1.2.2.4 Membership Function

In classical set theory, either an element belongs to a set or not ($[1, 0]$). The rough membership function quantifies the degree of relative overlap between the set X and the equivalence class $R(X)$ to which X belongs. This can be defined as $\mu_X^R : U \rightarrow \langle 0,1 \rangle$ where

$$\mu_X^R(x) = \frac{|X \cap R(x)|}{R(x)} \tag{2.6}$$

Figure 2.4 shows the meaning of the rough membership function.

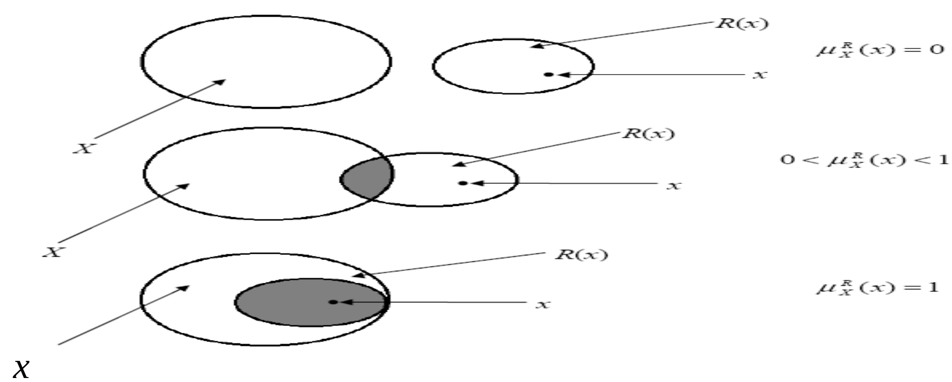


Figure 2.4. Rough membership function (Suraj, 2004).

As Figure 2.4 shows, in the first two circles, the degree of relative overlap between the set X and the equivalence class $R(X)$ is zero, since there is no intersection between them. In the second set of circles, it is not clear whether x is a member of X (or x is in X), so the degree is between zero and one. In the last set of circles, certainly x is in X, so the degree is one.

2.1.2.2.5 2.1.2.2.5 Dependency of Attributes

Suppose that U (universe) is a finite set of objects and R is a finite family of equivalence relations on U. Dependency in rough set theory means how some concepts of knowledge $K = \langle U, R \rangle$ can be expressed by other concepts of knowledge $K = \langle U, R \rangle$ (Pawlak, 1991). In other words, the set of attributes $R \subseteq Q$ depends upon the set of attributes $P \subseteq Q$ in S , if $IND(P)(indiscernibility\ relation\ P) \subseteq IND(R)(indiscernibility\ relation\ R)$

If there are two sets of attributes C and D (in the example, which is shown later, consider the two attributes sets as condition and decision), then to find the degree of dependency of attribute set C on attribute set D, the following formula can be used:

$$\gamma(c) = \frac{\left|\sum_{i=A\ and\ R} \underline{P}C_i\right|}{U} \tag{2.7}$$

where $|x|_c = \{C_A, C_R\}$, and U is the set of objects.

That measure of dependency expresses the degree of functional dependency of attribute set C on attribute set D. Sometimes this is called the quality of approximation.

2.1.2.2.6 Reduction of Attributes

If the knowledge is $K = \langle U, R \rangle$ and two different families of equivalence relations, R and R' (a subfamily of R), may give the same family of elementary sets, then there is a possibility of reducing R' while preserving the family of elementary sets, without losing a part of knowledge. In other words, the reduced set of attributes provides the same quality of sorting as the original set of attributes. The minimal subset $R \subseteq P \subseteq Q$, such that $\gamma P(У) = \gamma R(У)$, is called the У reduct of P and is denoted by $REDУ(P)$. R represents the reduced set of attributes that are important in decision-making. In other words, "The reduct in rough set theory is the minimal and sufficient condition for attributes of an object belonging to a certain group to be discerned from among all other objects" (Yanagisawa & Fukuda, 2007). The intersection of all У reducts is called the У −core of P ($COREУ\ (P)\ = \cap REDУ\ (P)$). The core is a collection of the most significant attributes in the system.

2.1.2.2.7 Decision Rule Synthesis

$DesC(X_i) \Rightarrow DesD(Y_j)$ is called a (C, D)-decision rule (condition, decision). In addition, the logical statement "If . . . then" can represent the decision rule, which relates the description of condition to decision classes. Decision rules can be deterministic $(iff\ X_i \subseteq Y_j)$ or nondeterministic. Nondeterministic rules are the results of an estimated description of decision classes. In other words, based on available knowledge, one is not able to decide if some objects belong to a given category or not.

2.1.2.3 An Example

To illustrate the mechanism of the rough set theory, a multicriteria sorting problem is considered as an example. This problem is solved using the following six-step procedure, which is also used in Chapter 3, to find the most influential user characteristics and also the most influential product features:

1. Construct information system table.
2. Define partitions.
3. Define indiscernible sets.

4. Calculate C-lower approximation (C refers to condition or attributes), C-upper approximation, and C-boundary (doubtful region) of set Y. Also, calculate the accuracy of approximation of set Y and quality of approximation (quality of sorting) of partitionΥ.
5. Construct minimal subsets of independent criteria (reducts subsets).
6. Generate decision rules from the reduced decision table.

The problem under consideration is how to select qualified candidates among those who want to obtain admission to a university. To gain admission, candidates must submit their application packages, including secondary school certificate, curriculum vitae, and letter of recommendation, from a previous school. These documents are evaluated by the admissions committee (decision-makers) who considers seven criteria. These criteria and their scales are as follows (ordered from best to worst value; for example in mathematics, scores can be 5 excellent, 4 medium, and 3 weak):

C1: Score in mathematics (5, 4, 3)

C2: Score in physics (5, 4, 3)

C3: Score in English (5, 4, 3)

C4: Mean score in other subjects (5, 4, 3)

C5: Type of secondary school (1, 2, 3)

C6: Motivation (1, 2, 3)

C7: Opinion from pervious school (1, 2, 3)

In this problem, specific decision-making rules are generated based on available knowledge to be used for making future decisions.

1. Construct information system table:

Table 2.2 shows the information system table. As shown in this table, the condition attributes are $C = \{C_1, C_2, C_3, C_4, C_5, C_6, C_7\}$, and the decision attribute is $D = \{d\}$.

TABLE 2.2 INFORMATION SYSTEM TABLE (PAWLAK & SLOWINSKI, 1994)

Candidate	Criteria							Decision
	C_1	C_2	C_3	C_4	C_5	C_6	C_7	
X_1	4	4	4	4	2	2	1	A
X_2	3	3	4	3	2	1	1	R
X_3	3	4	3	3	1	2	2	R
X_4	5	3	5	4	2	1	2	A
X_5	4	4	5	4	2	2	1	A
X_6	3	4	3	3	2	1	3	R
X_7	4	4	5	4	2	2	2	A
X_8	4	4	4	4	2	2	2	A
X_9	4	4	4	4	2	2	2	R
X_{10}	5	3	5	4	2	1	2	A
X_{11}	5	4	4	4	1	1	2	A
X_{12}	5	3	4	4	2	2	2	A
X_{13}	4	3	3	3	3	2	2	R
X_{14}	3	3	4	3	2	3	3	R
X_{15}	4	5	5	4	2	1	1	A

2. Define Partitions:

In this problem, there are two partitions—Y_A and Y_R ($\Upsilon = \{Y_A, Y_A\}$). Y_A is the set of candidates that are admitted by the committee, and Y_R is the set of candidates that are rejected. From the information table: $Y_A = \{x_1, x_4, x_5, x_8, x_{10}, x_{11}, x_{12}, x_{15}\}$ and $Y_R = \{x_2, x_3, x_6, x_9, x_{13}, x_{14}\}$

3. Define Indiscernible Sets

In this problem, there are 13 C-elementary indiscernible sets (C refers to condition or attributes): $\{x_4, x_{10}\}$, $\{x_8, x_9\}$, and 11 other sets, each of which includes one candidate (for example, both candidates x_8 and x_9 have the same conditions).

4. Identify Lower, Upper Approximation Sets, Doubtful Region, the Accuracy, and Quality of Approximation of Set Y

Based on what was discussed in section 1.3, in this step, the objects (candidates) that certainly belong to set Y_A are defined. This candidate set shows that the lower limit of Y_A is $\underline{C} = \{x_1, x_4, x_5, x_7, x_{10}, x_{11}, x_{12}, x_{15}\}$. In addition, the candidates that possibly belong to set Y_A are defined. This set of candidates represents the upper limit of Y_A, which is $\overline{C} = \{x_1, x_4, x_5, x_7, x_8, x_9, x_{10}, x_{11}, x_{12}, x_{15}\}$, where x_8 and x_9 are the candidates that have same conditions, but one of them is rejected and one of them is accepted. The difference between the upper and lower approximation constitutes the C-boundary (doubtful region) of set Y_A, which is Bnc (Y_A)= $\{x_8, x_9\}$. In addition, the same can be done for the rejected set Y_R. Results are as follows:

$\underline{C}Y_R = \{x_2, x_3, x_6, x_{13}, x_{14}\}$

$\overline{C}\, Y_R = \{x_2, x_3, x_6, x_8, x_9, x_{13}, x_{14}\}$

$Bnp\ (Y) = \{x_8, x_9\}$

Based on formulas that were mentioned in section 2.1.2.2.3, the accuracy of approximations of sets Y_A and Y_R by C are equal to 8 (cardinality of lower approximation) divided by 10 (cardinality of upper approximation), which are 0.8 and (5/7) = 0.71, respectively. The quality of approximation (quality of sorting) of decision by C is equal to ((8+5)/15) = 0.87.

As mentioned, x_8, x_9 are candidates that have the same conditions, but one of them is rejected and the other is accepted. This shows that the decision is inconsistent with the evaluation of candidates by criteria C. Therefore, the committee may want them to provide more information as an additional discriminatory criterion or to interview them (by creating a third category of candidates—those who should be interviewed).

5. Construct Minimal Subsets of Independent Criteria (reducts subsets)

Based on what was mentioned in section 2.1.2.2.6, the redundant criteria from the whole set of criteria could be eliminated so that removing them does not affect decision-making. Minimal subsets of independent criteria are obtained in such a way that they have the same quality sorting as the whole set condition attributes (C).

"Finding a minimal reduct among all reducts is NP hard." (Suraj, 2004) In this problem, there are three such reduct sets:

RED_{Υ}^{1}(C) = $\{c_2, c_3, c_6, c_7\}$

RED_{Υ}^{2}(C) = $\{c_1, c_3, c_7\}$

RED_{Υ}^{3}(C) = $\{c_2, c_3, c_5, c_7\}$

Decision-makers can take just use the reduced set of criteria to make a decision, exactly the same as when considering all criteria. The core set of all reducts set can be obtained by finding the intersection of all of them, which is as follows:

$$CORE_{\Upsilon}(C) = RED_{\Upsilon}^{1}(\mathrm{C}) \cap RED_{\Upsilon}^{2}(\mathrm{C}) \cap RED_{\Upsilon}^{3}(\mathrm{C}) = \{c_3, c_7\}$$

5. Generate Decision Rules from the Reduced Decision Table

If reduct set 2 ((RED_{Υ}^{2}(C)) is considered the reduct set in this problem, then the decision table can be reduced to criteria c_1, c_3, c_7. Then the decision rules shown in Table 2.3 are generated from the reduced decision table.

TABLE 2.3 DECISION RULES GENERATED FROM THE REDUCED DECISION TABLE

Rule No.	If	C_1	C_3	C_7	Decision
1	If	5			Admit
2	If		5		Admit
3	If	4		1	Admit
4	If	4	4	2	Admit or Reject
5	If	3			Reject
6	If		3		Reject

As Table 2.3 shows, five rules are deterministic and one is nondeterministic (rule no. 4), because x_8 and x_9 belong to the boundary region (doubtful region). Now, based on the above decision rules, the policy can be expressed as follows:

"Admit all candidates having a score of 5 in mathematics or in English. Also, admit those who have a score of 4 in mathematics and in English, and a very good recommendation from a previous school. Invite the candidate to an interview if he/she has a score of 4 in mathematics and in English, but only a moderate opinion from a previous school. Candidates having a score of 3 in mathematics or in English are to be rejected."

In this section, a rough set model was developed to obtain some decision rules that create decision policy. In fact, in this problem, the rough set method used available information that was obtained from decision-makers (information system). Based on this information, the rough set method tried to reach fundamental, straightforward, and consistent decision rules and decision policy that could be used as an instruction for decision-makers for similar future decision-making cases. "We are able to drive implicit facts from explicit and unquestionable facts (knowledge) about a decision situation" (Suraj, 2004).

2.1.3 Traditional Product Development Method

The approach explained in this chapter is based on rough set theory. As mentioned earlier in this chapter, usually the product development team divides a heterogeneous market into different customer segments and targets one or more of them; then, they determine the requirements and expectations of each customer segment, prioritize the requirements, and use them to find the quality elements. This is shown in Figure 2.5, and each step of the traditional approach is briefly explained.

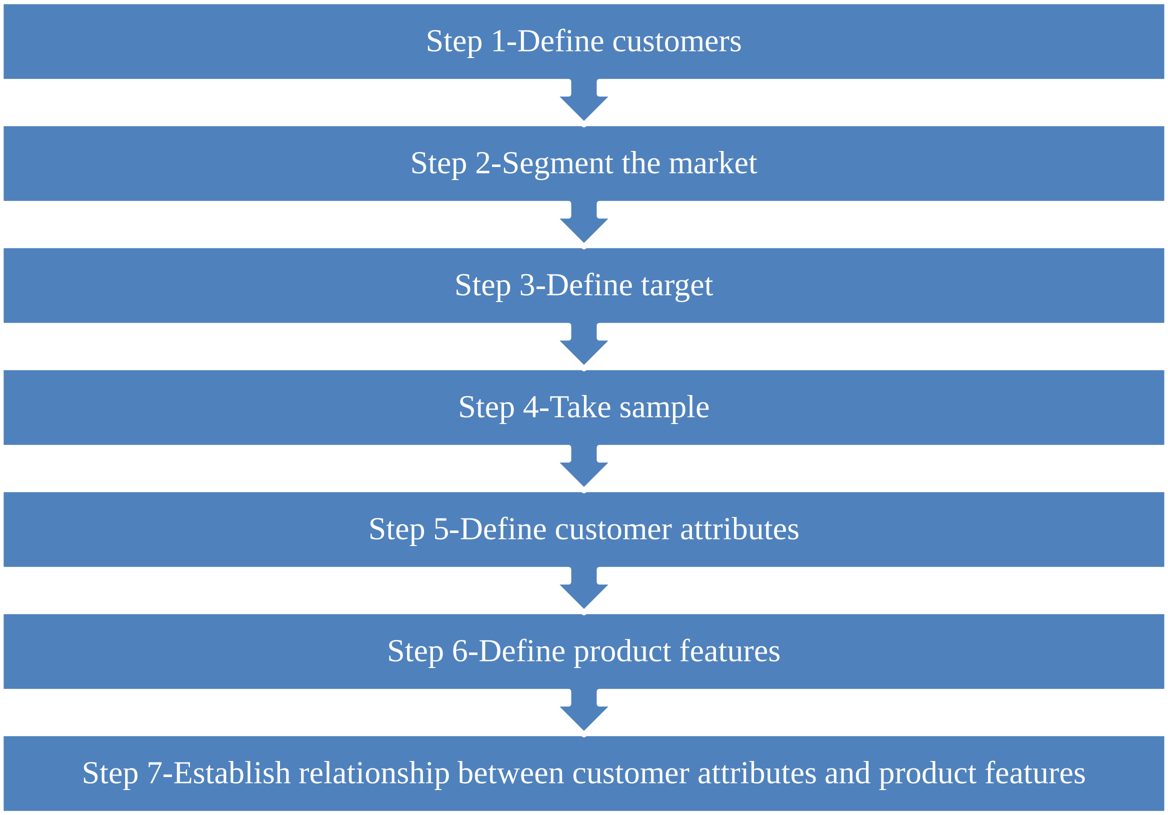


Figure 2.5. Traditional product development approach using customer requirements.

2.1.3.1 Define Customers

In order to satisfy customers, a company should first define the customers of its products and then concentrate on their requirements. There are many methods to define group(s) of customers to be considered in the product design process. Juran (1992) suggests the Pareto analysis method. This method can help the company define its customers more precisely. Factors that can be used as a basis for this analysis could be sales volume or strategic importance. Also, Shillito (2001) proposes other tools, such as the customer chain concept and customer morphology. For example, customer morphology will focus on the world in which the product is used.

2.1.3.2 Segment the Market

Usually customers are segmented by the corporate finance group so the company can track the financial results of each segment and so that segmented customers can be targeted easily with advertising and marketing programs (Ulwick, 2005). Moreover, market segmentation can be used to define targeted groups of customers for product development projects. Customers can be segmented based on age, gender, income, socio groups, family life cycle, lifestyle, product usage, and so on (Hill & Alexander (2000); Armstrong & Kotler (2006); Weinstein, (2004)). In market segmentation, it is very important to identify what kind of segmentation results are the largest customer requirement differences between groups. In fact, market segmentation should be such that the homogeneity within segments and the heterogeneity between groups are increased.

Using an air conditioner as an example, segmenting customers based on geographical location where users live is meaningful, but segmenting them based on religion or gender does not make sense, because there are no significant differences between the customer requirements of different users who use an air conditioner and those who have different religions or are of a different gender.

2.1.3.3 Define Target

In this step, depending upon the company's strategy, one or more segments of the market could be selected. Two important factors should be considered when selecting a target: attractiveness of the segment, and fit between the segment and the firm's objective, resources, and capabilities. Attractiveness of the segment includes size of the segment, competition in the segment, and growth rate of the segment. However, the company should evaluate each segment in terms of whether or not the firm can offer superior value to the customers in the segment, access to distribution channels required to serve the segment, and so forth (http://www.netmba.com/marketing/market/target/).

2.1.3.4 Select Sampling Method

In this step, the most appropriate sampling method is selected. Three important factors should be considered in sampling: sampling reliability (precision, confidence level, and variance), sample size, and sampling plan. The sampling plan could be a random sample (probability sample), including a simple random sample, stratified random sample, and cluster sample; or the sampling plan could be a nonrandom sample, including convenience sample, judgment sample, and quota sample (Hill & Alexander, 2000).

2.1.3.5 Define Customer Attributes

As mentioned previously, customer attributes can be divided into two types: concrete and abstract. Functional requirements and Kansei can be considered abstract attributes. One of the main parts of all product development approaches is defining and prioritizing either customer needs or customer feelings. The ways in which to do this are almost identical. In gathering customer demands, recognizing what customers really want is critical. In some cases, customers offer solutions to some of their unsolved problems. Alternately, sometimes customers express their needs in terms of problems. Obviously, these are not their real requirements. Real customer requirements are behind those needs. For instance, consider that a customer wants a splashguard as standard equipment on his van. This is one of the solutions for his/her real needs, which is to prevent rust around fender wells (Shillito, 1994). These kinds of requirements should be explored using the right questions.

Identifying customer requirements is a process. One well-known process is voice of the customer (VOC) (Shillito, 2001). This process could be attached to other processes, such as quality function deployment, or the customer satisfaction and loyalty measurement (CSLM).

Usually the output of the VOC process is the input of the QFD and CSLM processes. The aim of VOC is to provide a prioritized list of customer requirements in an efficient way so that these requirements accurately represent customers' needs. This process is briefly explained below.

2.1.3.6 Define Scope and Objectives of VOC

In this step, clear and realistic boundaries and objectives for the VOC process should be defined. The following questions should be answered: Why is this project being done? What is the purpose of gathering customer requirements? Is this project part of another project? Will this project help designers design a revolutionary or evolutionary design? Will this project increase customer satisfaction and loyalty? Will it reduce costs; increase productivity; or improve quality, reliability, or manufacturability? What should be included in this survey? What is not included in this study? What are the boundaries, and controllable and noncontrollable factors? (Hill & Alexander (2000); Shillito (2001)). Answering these questions will define the scope and purpose of the VOC study.

2.1.3.7 Use VOC Collecting Techniques

There are many methods to communicate with customers and gather their requirements. Most methods used in VOC actually come from a survey-sampling context. Scheaffer (1996) explains the methods of data collection and considers them as sources of errors of observation. These methods are personal interviews, telephone interviews, self-administered questionnaires, and direct observation. Among these techniques, direct observation is usually less used in surveys that do not involve measurements on people. For example, health information can be obtained from hospital records, or income information can be gotten from employer's records. In addition to the above techniques, Juran (1992) adds three more techniques: visits to customers, partnerships, and focus groups. Anton and Perkins (1997) introduced computer-assisted telephone (CAT) surveys as another technique to collect VOC. Immediately following the completion of a call, a predetermined sample of callers is automatically transferred to a CAT survey system. This system offers the survey introduction in a recorded human voice and begins asking the survey questions. Actually, the CAT system acts as a human telephone interviewer. Perhaps the most complete list of VOC collection techniques is that provided by Shillito (2001). These techniques are surveys, mailings, comment cards, interviews, phone calls, focus groups, location studies (like an industrial engineering survey), direct observations, visitations to the site of product usage, internal brainstorming, commercially prepared stock reports, panels, electronic databases and searches, service calls, 1-800 hotlines, in-context customer visits, and the Internet. Some of the most important of those techniques are explained here briefly.

Location studies are direct visits that allow direct observation of operations and sometimes direct interaction with operators or the process. Sometimes the product development team can brainstorm customer requirements. Contextual inquiry or in-context customer visits is a process that involves a visit to the customer in the real-work environment. The aim here is to observe and participate in the customer's experience on his/her turf as it relates to product and service that VOC is going to be gathering. This concept is similar to a Gemba visit, a Japanese word that simply implies a customer visit. The idea of Gemba originated from the occurrence of a problem, whereby engineers would go to the site to understand the full impact of the problem, gathering data from all sources. In the VOC context, Gemba is the place where the customer uses the product, and the customer can provide realistic requirements. A Gemba visit has many advantages. For example, sometimes a customer expresses wishes or needs while working. These are needs that could be forgotten in a structured interview (http://en.wikipedia.org/wiki/Gemba).

The source of VOC data could be internal (such as marketing, sales, or customer service), direct customer interaction, trade shows, sales calls, service repair calls, literature (trade and consumer), complaints/warranty records, visitations, or electronic data bases and searches ((Shillito, 2001).

2.1.3.8 Use Exploratory Research

Exploratory research should be performed to gain preliminary knowledge about customer requirements. This research can be done by design and involve an in-depth interview or can use a focus group to explore customers' hidden attitudes and perceptions. The results of this step can be used in designing the questionnaire (Hill & Alexander, 2000).

2.1.3.9 Design a Questionnaire

In this step, a questionnaire is structured. This can be done by understanding the basic principles of asking questions and using different kinds of rating scales to design the questionnaire: Likert scales, verbal scales, semantic differential scales, numerical rating scales, ordinal scales, or Simalto scales. Anton & Perkins (1997) and Hill and Alexander (2000) provide all necessary instructions.

2.1.3.10 Use Piloting

Before performing a sampling, the designed questionnaire should be piloted to observe and test how well it works in practice. In piloting, wherever respondents hesitate or appear puzzled should be monitored and recorded (Hill & Alexander, 2000). The questionnaire should be modified based on piloting results.

Affinity diagrams and tree diagrams can be used to structure the data. In addition, there are many other methods to determine how important each requirement is to the customer. Simple and alternative ranking, pair comparison, and direct magnitude estimation are the most common methods. Each of these techniques has advantages and disadvantages (Shillito, 2001).

2.1.3.11 Define Product Features

As mentioned previously, the product development team will define product characteristics. Technical documents and magazines, relevant literature, manuals, and experts could be some of the relevant sources.

2.1.3.12 Establish Relationship Between Customer Attributes and Product Features

This step is the central part of all product development methods. It clarifies which product features will impact customer requirements. Many methods are used to define this relationship. Either qualitative estimation methods (usually estimated by the product development team), which are used mostly in QFD and type one Kansei Engineering, or more sophisticated methods, such as multivariate analysis or artificial neural network, are used. The outcome of these methods is the most influential product features on customers' functional requirements, perceptions, and feelings.

2.2 Literature Review

2.2.1 Multiple Users in Product Development Approaches

As mentioned in section 2.1.3, one of the main components of the product development approach is customers' requirements, needs, or feelings. The aim of this approach is to find the most influential product features influencing customer requirements. Customer requirements can come from one user or many users who are in the same group. The problem that generally will arise is the large amounts of data (customer requirements) that could be gathered from one customer or multiple customers. This leads designers to structure and quantify customer requirements in order to reduce and focus the number of customer requirements so that designers can respond more effectively to those requirements. In other words, these structuring and quantifying tools can direct designers to those critical elements that will lead them to better product and service design.

Many techniques have been developed to structure the data, such as an affinity diagrams, interrelationship diagraphs, or tree diagrams (Shillito, 2001). The purpose of these techniques is to see relationships among customer requirements and the hierarchy of their importance. In addition, many techniques are available to quantify customer requirements such as Simple ranking, alternative ranking, pair comparison, direct magnitude estimation, and category scaling (Shillito, 2001). For example, dual scaling is one of the techniques that have been used in some Kansei Engineering applications to calculate the relative importance of Kansei. Nishisato developed this technique in 1980. Dual scaling captures linear and nonlinear relations among variables and is an “optimal method to extract a maximum amount of information from multivariate categorical data” (Kaplan, 2004).

Although one of the reasons to group customers is to reduce the variation within a customer group, elimination of all variation is not possible. The point of the above techniques is to ignore this variation by listing all customer requirements that come from all customers in a specific group and then quantify (prioritize) them for perceived importance to customers through different methods. This process produces a reduced set of customer requirements or perceptions by selecting the most important ones among all requirements of all customers in a particular group. During this combining, some requirements that might be important for subgroups of customers could be ignored or considered less important than other requirements corresponding to the dominant customer subgroup, since, for example, taking the average or other central tendency measures are based on the relative frequency of the attributes.

In fact, all of these processes are based on the assumption that customers in the same group have almost the same requirements and that there is no significant difference among their requirements. Also, in terms of importance, it is assumed that the average importance of customer requirements represents the importance of that particular requirement for all customers. Nevertheless, obviously this average cannot reflect each individual perception. Figure 2.6 shows the issue of multiple customers in customer-oriented product development.

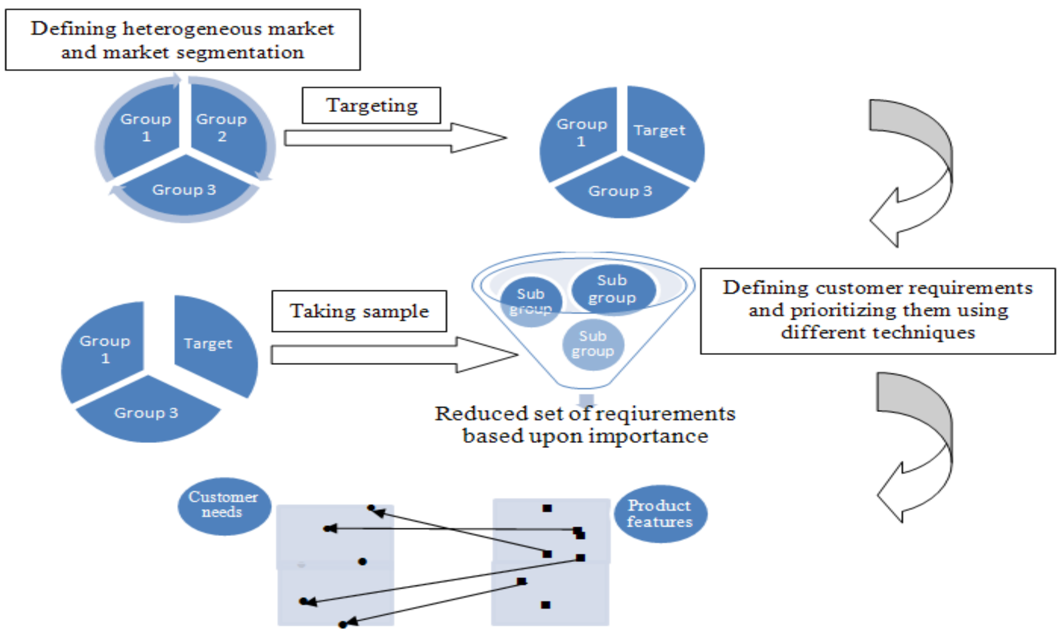


Figure 2.6. Multiple customers in product development approach.

As an example of this kind of process, Ahmady and Zegordi (2002) added a pre-house of quality (HOQ) matrix, called the "evaluation of the importance of the customer requirement matrix," to a QFD process. The requirements of several of the most important targeted customer groups who have different importance for designers were placed in this matrix. This matrix calculates the overall customer requirement importance when there are correlations between customer requirements. The importance of each customer group for designers is not the same, and the importance of each customer requirement for each group, which is taken from a central tendency measure calculation, is defined. The output of this matrix is the input of the HOQ matrix in the QFD process.

In summary, product development approaches consider a sample of customers as one unit, regardless of differences that exist between customers' points of view. In addition, they take into consideration one central tendency measure or rank as a representation of importance of a particular customer requirement for all customers who are in the same group. This makes these approaches inaccurate. Figure 2.6 illustrates the traditional product development approach along with multiple customers' issues.

2.2.2 Multiple Users in Kansei Engineering

The aim of Kansei Engineering is to incorporate customers' feelings into product function and design. In this context, customers could be one user or a group of users. Therefore, both single users and multiple users have been considered in KE. Figure 2.7 shows different types of Kansei Engineering that are classified in terms of either a single user or multiple users, and then each category is explained briefly.

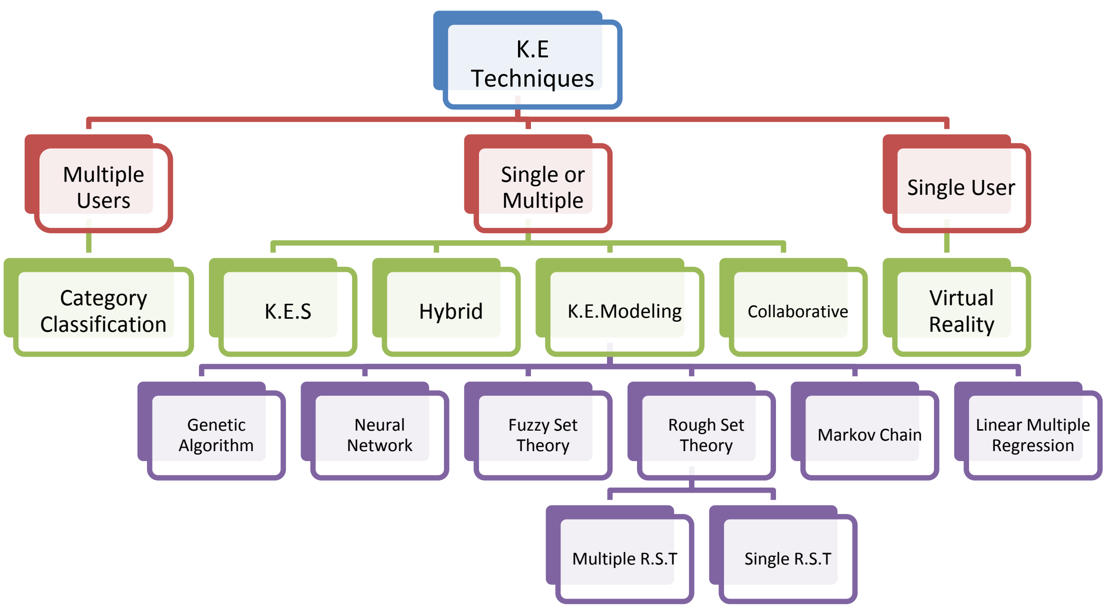


Figure 2.7Kansei Engineering classification based on single user and multiple users.

As Figure 2.7 shown, “virtual reality” can take into account a single user’s feelings, although theoretically it could be used for several individuals. In other words, it can help designers to customize a specific product for a particular user (Oyama, 1997). For example, if virtual reality Kansei Engineering was used to customize a kitchen for a customer, it could decrease the number of consultations and consultation time between the designer and the customer in the design process (Imamura, Nomura, Tamura, & Nagamachi, 1996).

Kansei Engineering type one uses a qualitative treatment to link Kansei words and product features. In fact, the principles that are used in this method are almost the same as the QFD method (Schütte et al., 2004). In both methods, feelings (in Kansei Engineering) or functional requirements (in QFD) are taken from multiple users. KE type one is capable of considering a single user’s needs. Usually the importance of customer feelings is calculated across multiple customers using different methods. For example, Petiot and Yannou (2004) wanted to design a customer-oriented glass. They asked ten people to assess each sample of glass on a seven-point Likert scale with respect to a specific Kansei. One subject, whose assessment was very different from the rest of the groups, was removed from their study. Then for each glass and each Kansei, they calculated the average value of the assessment for ten subjects and considered that average as the perceived importance for all users. This quantified data was used as the input of the synthesis step in Kansei Engineering.

Another example is what Grimsaeth (2006) did to improve the t-shape portable/cordless/ rechargeable drills that are used as home tools. The target group was younger adults between 20 and 30 years old. It was found that only a few products on the market were directly aimed at this particular market segment. Different sources, such as magazines, manuals, and specialty users, were used to gather Kansei words. Participants were ten design students, five of each gender, who were asked to evaluate each sample drill. Again, the mean values for the 25 Kanseis describing each of the thirteen drill samples were considered. No significant differences between male and female ratings were identified.

Many other applications and explanations regarding the Kansei Engineering type one approach can be found (Schütte et al. (2004); Grimsaeth (2006); Nagamachi (1995); Nagamachi (1996); Nagamachi (1997); Nagamachi (1999); Nagamachi (2000); Nagamachi (2002); Petiot & Yannou (2004); Hon Yee Siu & Ho (2005); and Horiguchi & Suetomi (1995)).

The Kansei Engineering system, hybrid Kansei Engineering, and the collaborative system use almost the same approach to deal with the multiple-user issue. These three types of KE can handle both single-user (Matsubara & Nagamamachi, 1997) and multiple-user problems. As mentioned in the background section, all three methods have a computerized expert system that supports the transfer of customers' Kansei and perceptions into physical elements. The method that is usually used in the synthesis step in these three methods is Hayashi's quantification theory type I. This method is a kind of multivariate linear regression analysis for qualitative data, which determines the correlation between Kansei word rankings as dependent variables and product features as independent variables. The method defines the amount of negative and positive influence of product features on each Kansei factor (Grimsaeth, 2006). Here, users or subjects are asked to evaluate photographs of different product samples with respect to a seven-point semantic differential scale, which is a bipolar-scale rating measure. The method considers all users' data as one sample and assumes that there are no significant variations within the sample (Nagamachi (1994); Shimuzu, Sadoyama, Kamijo, Hasoya, et al. (2004); Bianchi-Berthouze,(2001); Matsubara & Nagamamachi, 1997).

Kansei Engineering modeling uses mathematical approaches including those for linear data (such as linear regression analysis) and nonlinear data (such as fuzzy set, artificial neural network, and rough set theory). Using this kind of KE, a product can be designed for either one user or multiple users. For example, to improve the design of a keyboard switch, Kosaka, Nishitani, & Watanabe (2005), took a sample of 54 students and asked them to evaluate a specific keyboard switch with respect to their Kansei, while they changed each of the 15 keyboard characteristics to generate virtually different types of switches. Then they used the dual scaling method to define each Kansei value. This technique eliminates the “users” dimension in this problem. The values scaled by dual scaling for each Kansei were the input, and the set of characteristic values of a keyboard switch was the output of the neural network. Table 2.4 shows the format of input data of a neural network that can be used for training data to build the network. As shown, dual scaling calculates the score of each Kansei for each switch across different users using the method of reciprocal average (MRA) (Kaplan, 2004).

TABLE 2.4 FORMAT OF TRAINING DATA OF NEURAL NETWORK FOR KEYBOARD SWITCH DESIGN

	Kansei-Input (calculated by dual scaling)					Keyboard Switch Characteristics-Output				
	Semantic Scale					Normalized [0,1]				
	Very	More	Neither	Less	Little	Feature 1	Feature 2	…	Feature 15	Switches
Kansei 1	0.34	0.13	-0.05	-0.13	-0.24	0.2	0.3		0.6	Sw1
Kansei 2										Sw1
….										Sw1
Kansei 7										Sw1
Kansei 1										Sw2
….										Sw2
Kansei 7										Sw2
Kansei 1										Sw3
…..										….
Kansei 7										Sw3

Other applications of neural networks in Kansei Engineering have used the same approach (Ishihara, Ishihara, Nagamachi, Matsubara, and Yukihiro (1997); Shigekazu, Keiko, & Nagamachi (1996); Ishihara, Ishihara, & Matsubara, (1995)).

In summary, Kansei Engineering problems have three dimensions: subjects, concepts, and scales. Usually in a KE problem, a group of subjects (x-axis) judges a sample of concepts (y-axis) against semantic scales (z-axis) (Schütte et al., 2004). Figure 2.8 depicts three dimensions of a general Kansei Engineering problem in the format of a three-dimensional matrix. In this matrix, each cell contains a number, which is the judgment of a particular concept on a particular scale by a single subject. The research proposed here is looking for an approach to map the subject dimension of this matrix onto the scale-concept surface efficiently so that eventually the concept(s) are matched with different subjects' perceptions. Since most available approaches, including statistical approaches, need two-dimensional data, one of these dimensions is reduced using different methods, for example, by taking the mean over subjects.

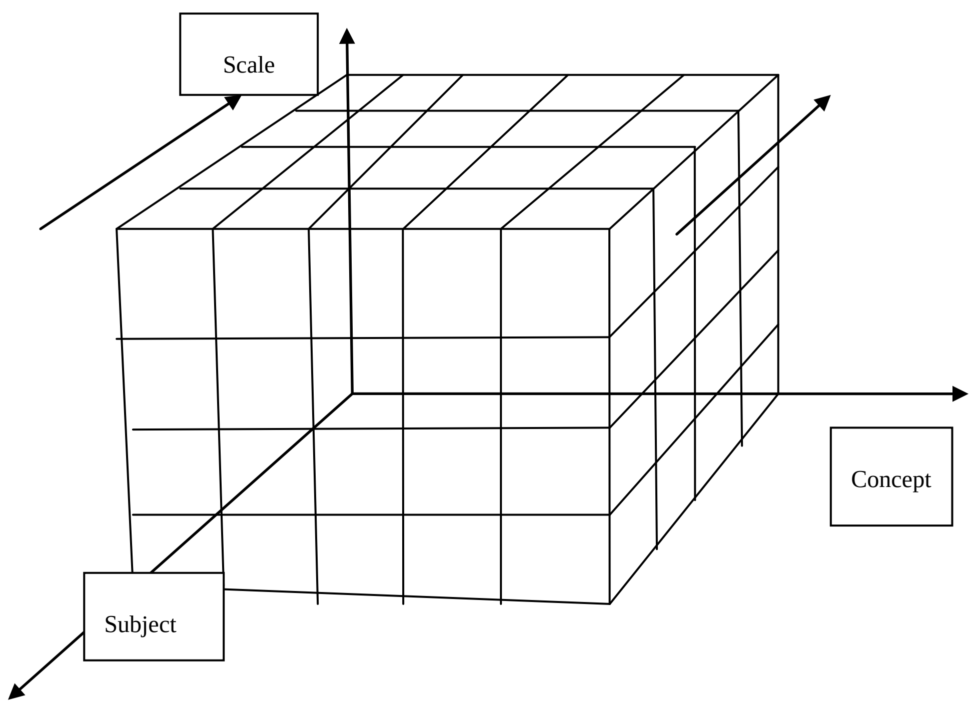


Figure 2.8. Three dimensions of a general Kansei Engineering problem
(Schütte, Eklund, Axelsson, & Nagamachi, 2004).

Specifically in Kansei Engineering and also in general, in customer-oriented product development approaches, after the relevant Kanseis or functional requirements have been gathered and rated on semantic scales, the number selected is reduced in a way that the remaining Kanseis properly represent the semantic space. This process has two main steps. First, data are collected from multiple users by taking a sample, preparing a list of customers' feelings, and rating them by users so that the perception of different subjects is reduced to the perception of one specific dummy subject. Then, the number of Kansei is reduced by selecting the most important ones.

2.2.3 Multiple Users in Kansei Engineering Using Rough Set Theory

A review of the state of the art showed that, so far, there are few applications of rough set theory in Kansei Engineering. Only five applications were found. In two of them, the products were customized for one user (Yanagisawa & Fukuda, 2007).

2.2.3.1 Single-User Applications

Yanagisawa & Fukuda (2005) developed a new method based on rough set theory to help a product's user be involved in the design process. In this method, interaction occurs between a user and a computer. First, the user evaluates different designs based on his or her Kanseis (feelings), and then the system refines the characteristics that are more favorable for the user using the reduct concept in rough set theory. These characteristics are transferred to the next generation of design, mixing with some random features, and again the user decides on the new generation of design and evaluates the characteristics. This process continues until a satisfactory design is achieved. This method was applied to designing and developing a three-dimensional cylinder, such as a cup. Then the results were compared with the results of applying interactive genetic algorithm for these kinds of products.

The core of this method uses the reduct concept in rough set theory. A reduct set can estimate the most attractive product features for a user. Although until recently reduct sets in rough set theory were used for design, many applications have used reduct for analysis. The difference between this method and other similar methods that use reduct for synthesis (in design) or analysis, is that this method uses reduct sets to create the interactive process that goes back and forth between synthesis and analysis. This proposed method is different for discrete data and continuous data. Discrete data in design are the characteristics that have limited possible options, such as large, small, square, or cubic. Continuous data are the features that have unlimited options, such as length or weight. The methods for discrete data and continuous data are almost the same with very few small differences. In general, in both methods, the features values of initial design samples are generated randomly. Then the user evaluates design samples as good, bad, or normal. The system calculates “good” and “bad” reduct sets considering the evaluation of customers for all sample designs. Then the system stores good reduct sets into an inheritance reduct set, and bad reduct sets are deleted from the inheritance reduct set. Now the inheritance reduct set contains only good reduct sets. Reduct sets regarding earlier evaluations are deleted from the inheritance reduct set. Then N reduct sets are chosen from among the good reduct sets in the inheritance reduct set, and new design samples are generated, mixing good reduct sets and some random features for which the reduct sets do not have those specific features. This process will continue until the user achieves an acceptable satisfaction level. To check the effectiveness of their method—interactive reduct evolutional computation (IREC)—the authors applied it to improve the design of a cylindrical-shaped cup or vase and conducted some experiments to compare IREC and other similar methods that used genetic algorithms in terms of two criteria: effectiveness and convergence of solution. Results showed that IREC could satisfy users more often than the interactive genetic algorithm (IGA) with less iteration. Convergence of a solution is an important indicator of system optimization. Convergences defined as the design process that came closest to the user's favorite product image, indicated that IREC converged better than IGA.

In summary, Yanagisawa & Fukuda (2005) proposed a method in which one user is involved with a design process where there is interaction between the user and the computer. Using this system, an individual user can design a product matching his or her feelings. Based on user input, the system creates an inheritance reduct set using rough set theory and then generates sample designs that are evaluated by the user. Once the new data is obtained from the user, the inheritance reduct set is updated based on new data. This guarantees that the user's favorite features are transferred to the subsequent generations, and the favorite product features contributing to the final design are distinguishable and traceable. These two aspects are two advantages of this method over other similar methods. Moreover, through some experiments, results show that with the IREC, user satisfaction level, speed of the solution search, and convergence of the solution are better than with another similar interactive method based on a genetic algorithm.

Yanagisawa & Fukuda (2007) also looked for a way to involve a user in the design process. They attempted to design a system that was capable of interacting with a user and giving the user the ability to see his or her design and try to improve it. This system is a computer-based system and uses rough set models. First, the system generates some design samples choosing features randomly and shows them to the user. The user evaluates these samples and scores each of them. Then the system creates reduct sets, does some computations, and generates new design samples. This process continues until there are few large differences between the initial design and the subsequent design, and the user achieves a particular satisfaction level.

Yanagisawa & Fukuda (2005) applied their proposed method (IREC) to the design and outline of an automotive sideboard. The design of this component involved curved shapes. Curve surfaces influence aesthetics and improve the aerodynamic performance. However, it is difficult for users to explain in words about curved shapes. IREC can find the value regarding the shape-of-curves parameters that matches with a user's imagination. Usually a human first sees the product in global features, such as roundness, volume, and convexity/concavity, and then views it in detail. Here, the authors enhance their IREC system to support global attributes as design characteristics to apply to the industrial design shaped by curves. Then global features are computed from the model parameters. Each design is evaluated, and the one closest to the user's feelings is selected. If the user cannot find any appropriate design, then the system generates other N samples, and this will continue until more than one sample is selected. Now there are two groups of designs: one group of "selected design" samples and another group of "not selected" design samples. Then the system calculates the reduct set of selected designs. This reduct set is considered a set that contains characteristics that users paid more attention to than others. This kind of calculation can handle nonlinear data. Then Global IREC (GIREC) calculates two kinds of sets: one for global features and another for model parameters. The combination of these two sets are stored in an inheritance reduct set, which is responsible for transferring the good and random features to the next inheritance reduct set (IRS) to use them for generating new design samples. Then the system generates new design samples, and this process will continue until the user is satisfied. In order to verify the effectiveness of GIREC, a comparative experiment of GIREC and IREC for designing an automobile sideboard was carried out. Results showed that in terms of user preference between the two systems, there was no difference between them, but due to introducing global featuring to the IREC, the speed of finding a design that roughly matched the user's Kansei for GIREC was greater than for IREC.

In summary, helping a user to externalize his or her feelings is always a challenging issue for designers. The system proposed in this paper helps a user to be individually involved in the design process. Simply by viewing the sample designs, the user evaluates the samples and scores them according to a three-class scale: good, normal, and bad. It is very easy for a user to distinguish design samples by using only these three words, thus not needing to explain his or her feelings in detail. The proposed system takes this simple input, interprets the results using a reduct set in rough set theory, and concentrates on global rather than detailed product features. The authors who proposed the IREC method in their previous article and in this study carried out a comparative experiment to compare the effectiveness of these two methods. Results showed that both methods could give the same satisfying results. Moreover, GIREC determines more quickly than IREC the design that roughly matches the user's Kansei, usually within the first two generations.

2.2.3.2 Multiple-User Applications

Three other papers were found to have used rough set theory in Kansei Engineering for multiple users. Nagamachi et al. (2006) used rough set theory in KE to treat nonlinear Kansei data, in addition to other similar approaches such as artificial neural network, genetic algorithm, and fuzzy logic. They noted that using these kinds of methods in KE in which Kanseis usually have nonlinearity properties will lead to more accurate results in comparison to multivariate statistical methods such as factor analysis; multiple regression analysis; quantification theory types 1, 2, 3, and 4; and cluster analysis. To prove this claim the authors compared the output of Kansei Engineering type one (as a standard) and rough set KE for developing a child's shoe. Results for upper approximation using the rough set method and quantification theory in Kansei

Engineering were very similar, while the results for low approximation using the rough set method led to a more inventive shoe.

Regarding the issue of multiple users in this paper, the authors did not mention anything directly, but it seems that they used the results of users' evaluations, which they obtained in the Kansei Engineering type one process, in rough set KE. In the execution of the Kansei evaluation experiment, a group of 26 young mothers who had kindergarten-aged children was asked to evaluate their feelings on each shoe (29 different shoes from five manufacturers). Then, all subjects discussed together which shoes matched their Kanseis, such as easy to wear, safe running, comfortable usage, easy to wash, etc. After this discussion, all subjects agreed on one specific number on a five-point Kansei semantic differential scale, so that there was no inconsistency between multiple users. Obviously, in this method, it is possible that some people in the group did not agree on the output of the group.

Also Shimohara & Shimazaki (2005) used a rough set model to find the most important combination of elements of walking space that affected each pedestrian's preferences for comfort. Once pleasant and unpleasant "decision rules" for each person sampled were obtained, the common (intersection) pleasant decision rules across all people were defined. Then those decision rules that were unpleasant for each person were removed from the common decision rules. Although this method is a good method to obtain agreement on walking space characteristics that influence all participants' perception, if a particular decision rule is pleasant for all people except one, then it is removed from the list of pleasant decision rules.

In another application of rough set theory in Kansei Engineering, Nishino, Nagamachi, & Tanaka (2006) proposed an approach for situations where there is no lower approximation of classification. This means that the decision classes embody considerable inconsistency and uncertainty. As will be explained in Chapter 3, once multiple users evaluate the same product, their evaluation processes include the cognition of interactions between elements of the product and the ambiguity of decisions arising from individual differences. Therefore, these differences create a situation where one cannot say for sure which particular product belongs to which user perception category. In this case, there is no lower approximation of a classification. Therefore, for such a case, a rough set method based on both the Bayesian rough set (BRS) and variable precision Bayesian rough set (VPBRS) was developed. In this approach, it was assumed that decision classes of human evaluation occur with different prior probabilities. Then a lower approximation of decision classes is defined by introducing the information gain. The information gain is considered the difference between prior and posterior probabilities. It is assumed that the information gain evaluates the impact of the set of conditional attributes (product features) on a decision class relative to its prior probability. The positive region is the region that includes all products that would possibly belong to a specific class of users' impressions. Any product for which the corresponding gained information is greater than the specific threshold is placed in this region. On the other hand, the negative region is where the products would not possibly belong to a specific perception class. Accordingly, the boundary region in this paper is defined as the region where products do not belong in either the positive or negative regions. After these regions are defined, the decision rules from these approximate regions are extracted by using a process that has two stages. In the first stage, certain decision rules are defined. Then, in the second stage, rule evaluation factors are used to extract the rules. This proposed approach was applied to coffee taste design in a coffee company. The aim was to find effective decision rules and to develop coffee manufacturing conditions to produce a new coffee taste fitted to customers, according to extracted decision rules. Results showed that this approach was applicable and powerful for practical problems in Kansei Engineering.

In summary, a review of the state of the art showed that, generally, rough set theory has not been applied in Kansei Engineering in many cases. Single-user and multiple-user situations were reviewed in the literature review.

As in other product development approaches such as QFD, in most reviewed cases, uncertainty and inconsistency in input data are not seen as a challenge since the group of customers are considered one unit. However, the subjectivity of functional and psychosocial customer requirements affects these approaches primarily in two ways. First, inconsistencies within customer points of view regarding a typical product affect customer importance ratings. Also, the effect of inconsistency of multiple users is reflected in the competitive assessment part of the approaches, where different customers assess the performance of the current organization’s and competitors’ products with respect to each customer’s requirements. Either using a ranking system or taking an average can result in some sort of bias in the approaches process and invalidate the results.

The approach proposed in the Chapter 3 not only provides a method to use rough set theory in Kansei Engineering but also offers a way to enhance the customer-oriented product development approaches when there are multiple users. In the next chapter a rough set-based approach is proposed to identify the natural classification of product users based on the similarity of their assessments regarding functional and psychosocial requirements (Kanseis).

CHAPTER 3

PROPOSED WORK

This chapter will define the problem using notations of rough set theory, and describes the approach to solve the problem. To show clearly how the suggested method works, an example using fictional data is shown. In order to show the computational aspect of the problem, the simplified problem is solved manually.

3.1 Problem Definition

In order to develop originating requirements for complex products for costumers with diverse needs and point of view, it is typical to consider the concrete attributes more than the abstract attributes. Many approaches have been developed to convert abstract attributes to concrete attributes. Customer diversity is a subject that has not been discussed in previous approaches. The research here looks for a way to resolve some uncertainties that might be caused by customer diversity using a product development approach such as Kansei Engineering.

As stated in Chapter 1, markets are usually heterogeneous. Consequently, each user may perceive a product differently. Since pure customization and generalization are wasteful and/or costly, companies should position their product design between these two extremes. This creates a challenge for designers to maximize product value for customers while considering the company's interest. This problem may be stated in technical terms as follows.

Suppose that in the process of product development, a sample of "P" customers is asked to evaluate "N" products. This finite set $U \neq \emptyset$ (the universe) of objects (products) is the set in which the designers are interested. Any subset $X \subseteq U$ of the universe is called a concept or category in U. In addition, any family of concept in U is referred to as abstract knowledge or, simply, knowledge.

For each product, at most, "M" features can be considered (hereafter referred to as "condition attributes"). Customers are asked to rank each product with respect to one specific Kansei in "K" different levels (referred to as "classes"). In general, these products can be classified based on different levels of one or several features, or the customers' feelings (Kanseis) that are evaluated by multiple users. In this study, the latter is considered. To make this clear, an example follows. Suppose the set of product blocks U are classified according to "w" different users' Kanseis or perceptions. By these classifications, "w" equivalence relations are defined as R_1, R_2… R_w. For example, if $U = \{x_1, x_2, x_3, x_4, x_5, x_6, x_7, x_8\}$, and these products are evaluated based on three different users' Kanseis, such as reliability (less reliable, medium reliable, more reliable), harmoniousness (less, medium, more), or rigidity (bad, good), then three equivalence relations $R1, R2, R3$ could be defined by users as follows:

$$U/\,R1 = \{\{x_1, x_3, x_7\}, \{x_2, x_4\}, \{x_5, x_6, x_8\}\}$$

$$U/\,R2 = \{x_1, x_5\}, \{x_2, x_6\}, \{x_3, x_4, x_7, x_8\}\}$$

$$U/\,R3 = \{\{x_2, x_7, x_8\}, \{x_1, x_3, x_4, x_5, x_6\}\}$$

These are considered elementary concepts (categories) on the knowledge base $K = (U, \{R_1, R_2, R_3\})$. In addition, basic categories are set theoretical intersections of the elementary categories. For example set $\{x_1, x_3, x_7\} \cap \{x_3, x_4, x_7, x_8\} = \{x_3, x_7\}$ is $\{R_1, R_2\}$, which is a basic category that is less reliable and more harmoniousness. Alternately, set $\{x_1, x_3, x_7\} \cap \{x_3, x_4, x_7, x_8\} \cap \{x_2, x_7, x_8\} = \{x_7\}$ is exemplary $\{R_1, R_2, R_3\}$, which is a basic category that is less reliable, more harmoniousness, and has bad rigidity. Furthermore, set $\{x_1, x_3, x_7\} \cup \{x_2, x_4\} = \{x_1, x_2, x_3, x_4, x_7\}$ is $\{R_1\}$, which is a category that is less reliable or medium reliable (but not more reliable).

Going back to the problem definition in this specific problem, it is assumed that multiple users will evaluate the products with respect to one Kansei in "K" different levels. For example, if $U = \{x_1, x_2, x_3, x_4, \dots, x_{14}, x_{15}\}$ and these products are evaluated based on one user's Kansei, such as reliability at three different levels (less reliable, medium reliable, more reliable), then there is one equivalence relation with K = 3 equivalence classes, as follows:

U/Reliability = $\{\{x_1, x_3, x_7, x_9, x_{10}\}, \{x_2, x_4, x_{11}, x_{12}\}, \{x_5, x_6, x_8, x_{13}, x_{14}, x_{15}\}\}$

in which each set corresponds to each level of reliability.

In summary, the problem of this study has two main sets: a set of products, which is

$$X = \{x_i | i = 1, \dots, N\}$$

where

N = number of similar products that are to be evaluated by consumers

x_i = product i (alternative or candidate); i = 1…N

and a set of product features which is $C = \left\{c_j \middle| j = 1, \dots, M\right\}$ where

M = maximum number of product features corresponding to each product (condition attributes)

c_j = product feature j; j = 1… M

These two sets constitute a table, or information system. In addition, in many applications there is an outcome of classification that is known. This "a posteriori" knowledge is expressed usually by one distinguished attribute called the decision attribute. A decision system is any information system of the form $(U, A \cup \{d\})$, where $d \notin A$ is the decision attribute.

Any product feature of C has a domain of its design feature values Vc_j such that $c_j : X \to V_{c_j}$ for every $c_j \in C$. The decision attribute may take several values, although binary outcomes are rather frequent (Komorowski, Polkowski, & Skowron, 1998). Here, the decision made by each user for each product can be “accept, reject,” “good, bad,” “reliable, unreliable,” etc. However, instead of making a decision by using binary perception, users may rate an attribute on a semantic differential scale on which they are asked to choose where their position lies between two bipolar words, or a range of words or numbers across a bipolar position.

For algorithmic reasons, the knowledge of this specific problem can be presented in the form of an information system $S = \langle X, C, V, q \rangle$, where $X \subset U$ is a finite set of products, $C \subset Q$ is a finite set of product features, $V = \bigcup_{c_j \in C} V_{c_j}$, Vc_j is a domain of the product feature cj, and $q: X \times C \rightarrow V$ is a total function such that $q(x, c) \epsilon V c_j$ for every $c \in C$ and $x \in X$ (called information function). In another words, this information function is $q\ (x_i, c_j)$ = score of each product (x_i) with respect to each product’s feature (c_j) that the product development team gives $(i = 1 \dots N;\ j = 1 \dots M)$. This information is shown in Table 3.1.

TABLE 3.1 INFORMATION SYSTEM TABLE

Product	Product Feature				User Impression (Decision)
	c_1	c_2	……	c_M	d
x_1	$q(x_{1,}c_1)$	$q(x_1,c_2)$		$q(x_{1,}c_1)$	
x_2	$q(x_{2,}c_1)$				
x_i . . .					
x_N	$q(x_{N,}c_1)$			$q(x_N,c_M)$	

As mentioned previously, there could be many users for each product. Therefore, it would be very common for each person with his/her point of view to assign each product (xi) to any impression class d. Therefore, each row in Table 3.1 could be split into P (number of sample size) rows to reflect customers’ preferences, as shown in Table 3.2. If all users classify a particular product to the same class, there is no uncertainty. For example, as the first row of Table 3.3 shows, all users may assign product 1 to the same impression class (i.e., very good). On the other hand, users may have different perceptions for each product.

TABLE 3.2 INFORMATION SYSTEM TABLE WITH MULTIPLE USERS

Product	Product Feature	User	User Impression (Decision) d
x_1		User 1	
		User 2	
		User 3	
		User 4	
		…	
		User P	
x_2			
x_i . . .			
x_N			

Row 2 in Table 3.3 indicates an uncertain situation in which different users evaluate the same product differently.

TABLE 3.3 UNCERTAINTY IN MULTIPLE USERS

Product	Product Feature	User	User Impression (Decision) d
x1		User 1	very good
		User 2	very good
		User 3	very good
		User 4	very good
		…	
		User P	very good
x2		User 1	very good
		User 2	good
		User 3	bad
		User 4	not good–not bad
		…	
		User P	very bad
xi . . .			
xN		User 1	good
		User 2	bad
		User 3	good
		User 4	very bad
		…	
		User P	very good

In summary, if all users assign a particular product to a specific class, then there would be very clear boundaries between classes. In other words, in this case, there exists a fully correct and certain classification derived from a decision table. Otherwise, if some users assign the product to one class and some to another, one cannot say for certain to which class the product belongs, so the boundaries between classes would be unclear. In this case, when human evaluation data involves ambiguity, abundance of information, approximation, different beliefs, and conflicting evidence, it is very difficult to derive effective decision rules from such data.

Since a proposed approach based on rough set theory is used to resolve these kinds of uncertainties and inconsistencies, it is useful to mention that in rough set theory context, the original rough set approach is restricted to where there is no inconsistency between users' points of view. This is a situation in which there is no lower approximation (for example, a set of products that all users evaluate as reliable) of a classification, and the decision classes embody considerable ambiguity (Nishino et al., 2006).

Although in a heterogeneous market where designers use market segmentation techniques to classify their customers and optimize their design, inconsistency and uncertainty is inherent in this type of classification, so the objective is to minimize its negative effects. Nonetheless, confusion for designers, which creates uncertainty for them and conflict in user perceptions (inconsistency), can still exist. Part of this ambiguity and inconsistency can come from lack of harmony between multiple users' thoughts and/or feelings (which still exist even when they are grouped) or from inconsistent assignments and substations in market segmentation. In such a case, there would be no lower approximation of a classification or only very few elements in a lower approximation of some decision sets. In the former, it is very difficult to derive effective decision rules from such human evaluation data, while in the latter, the if-then rules extracted from these few elements might be unreliable. Therefore, resolving this kind of uncertainty and inconsistency is the subject of this study.

3.2 Approach

The approach used is a combination of traditional and rough set product development techniques. This procedure has two main stages. The objective of stage I is to identify the influential user characteristics on user perceptions and generate the rules that identify discernible classes of users. The objective of stage II is to determine the influential product features on perceptions of a selected discernible class of users along with product design rules. The approach is shown Figure 3.1.

3.2.1 Procedure

In this section, each step of the approach is explained in detail. In the next section, an example is used to show how the proposed procedure can be applied.

3.2.1.1 Stage I

3.2.1.1.1 Steps 1 to 3: Identify and Group Customers, and Define Targeted Customer Groups

Steps one to three have been explained in Chapter 2 (under the traditional product development methods section). During these steps, the customers of a specific product are defined and placed in groups based on particular factors such as age, education, etc. According to the company’s interest, one group can be targeted. Then an appropriate sample is taken from the target group.

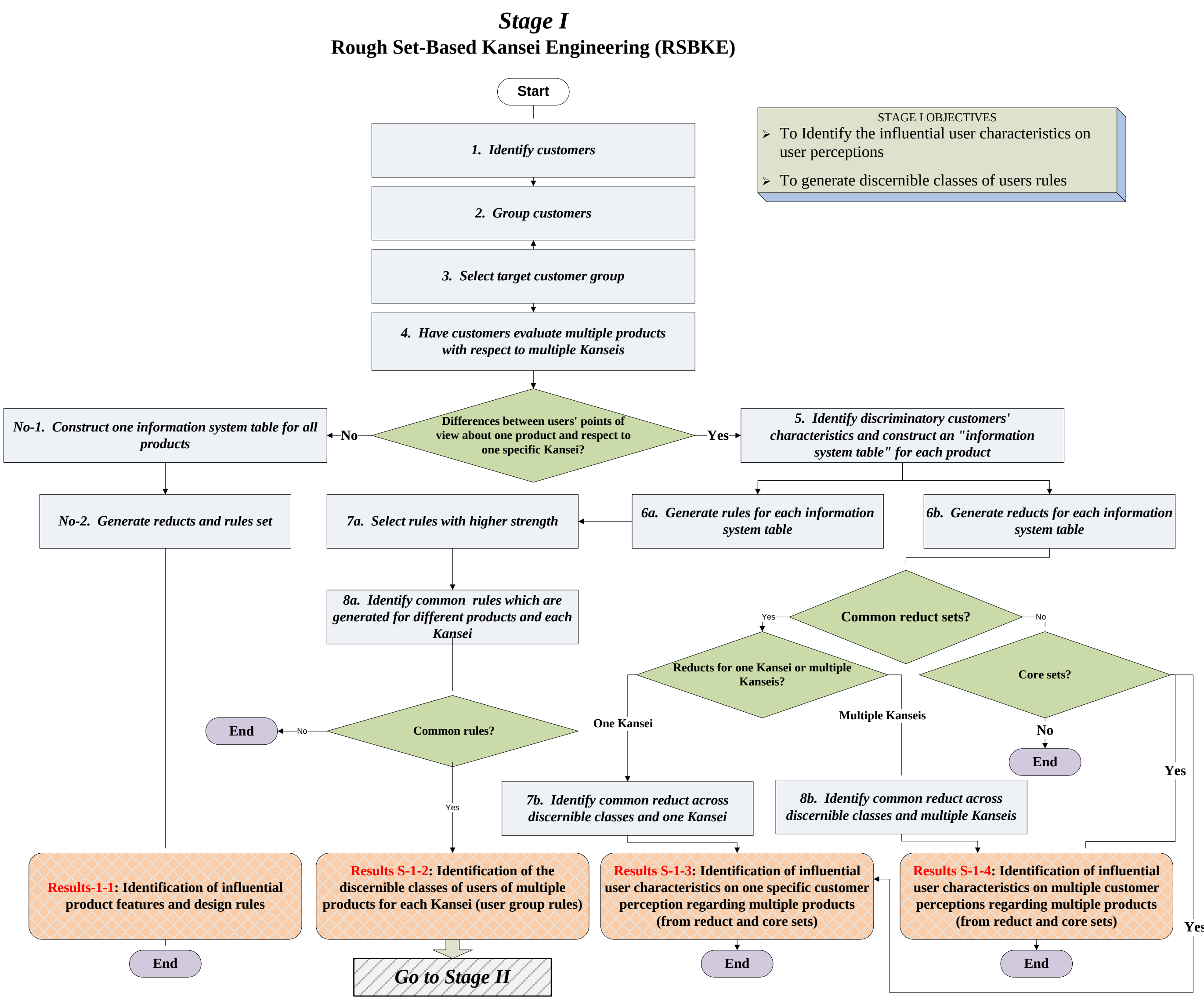


Figure 3.1. Two-stage approach.

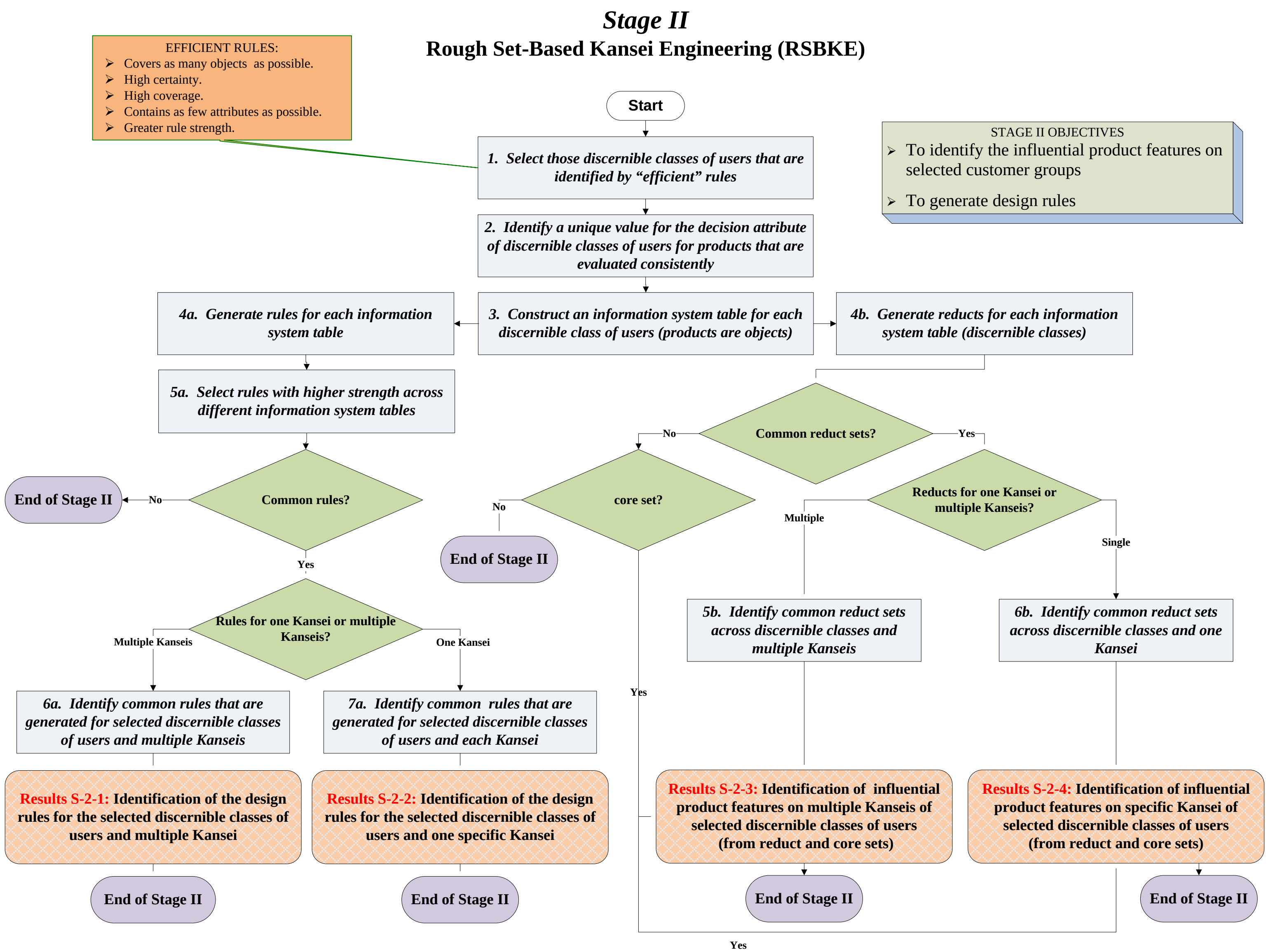


Figure 3.1 Two-stage approach (continued).

3.2.1.1.2 Step 4: Have Customers Evaluate Multiple Products

In this step, a group of products that are similar in function are shown to the users. Users evaluate them using a Likert or semantic differential scale with respect to some specific Kansei words such as harmoniousness (feeling that the components of a product are well matched or in harmony) or rigidity (feeling that a product looks stout, stable, and secure). Finally, the results are placed in the user impression class column in the information system table.

3.2.1.1.3 Step 5: Find Inconsistency and Look for Discriminatory User Characteristics

When users are asked to evaluate products, two sets of factors influence their decisions. The first set is based on what they see, sense, hear, and so on. These are the features of the product that influence user perception about that product. The purpose of this study is to find the most influential product features on user perception. The second set of factors comes from natural, intrinsic, and instinct user characteristics (who the users really are) and the consequences of the effects of their environment on them. These inherent and acquired characteristics of users, such as demographic, geographic, psychographic, and behavioralistic characteristics, could influence the users' decision-making processes.

If all users evaluate the selected products in the same way, there is no inconsistency among their perceptions and thus no need to categorize the customers. Then, designers can directly construct an information system table that includes all product evaluation results and can generate reducts and rules sets to identify product features (Steps N-1 and N-2). However, sometimes users with the same characteristics have different points of view with respect to a particular product. If there is such an inconsistency between multiple users' points of view, the reasons for such inconsistencies should be investigated. For example, users may not have been able to perceive a product appropriately, so maybe it is better to give them another chance to evaluate it. Nevertheless, they might actually belong to a different class of customers than was first assumed. In this case, some other distinguishable user characteristics should be searched until the users become discernible. If there are N different products which are evaluated by P customers, one separate information system can be constructed for each product. Therefore, each N information system table will have P rows representing the number of users who evaluate each product.

3.2.1.1.4 Step 6-A: Generate Rules

In this step, for each information system table constructed in the previous step, the corresponding rules are generated. The “if-then” rules determine the effect of specific values of user characteristics on a particular value of the user’s decision attribute, e.g., if the user is male, then the related product is evaluated as attractive.

3.2.1.1.5 Step 6-B: Generate Reduct Sets and Define Common Reducts

For each information system table, reduct sets are generated. Each reduct set represents influential user characteristics on user perceptions regarding the corresponding product. There might be one or multiple reduct sets for each product. Once all reducts are generated, the common reduct sets are identified for one specific Kansei or multiple Kanseis. If there is no common set among generated reduct sets, then taking the common set of core sets is suggested. Since the core set is usually smaller than the reduct sets, this will expand the boundary and decrease the accuracy of estimation. At this point, the first collection of results is identified as follows:

- Result S-1-1: Influential user characteristics on one specific customer perception regarding multiple products.
- Result S-1-2: Influential user characteristics on multiple customer perceptions regarding multiple products.

3.2.1.1.6 Steps 6-A, 7-A, and 8-A: Generate Rules and Identify Common Rules

In these steps, a set of rules is generated for each information system table, and those rules with higher strength are selected. The threshold of strength depends on the maximum strength, designer judgment, and number of rules that can be selected. For example, if the maximum rules strength is 0.2 and the threshold is tuned to 0.3, then no rules are selected. Each rule represents one discernible class of users for each product. By taking the common rules across products, the cluster of rules that identify the discernible class of users for a specific set of products is extracted. Each cluster has a unique value of decision for each product.

3.2.1.2 Stage II: Identify the Most Influential Product Features on User Decisions

3.2.1.2.1 Step 1: Select Discernible Classes Identified by Efficient Rules

This stage starts with taking the output of the previous stage, which is different clusters of rules, each identifying discernible class of users with consistent decisions regarding multiple products. Those clusters identified as “efficient rules” are selected. An efficient rule is defined as 0. Selecting discernible classes among many clusters and the number of selected clusters is the company’s decision.

3.2.1.2.2 Step 2: Identify the Decision Attribute Value of a Discernible Class of Users for Each Product

Each discernible class of users has a unique and consistent perception regarding specific products. In fact, products are categorized based on consistent user decisions. While a class of users may make a consistent decision about one specific product, no such consistent discernible class of users can be identified for another product.

3.2.1.2.3 Step 3: Construct an Information System for Each Class of Users

In this step, the information of the previous step is used to construct an information system table for each discernible class of users.

3.2.1.2.4 Steps 4-A and 5-A: Generate Rules and Select Strong Rules

Using the rough set method, design rules are generated for each information system table. In the next step, rules with a corresponding strength that is more than a specific threshold are selected. Again, the threshold of strength depends upon the maximum strength, designer judgment, and number of rules that can be selected.

3.2.1.2.5 Steps 6-A and 7-A: Define Common Rules for One or Multiple Kansei(s)

Common rules are identified from among the strong rules that were generated for different information system tables. If common rules are selected across a different discernible class for one specific Kansei, then the result is the design rules for the selected discernible classes of users and that specific Kansei. Otherwise, by selecting rules across different discernible classes and multiple Kanseis, the results are the design rules for the selected discernible classes of users and multiple Kanseis.

3.2.1.2.6 Steps 4-B, 5-B, and 6-B: Generate Reducts and Define Common Reducts for One or Multiple Kansei(s)

In this step, reduct sets are generated for each information system table, and those that are common are selected. The common reducts identify the influential product features on one or multiple Kansei(s), depending upon if they are selected across discernible classes for one or multiple Kansei(s).

The second collection results of stage II are as follows:

From the reduct sets:

- Stage II-1: Determine influential product features on multiple Kanseis of selected discernible classes of users.
- Stage II-2: Determine influential product features on one specific Kanseis of selected discernible classes of users.

From the rules:

- Stage II-3: Design rules for the selected discernible classes of users and one specific Kansei.
- Stage II-4: Design rules for the selected discernible classes of users and multiple Kanseis.

3.3 Demonstration

An example is given to show how the procedure can be applied. To simplify the problem, assume that the structure of the example problem is the same as the one mentioned in Chapter 2. The new definitions and changes are explained below.

3.3.1 Problem Statement

Suppose that a designer of a cell phone is interested in identifying cell phone attributes that impact user perceptions regarding reliability. Also, suppose that the set of product features that might influence user perception regarding the feeling of “reliable” are as follows:

c_1: Shape—5 (sharp rectangular), 4 (medium rectangular), 3 (curved rectangular)

c_2: Color—5 (red), 4 (blue), 3 (white)

c_3: Size—5 (large), 4 (medium), 3 (small)

c_4: Weight—5 (heavy), 4 (medium), 3 (light)

c_5: Balance—1 (low), 2 (medium), 3 (high)

c_6: Texture—1 (rough), 2 (medium), 3 (fine)

c_7: Translucency—1 (low), 2 (medium), 3 (high)

The product development team defines the technical value of each cell phone’s specification. For example, cell phone number one may have these features: c_1: (5) sharp rectangular, c_2: (5) red, c_3: (4) medium, c_4: (3) light, c_5: (3) high, c_6: (1) rough, and c_7: (1) low.

3.3.2 Procedure

3.3.2.1 Stage I: Identifying Influential User Characteristics and Discernible Classes of Users

3.3.2.1.1 Steps 1 to 4: Define Targeted Customer Group

Here it is assumed that steps 1 to 3 of the proposed procedure have been followed and that the results are available. In addition, users have been segmented according to some meaningful primary factors such as nationality, gender, and brand awareness. For example, users who are male, American, and familiar with cell phone brands are used in this study. In step 4, suppose that 15 brands of a specific type of cell phone are shown to 15 users who are asked to evaluate them with respect to a specific Kansei (in this case, the feeling “reliable”).

3.3.2.1.2 Step 5: Find Inconsistencies

Step 5 involves checking for differences between users’ different points of view regarding the same product. It is assumed that there are inconsistencies among users’ points of view. Therefore, there are 15 such tables, one for each product, which includes multiple user evaluations. An example of such evaluations is shown in Table 3.4, where inconsistencies are obvious; where the reliability of the product for users 1, 4, 5, 7, 8, 10, 11, 12, and 15 is acceptable; and where the reliability of the same product for users 2, 3, 6, 9, 13, and 14 is not acceptable. Also, users 4 and 10 are indiscernible. However, users 8 and 9 (indicated by ○) are indiscernible in terms of their characteristics, but they evaluate the product differently.

Looking for discriminatory user characteristics, it is assumed that seven user characteristics (as secondary characteristics) impact their decisions, as shown in Table 3.4. These characteristics can be selected from among many characteristics that are believed to have an impact on user feelings. In such a case, different methods, such as multiple regression analysis, can be used to reduce the number of characteristics to a manageable number. Each user is instructed to accept or reject these 15 products in terms of reliability. User characteristics are identified below and shown in Table 3.4:

Demographic

c_1: Age—5 (teens), 4 (twenties), 3 (thirties)

c_2: Income—5 (over 50K), 4 (20K–50K), 3 (less than 20K)

c_3: Education—5 (high school), 4 (undergraduate), 3 (graduate)

Geographic

c_4: Climate—5 (hot and humid), 4 (hot and dry), 3 (mild)

c_5: Living Place—1 (urban), 2 (suburban), 3 (rural)

Behavioralistic

c_6: Usage Rate—1 (low), 2 (moderate), 3 (high)

c_7: User Status—1 (potential), 2 (first time), 3 (regular)

TABLE 3.4 INFORMATION SYSTEM FOR MULTIPLE USERS FOR ONE SPECIFIC PRODUCT

Users	Criteria							Decision
	c_1 (Age)	c_2 (Income)	c_3 (Education)	c_4 (Climate)	c_5 (Living Place)	c_6 (Usage Rate)	c_7 (User Status)	
x_1	4	4	4	4	2	2	1	A
●x_4	5	3	5	4	2	1	2	A
x_5	4	4	5	4	2	2	1	A
x_7	4	4	5	4	2	2	2	A
○x_8	4	4	4	4	2	2	2	A
●x_{10}	5	3	5	4	2	1	2	A
x_{11}	5	4	4	4	1	1	2	A
x_{12}	5	3	4	4	2	2	2	A
x_{15}	4	5	5	4	2	1	1	A
x_2	3	3	4	3	2	1	1	R
x_3	3	4	3	3	1	2	2	R
x_6	3	4	3	3	2	1	3	R
○x_9	4	4	4	4	2	2	2	R
x_{13}	4	3	3	3	3	2	2	R
x_{14}	3	3	4	3	2	3	3	R

Based on the above discussion, the condition attributes are $C = \{c_1, c_2, c_3, c_4, c_5, c_6, c_7\}$ = {age, income, education, climate, living space, usage rate, user status}, and the decision attribute is D = {d} = {Reliable}.

3.3.2.1.3 Steps 6, 7, and 8: Generate Reducts and Rules, and Identify Common Reducts and Rules

Reduct sets and rules are generated based on the rough set algorithm, as explained in Chapter 2. A brief explanation is provided here.

3.3.2.1.3.1 Define Partitions

In this problem, there are two partitions, Y_A, and Y_R $(Y = \{Y_A, Y_R\})$. Y_A is the set of users that evaluates a specific cell phone as reliable, and Y_R is the set of users that evaluates the cell phone as not reliable. This problem may have an additional set, medium reliability, which is not considered here. As shown in Table 3.4:

$$Y_A = \{x_1, x_4, x_5, x_7, x_8, x_{10}, x_{11}, x_{12}, x_{15}\} \text{ and } Y_R = \{x_2, x_3, x_6, x_9, x_{13}, x_{14}\}$$

3.3.2.1.3.2 Define Indiscernible Sets

In this problem, there are 13 C-elementary indiscernible sets, where C refers to the condition attributes. These indiscernible sets include $\{x_4, x_{10}\}$,$\{x_8, x_9\}$, and 11 other sets, each including one user, such as $\{x_1\}$, $\{x_5\}$, etc. It should be noted that in $\{x_4, x_{10}\}$,$\{x_8, x_9\}$, both users x_8 and x_9 have the same conditions, which are shown by the same signs in Table 3.4.

3.3.2.1.3.3 Identify Lower and Upper Approximation Sets and Doubtful Region, and Calculate Accuracy and Quality of Approximation of Set *Y*

The objects (users) that certainly belong to set YA should be identified. This user set shows the lower limit of YA, which is $\underline{C}\mathrm{Y_A} = \{x_1, x_4, x_5, x_7, x_{10}, x_{11}, x_{12}, x_{15}\}$.

In addition, users that possibly belong to set Y_A should be identified. This user set represents the upper limit of Y_A, which is $\overline{C} = \{x_1, x_4, x_5, x_7, x_8, x_9, x_{10}, x_{11}, x_{12}, x_{15}\}$. x_8 and x_9

are the users that have the same conditions (characteristics), but one of them accepts the product as a reliable product and one of them rejects it. The difference between upper approximation and lower approximation constitutes the C-boundary (doubtful region) of set Y_A, which is B_{n_c} (Y_A) $=$ $\{x_8, x_9\}$. In fact, x_8 and x_9 are the users that cannot be classified to Y with certainty using the set of user characteristics (C). The same thing can be done for the "not reliable" set Y_R. Results are shown below:

$\underline{C}Y_R = \{x_2, x_3, x_6, x_{13}, x_{14}\}$

$\overline{C}\ Y_R = \{x_2, x_3, x_6, x_8, x_9, x_{13}, x_{14}\}$

B_{n_c} (Y_R) = $\overline{C}\ Y_R$ - $\underline{C}Y_R$=$\{x_8, x_9\}$

Based on formulas that were mentioned in Chapter 2, section2.1.2.2.3, equation (2.4), the accuracy of approximation of sets Y_A and Y_R by C is equal to 8 (cardinality of lower approximation) divided by 10 (cardinality of upper approximation), which is 0.8 and (5/7) = 0.71, respectively. In addition, the quality of approximation (quality of sorting, section 2.1.2.2.3, equation (2.5)) of the decision by C is equal to ((8+5)/15) = 0.87.

Users x_8 and x_9 are inconsistent when making decisions about cell phones with respect to criteria or condition C. As mentioned previously, if there is such an inconsistency, the reasons should be investigated, for example, giving them another chance to evaluate the product. Nevertheless, each might belong to a different class of customers. Therefore, in this case, the designers may want the users to evaluate the product again (by creating a third category of users, those who should be asked again). Or the inconsistency might be some other user characteristics that have influenced user decisions but have not been considered in the condition set. In other words, other distinguishable user characteristics should be searched until the users become discernible and the uncertainty is cleared. This may provide some hints to designers for finding those characteristics that really affect a user's decision process.

3.3.2.1.3.4 Construct Minimal Subsets of Independent Criteria (Reducts Subsets)

The redundant criteria from the whole set of criteria can be eliminated such that the decision-making results will not be affected. Minimal subsets of independent criteria are obtained in such a way that they will have the same quality of sorting as the entire set of condition attributes (C).

In this problem, there are three such reduct sets:

$$RED_{\Upsilon}^{1}(C) = \{c_2, c_3, c_6, c_7\}$$

$$\mathrm{RED}_{\Upsilon}^{2}(\mathrm{C}) = \{\mathrm{c}_1, \mathrm{c}_3, \mathrm{c}_7\}$$

$$\mathrm{RED}_{\Upsilon}^{3}(\mathrm{C}) = \{\mathrm{c}_2, \mathrm{c}_3, \mathrm{c}_5, \mathrm{c}_7\}$$

Any of the above sets can influence the user's decision process. If set one is considered a reduced set of criteria, which are {c_2 = income, c_3 = education, c_6 = usage rate, c_7 = user status}, then these user characteristics play a major role in the decision process of perceiving the product as reliable or not. In addition, the core set of all reduct sets can be obtained by finding the intersection of all reduct sets:

$$CORE\Upsilon\,(C) = RED_{Y}^{1}(C) \bigcap RED_{Y}^{2}(C) \bigcap RED_{Y}^{3}(C) = \{c_3, c_7\}$$

It can be said that when users decide about whether or not a product is reliable, certain characteristics more than others influence their decision. It is interesting that criterion $c_{4,}$ which is "climate," has no influence at all on the decision because it is not represented in any reduct set. Also the core, which is {c_3 = education, c_7 = user status}, is the most essential part of set C and cannot be eliminated without disturbing the ability of approximating the decision.

Now, a group of customers should be selected. If the second reduct set ($RED_{Y}^{2}(\mathrm{C})$) is considered as the selected reduct set in this problem, then the decision table is reduced to criteria c_1, c_3, and c_7, according to the approach by Pawlak and Slowinski (1994). Then the following decision rules are generated from the reduced decision table. Table 3.5shows these decision rules. Five rules are crisp and one is fuzzy (rule no. 4), because x_8 and x_9 belong to a boundary region (doubtful region).

TABLE 3.5 DECISION RULES FOR USER CHARACTERISTICS THAT AFFECTDECISION OF MULTIPLE USERS

Rule No.	If	c_1 (Age)	c_3 (Education)	c_7 (User Status)	Decision
1	If	5			Accept
2	If		5		Accept
3	If	4		1	Accept
4	If	4	4	2	Accept or Reject
5	If	3			Reject
6	If		3		Reject

Now, based on the above decision rules, the policy of customer categorization could be expressed as follows:

> "All users who are teens or have a high school level of education see the cell phone as reliable. Those potential users who are in their twenties consider the cell phone as reliable. First-time undergraduate users in their twenties should be surveyed again, or other characteristics should be investigated so that their class might be divided into several discernible classes depending upon new criteria. Users in their thirties or with a graduate level of education see the cell phone as not reliable."

If the same analysis using the same users is implemented for the rest of the products $\{X_2, \ldots, X_{15}\}$, then Table 3.6 can be constructed.

TABLE 3.6 REDUCT SETS AND CORE SET OF EACH PRODUCT EVALUATION

Product	Reduct Sets	Core Sets
x_1	$\{c_2, c_3, c_6, c_7\};\{c_1, c_3, c_7\},\{c_2, c_3, c_5, c_7\}$	$\{c_3, c_7\}$
x_2	$\{c_2, c_3\};\{c_1, c_3, c_7\},\{c_2, c_3, c_5, c_7\}$	$\{c_3\}$
x_3	$\{c_2, c_3, c_6, c_7\};\{c_1, c_3, c_7\},\{c_5, c_7\}$	$\{c_7\}$
x_4	$\{c_2, c_6, c_7\};\{c_1, c_3, c_7\},\{c_2, c_5, c_7\}$	$\{c_7\}$
x_5	$\{c_2, c_3, c_6, c_7\};\{c_1, c_3, c_7\}$	$\{c_3, c_7\}$
x_6		
……..		
x_{13}		
x_{14}	$\{c_1, c_3, c_7\};\{c_2, c_3, c_5, c_7\}$	$\{c_3, c_7\}$
X_{15}	$\{c_1, c_2, c_3, c_6, c_7\};\{c_1, c_3, c_7\}\{c_2, c_3, c_5, c_7\}$	$\{c_3, c_7\}$

As shown in Table 3.6, one common set of the reduct set, which is $\{c_1, c_3, c_7\}$, can be found. This means that the entire knowledge about users (all user characteristics) is not necessary to define categories in the knowledge available. Instead, a particular set of user characteristics affect their perception about the reliability of all considered products.

In addition, the common rules across different information systems can be identified to describe the discernible classes of users. If there would be no common rules, then those rules that are common across part of a product set can be identified. The detail of this step is demonstrated using real data in Chapters 4 and 5.

3.3.2.2 Stage II: Identifying Most Influential Product Features on User Decision

3.3.2.2.1 Step 1: Select Discernible Classes Identified by Efficient Rules

Now the designers can choose one specific group among users who have the same value of c_1, c_3, and c_7 characteristics. Some of these groups could be as follows:

- Teenagers with a high school level of education and who are regular users.
- People in their twenties with an undergraduate education and who are potential users.
- Users in their thirties with a graduate level of education and who are first-time users.

Many other groups could be defined using three user characteristics: age, education, and user status, depending upon the situation. Suppose a company chooses high school teenage students who have regular usage and are male, American, and aware of the cell phone brands in this study, since they are consistent and are the most profitable group. As mentioned in the if-then rules, all users who belong to this group view cell phone number one as a reliable cell phone. However, this group may view product number two as an unreliable cell phone, etc.

3.3.2.2.2 Steps 2 and 3: Identify Decision Attribute Value of Discernible Class of Users for Each Product, and Construct Information System for Each Class of Users

Assume that the aforementioned group rated all 15 products as shown in Table 3.7. The product features are as follows:

c_1: Shape—5 (sharp rectangular), 4 (medium rectangular), 3 (curved rectangular))

c_2: Color—5 (red), 4 (blue), 3 (white)

c_3: Size—5 (large), 4 (medium), 3 (small)

c_4: Weight—5 (heavy), 4 (medium), 3 (light)

c_5: Balance—1 (low), 2 (medium), 3 (high)

c_6: Texture—1 (rough), 2 (medium), 3 (fine)

c_7: Translucency—1 (low), 2 (medium), 3 (high)

TABLE 3.7 INFORMATION SYSTEM: "HIGH SCHOOL TEENAGER STUDENTS WITH REGULAR USAGE" GROUP

Product	Criteria							Decision
	c1 (Shape)	c2 (Color)	c3 (Size)	c4 (Weight)	c5 (Balance)	c6 (Texture)	c7 (Translucency)	
x1	4	4	4	4	2	2	1	A
x4	5	3	5	4	2	1	2	A
x5	4	4	5	4	2	2	1	A
x7	4	4	5	4	2	2	2	A
x8	4	4	4	4	2	2	2	A
x10	5	3	5	4	2	1	2	A
x11	5	4	4	4	1	1	2	A
x12	5	3	4	4	2	2	2	A
x15	4	5	5	4	2	1	1	A
x2	3	3	4	3	2	1	1	R
x3	3	4	3	3	1	2	2	R
x6	3	4	3	3	2	1	3	R
x9	4	4	4	4	2	2	2	R
x13	4	3	3	3	3	2	2	R
x14	3	3	4	3	2	3	3	R

Table 3.7 refers to the customer-product information system and shows the decision attribute value of a discernible class of users for each product. In addition, Table 3.7 implies that the inconsistency that existed between users has been resolved, since all users of the selected group are consistent for each product. The number of customer-product information systems will vary, depending on how many classes of users are selected. In this example, one group is selected. Using the rough set theory approach, the information in Table 3.7 is used to identify the most influential product features.

3.3.2.2.3 Steps 4, 5, 6, and 7: Generate Rules and Define Common Rules for One or Multiple Kansei(s)

At this point, the last process of the main procedure, finding the most influential product features results in the following reduct sets, decision rules, and policy. Since here one group of users is selected, the common rules across different groups cannot be shown. However, the detail of this step regarding common rules is illustrated in Chapters 4 and 5.

In this problem, there are three reducts:

$RED^1_{\Upsilon}(C) = \{c_2, c_3, c_6, c_7\}$

$RED^2_{\Upsilon}(C) = \{c_1, c_3, c_7\}$

$RED^3_{\Upsilon}(C) = \{c_2, c_3, c_5, c_7\}$

This means that if set one is considered a reduced set of criteria {c2, c3, c6, c7}, then it can be said that the selected users take into account only these criteria to make a decision about cell phones regarding reliability perception. In addition, the core set of all reduct sets can be obtained by finding the intersection of all reduct sets:

$$\mathrm{CORE}_{\Upsilon}(C) = RED^1_{\Upsilon}(C) \cap RED^2_{\Upsilon}(C) \cap RED^3_{\Upsilon}(C) = \{c_3, c_7\}$$

If the second reduct set (RED_{Υ}^{2}(C)) is considered a reduct set in this problem, then the decision table can be reduced to criteria c_1, c_3, and c_7. Then the following decision rules are generated from the reduced decision table, which is shown in Table 3.8.

Five rules are crisp and one is fuzzy (rule no. 4), because x_8 and x_9 belong to the boundary region (doubtful region). Now, based on the above decision rules, the policy can be expressed as follows:

The user considers having a score of “5” in shape or size as more reliable all cell phones, which means that the cell phones whose shapes are sharp rectangular and whose size are big, look more reliable to users. Reliability is also perceived in phones that have score of “4” in shape (medium rectangular) and in size (medium in size) but high in translucency. In the case of a score of “4” in shape and in size but only medium in translucency, phones are evaluation again. Cell phones having a score of “3” in shape (curved rectangular) or in size (small) are considered to be the least reliable.

TABLE 3.8 DECISION RULES FOR PRODUCT FEATURES THAT AFFECT DECISION GROUP

Rule No.	If	c_1 (Shape)	c_3 (Size)	c_7 (Translucency)	Decision
1	If	5			Reliable
2	If		5		Reliable
3	If	4		1	Reliable
4	If	4	4	2	Reliable or Unreliable
5	If	3			Unreliable
6	If		3		Unreliable

Therefore, it was discovered that size, shape, and translucency impact the perception of the “high school teenager students with regular usage” group of users in regards to reliability.

3.4 Computation

Since the size of the above problem is too large to be solved manually in this report, its dimension is reduced to show the computational aspect. Furthermore, sometimes objects may have the same features. For example, in the above problem, some of the users may have the same perception with respect to the same product. In such a case, the size of the data can be reduced by storing only one representative object. Therefore, in this problem, to reduce the dimension of problem, it is assumed that there are eight discernible users along with four condition attributes (the original problem had 15 users with seven condition attributes). As Table 3.9 shows, eight users are asked to evaluate one specific product, and it is assumed that four characteristics impact their decision.

TABLE 3.9 INFORMATION SYSTEM FOR MULTIPLE USERS FOR ONE SPECIFIC PRODUCT

User	Criteria				Decision
	c_1 (Age)	c_2 (Education)	c_3 (Climate)	c_4 (Usage Rate)	
x_1	4	4	4	3	A
x_4	5	3	5	3	A
x_5	4	4	5	1	A
x_7	4	4	5	2	A
x_8	4	3	4	2	A
x_2	3	3	4	3	R
x_3	3	4	3	3	R
x_6	3	4	3	1	R

The user characteristics that are considered are listed below:

Demographic

c_1: Age—5 (teens), 4 (twenties), 3 (thirties)

c_2: Education—5 (high school), 4 (undergraduate), 3 (graduate)

Geographic

c_3: Climate—5 (hot and humid), 4 (hot and dry), 3 (mild)

Behavioralistic

c_4: Usage Rate—1 (low), 2 (moderate), 3 (high)

$$X = \{x1, x2\ , x3\ , x4\ , x5\ , x6\ , x7\ , x8\ , x9\}$$

C = {c_1, c_2, c_3, c_4} = {age, education, climate, usage rate}

V_{Age} = {5 (teens), 4 (twenties), 3 (thirties)}

$V_{Education}$ = {5 (high school), 4 (undergraduate), 3 (graduate)}

$V_{Climate}$ = {5 (hot and humid), 4 (hot and dry), 3 (mild)}

$V_{Usage\ Rate}$ = {1 (low), 2 (moderate), 3 (high)}

V_d = {Reject, Accept}

Given information system $A = (U, C)$, its reduct is a minimal set of attributes $B \subseteq C$, such that $INDA(B) = INDA(C)$. In this problem, to calculate the reduct sets, first a symmetric 8×8 matrix with the empty diagonal, called the indiscernibility matrix, with c_{ij} entries as given below, should be obtained:

$$c_{ij} = \{a \in C \big| a(x_i) \neq a(x_j)\} \tag{3.1}$$

for $i, j = 1,2, \ldots, 8$

Therefore, each entry consists of a set of attributes upon which x_i and x_j differ. Table 3.10 shows the indiscernibility matrix for Table 3.9.

TABLE 3.10 INDISCERNIBILITY MATRIX FOR TABLE 3.9

User	x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8
x_1	-	c_1, c_2	c_1, c_3	c_1,c_2, c_3	c_3, c_4	c_1,c_3,c_4	c_3, c_4	c_2,c_4
x_2		-	c_2, c_3	c_1, c_3	c_1,c_2,c_3,c_4	c_2,c_3,c_4	c_1,c_2,c_3,c_4	c_1,c_4
x_3			-	c_1,c_2, c_3	c_1, c_3,c_4	c_4	c_1,c_3,c_4	c_1,c_2,c_3,c_4
x_4				-	c_1,c_2, c_4	c_1,c_2,c_3,c_4	c_1,c_2, c_4	c_1, c_3,c_4
x_5					-	c_1, c_3	c_4	c_2,c_3,c_4
x_6						-	c_1,c_3,c_4	c_1,c_2,c_3,c_4
x_7							-	c_2, c_3
x_8								-

An indiscernibility function f_A for information system A is a Boolean of m Boolean variables a*1… a*m (corresponding to the attributes a1… am) defined as follows:

$$(a^*1, \ldots, a^*m) = \{\vee c^*_{ij} \big| 1 \leq j \leq i \leq n, c_{ij} \neq \emptyset\} \tag{3.2}$$

where $c^*_{ij} = \{a^* \big| a \in c_{ij}\}$, and n is the number of objects. In this problem, the indiscernibility function is as follows:

$f_A\ (c_1, c_2, c_3)$

$= (c_1 \vee c_2) \wedge (c_1 \vee c_3) \wedge (c_1 \vee c_2 \vee c_3) \wedge (c_3 \vee c_4) \wedge (c_1 \vee c_3 \vee c_4) \wedge (c_3 \vee c_4) \wedge (c_2 \vee c_4)$

$\wedge (c_2 \vee c_3) \wedge (c_1 \vee c_3) \wedge (c_1 \vee c_2 \vee c_3 \vee c_4) \wedge (c_2 \vee c_3 \vee c_4) \wedge (c_1 \vee c_2 \vee c_3 \vee c_4) \wedge (c_1$

$\vee c_4) \wedge (c_1 \vee c_2 \vee c_3) \wedge (c_1 \vee c_3 \vee c_4) \wedge (c_4) \wedge (c_1 \vee c_3 \vee c_4) \wedge (c_1 \vee c_2 \vee c_3 \vee c_4) \wedge (c_1$

$\vee c_2 \vee c_4) \wedge (c_1 \vee c_2 \vee c_3 \vee c_4) \wedge (c_1 \vee c_2 \vee c_4) \wedge (c_1 \vee c_3 \vee c_4) \wedge (c_1 \vee c_3) \wedge (c_4) \wedge (c_2$

$\vee c_3 \vee c_4) \wedge (c_1 \vee c_3 \vee c_4) \wedge (c_1 \vee c_2 \vee c_3 \vee c_4) \wedge (c_2 \vee c_3)$

It should be noted that each row in the above indiscernibility function corresponds to one row in the indiscernibility matrix. For example, the fifth row of the above equation says that the fifth user (x5) (or more precisely, the fifth equivalence class) may be discerned from the sixth one by any of the attributes of age or climate ($c_1 \vee c_3$), from the seventh one by usage rate (c_4), and from the eighth one by any attributes of education, climate, or usage rate ($c_2 \vee c_3 \vee c_4$).

To have a better understanding of map, minterms, and the relationship of the four variables, the minterm assignment in each square and the relationship of the four variables are indicated in Table 3.11 and Table 3.12, respectively. The set of all prime implicants of f_A determines the set of all reducts of A[1].

TABLE 3.11 MINTERM ASSIGNMENT

m_0	m_1	m_3	m_2
m_4	m_5	m_7	m_6
m_{12}	m_{13}	m_{15}	m_{14}
m_8	m_9	m_{11}	m_{10}

[1]"An implicant of a Boolean function f is any conjunction of literals (variables or their negations) such that if the values of these literals are true under an arbitrary valuation υ of variables then the value of the function f under υ is also true. A prime implicant is a minimal implicant. Here we are interested in implicants of monotone Boolean functions only, i.e., functions constructed without negation" (Komorowski, Polkowski, & Skowron, 1998).

TABLE 3.12 RELATIONSHIP OF FOUR VARIABLES

		c_3c_4					
					c_3		
c_1c_2			00	01	11	10	
		00					
		01					c_2
	c_1	11					
		10					
				c_4			

So to simplify the above equation, first, a four-variables map (Mano & Kim, 2004), which has 16 minterms[2] and 16 squares for four binary variables, is defined as shown in Table 3.13. Here, the rows and columns are numbered so that only one bit of the binary number changes in value between any two adjacent columns or rows. Using Matlab software, each square, which represents the output of the indiscernibility function, f_A, is calculated. As shown, the minterms of the function are marked with "1" in the map of Table 3.13. Using the map manipulation method (Mano & Kim, 2004), the optimized expression can be found. To do this, all prime implicants should be determined. Then the optimized expression is obtained from the logical sum of all the essential prime implicants, plus other prime implicants needed to contain remaining minterms not included in the essential prim implicant. The output optimized function indicates the reduct sets.

TABLE 3.13 FOUR-VARIABLE MAP OF PROBLEM

		c_3c_4			
		00	01	11	10
c_1c_2	00	0	0	0	0
	01	0	0	1	0
	11	0	1	1	0
	10	0	0	1	0

[2]A minterm is the product of N distinct literals, where each literal occurs exactly once. A literal is a single variable within a term which may or may not be complemented (URL: http://www.asic-world.com/digital/boolean2.html).

$m_7 + m_{15} = \overline{c_1} c_2c_3c_4 + c1c_2c_3c_4 = c_2c_3c_4(\overline{c_1} + c_1) = c_2c_3c_4$
$m_{15} + m_{13} = c_1c_2c_3c_4 + c_1c_2\overline{c_3}c_4 = c_1c_2\ c_4\ (c_3 + \overline{c_3}) = c_1c_2c_4$
$m_{11} + m_{15} = c_1\overline{c_2}c_3c_4 + c_1c_2c_3c_4 = c_1c_3c_4(c_2 + \overline{c_2}) = c_1c_3c_4$

Also, the optimized function is obtained as follows:

$f\ (c_1, c_2, c_3, c_4) = c_2c_3c_4 + c_1c_2\ c_4 + c_1c_3c_4$

Therefore, three reducts are obtained:

$\{c_2, c_3, c_4\}, \{c_1, c_2, c_4\}, \{c_1, c_3, c_4\}$

To extract the decision rules, according to the approach by Komorowski et al. (1998), another matrix, called the decision-relative indiscernibility matrix, should be obtained. This matrix shows how an object belonging to the corresponding decision class is discerned from other objects belonging to the other decision classes.

To construct this matrix, let $A = (U, C \cup \{d\})$ be an information system, and let $M\ (A) = (c_{ij})$ be its indiscernibility matrix. A decision-relative indiscernibility matrix of A, $Md(A) = (cd_{ij})$ is constructed assuming $cij = 0$ if $d(x_i) = d(x_j)$; otherwise, $cd_{ij} = c_{ij}$. The procedure to construct this matrix and its indiscernibility function is the same as the procedure and indiscernibility function of the indiscernibility matrix. Table 3.14 shows the decision-relative indiscernibility matrix of the problem. Again, this matrix is symmetrical and the diagonal is empty.

TABLE 3.14 DECISION-RELATIVE INDISCERNIBILITY MATRIX FOR TABLE 3.9

Users	x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8
x_1	-	c_1, c_2	c_1, c_3			c_1, c_3, c_4		
x_2		-		c_1, c_3	c_1, c_2, c_3, c_4		c_1, c_2, c_3, c_4	c_1, c_4
x_3			-	c_1, c_2, c_3	c_1, c_3, c_4		c_1,c_3,c_4	c_1, c_2, c_3, c_4
x_4				-		c_1, c_2, c_3, c_4		
x_5					-	c_1, c_3		
x_6						-	c_1,c_3,c_4	c_1, c_2, c_3, c_4
x_7							-	
x_8								-

The decision-relative indiscernibility function of the above matrix is as follows:

$f_M^d(c_1, c_2, c_3, c_4) = (c_1 \vee c_2) \wedge (c_1 \vee c_3) \wedge (c_1 \vee c_3 \vee c_4)$

$\wedge (c_1 \vee c_3) \wedge (c_1 \vee c_2 \vee c_3 \vee c_4) \wedge (c_1 \vee c_2 \vee c_3 \vee c_4) \wedge (c_1 \vee c_4)$

$\wedge (c_1 \vee c_2 \vee c_3) \wedge (c_1 \vee c_3 \vee c_4) \wedge (c_1 \vee c_3 \vee c_4) \wedge (c_1 \vee c_2 \vee c_3 \vee c_4)$

$\wedge (c_1 \vee c_2 \vee c_3 \vee c_4)$

$\wedge (c_1 \vee c_3)$

$\wedge (c_1 \vee c_3 \vee c_4) \wedge (c_1 \vee c_2 \vee c_3 \vee c_4)$

Using Matlab software and the map manipulation method, the map of the above equation is shown in Table 3.15. Therefore, the prime implicants and the optimized function are:

$m_{12} + m_{13} + m_{15} + m_{14} + m_8 + m_9 + m_{11} + m_{10} = c_1$

$m_{15} + m_{11} = c_2c_3c_4$

$$f\ (c_1, c_2, c_3, c_4) = c_2c_3c_4 + c_1$$

TABLE 3.15 MAP OF DECISION-RELATIVE INDISCERNIBILITY FUNCTION

		c_3c_4			
		00	01	11	10
c_1c_2	00	0	0	0	0
	01	0	0	1	0
	11	1	1	1	1
	10	1	1	1	1

Therefore, the following two reducts are obtained: {c2, c3, c4}, {c1}

If the reduct set $2(RED_A^2(C))$ is considered as a reduct set in this problem, then the decision table can be reduced to criteria c_1. Then the following decision rules are generated from the reduced decision table, as shown in Table 3.16.

TABLE 3.16 DECISION RULES GENERATED FROM REDUCED DECISION TABLE FOR SPECIFIC USERS GROUP

Rule No.	If	c_1	Decision
1	If	4 or 5	Accept
2	If	3	Reject

3.5 Conclusions

In brief, customer inconsistency caused by having multiple users for a product usually is a problem in customer-oriented product development approaches. The proposed approach to resolving this inconsistency and uncertainty has four major stages. First, the market is segmented initially. Second, the significant user characteristics and their product assessment with respect to either functional or nonfunctional (psychosocial) requirements are defined. Third, customers are categorized based on the similarity of their perception using reduct sets and the rule generation concept in rough set theory. Fourth, for each group of customers, the most influential product features on customers' feelings are defined. An example was given to demonstrate how the proposed approach works. In Chapter 4, website design data are obtained to demonstrate the proposed approach. To do this, the process of extracting the reduct sets, rules, and corresponding rough set measures using the rough set exploration system program are demonstrated. Rough set notions and equations required to apply the proposed approach are introduced. Also, to apply the proposed approach efficiently, software to integrate the reducts and rules is required. The output of this software is also illustrated in Chapter 4.

CHAPTER 4

IMPLEMENTATION

In this chapter, the application of the proposed approach is demonstrated through a website design project. The purpose here is to generate two sets of decision rules. The main objective of the problem under consideration is to identify clusters of consistent users and to identify the influential website features that can help designers customize the website design for each cluster of users.

The data used in this study is adopted from a study by Phillips (2007). The study was performed in the usability laboratory of the Department of Psychology at Wichita State University. To identify reduct sets and generate decision rules, the Rough Set Exploration System (RSES), version 2.2, was used. In addition, software was programmed by visual basic language to bundle common reducts and rules. This software is referred to as the "reduct-rule integrator" (RRI).

This chapter includes four major sections. First the nature of the data is explained. Then the process of preparing data to be used in the developed approach is described. Following that, the website design case is explained using rough set theory notation. Finally, the progression of data processing using RSES and RRI, along with the view of the output, is illustrated.

4.1 Data

4.1.1 Background Data (User Characteristics)

Sixty-three participants were selected to evaluate 24 different websites (previously identified as 12 having high appeal and 12 having low appeal) with respect to eight adjectives (Kanseis). Among those, three of them (interesting, imaginative, and attractive) were selected for consideration in this study. It was assumed that these websites are different realizations of one specific website, which is assumed to be a communication interface between the company and its users. For each participant, 55 characteristics of the users were recorded. These characteristics included the following: age; gender; scores from the Internet Experience Scale (IES), which measures the user's Internet experience level and includes the four main subscales of hedonic, utility, involvement, and skill; and participant response to visual appeal, or centrality of visual product aesthetics (CVPA) ((Phillips, 2007). In this study, it was assumed that 12 of those user characteristics had a greater relationship with the three user perceptions (Kanseis). These characteristics may come from a previous relationship analysis, such as regression analysis. These characteristics along with their measurement scales are listed below.

Demographic Characteristics

- Student/Community: Defines whether a participant is a student or a member of the community (student = 1, community = 2).
- Gender: Defines the gender of the participant (male = 1, female = 2).
- Age: Indicates the age of the participants (below 25 = 1, 25–40 = 2, 40–50 = 3, 50–60 = 4, over 60 = 5).

Self-Ratings Characteristics

- Frequency of Use: Defines how often the participant uses a computer (never = 1, 1 to 10 times = 2, monthly = 3, weekly = 4, daily = 5, hourly = 6).
- Participant's Entertainment: Defines how often a participant uses the Internet for entertainment purposes (low = 1, high = 2).
- Participant's Work: Defines how often a participant uses the Internet for work purposes (low = 1, high = 2).

- Participant's Communication: Defines how often a participant uses the Internet for communication purposes (low = 1, high = 2).
- Skill: Deals with the users' ability to problem-solve (continuous scale of 1–6, where high = 1, low = 6).
- Utility: Examines the Internet as a way to get information and finish the task (continuous scale of 1–6, where high = 1, low = 6).
- Hedonic: Focuses on activities that are more exploratory and entertainment-derived in nature (continuous scale of 1–6, where high = 1, low = 6).
- Involvement: Denotes the level that the users are being immersed while completing tasks on the Internet (continuous scale of 1–6, where high = 1, low = 6).
- Centrality of Visual Product Aesthetics (CVPA) (Phillips, 2007): Measures aesthetic importance to users in their purchasing decisions (continuous scale of 1–3, where 1 = lower reported importance of aesthetic appeal, 3 = higher reported importance of aesthetic appeal).

In addition, the scale of three dependent variables of perception of site aesthetics was measured using a bipolar 1–10 scale, with a positive Kansei closer to one and a negative Kansei closer to 10.

The data of user characteristics and their evaluation for one particular website (website no. one) with respect to three Kanseis (interesting [Int], attractive [Att], and imaginative [Im]) is partially shown in Table 4.1.

Table 4.1 contains different characteristics with different user ratings. It also contains discrete and continuous data. Furthermore, the entire data set includes some missing or incomplete data. In fact, these are the characteristics of real-world data. Each of these issues might cause some difficulties in analyzing the results. The problems and corresponding appropriate processes that the data must undergo are explained below. However, first, it is essential to provide some necessary definitions of knowledge and classification in order to provide the rough set theory-based framework for this specific problem.

TABLE 4.1 USER BACKGROUND CHARACTERISTICS AND EVALUATION OF KANSEIS FOR WEBSITE NO. ONE

Subject	Users Characteristics												Kansei		
	S/C	Gender	Age	Freq	Enter	Work	Com	Skill	Utility	Hedonic	Involvement	CVPA	Int	Att	Im
1	1	1	1	1	2	2	1	4.75	3.40	4.20	3.75	3	4	6	8
2	1	2	1	2	2	2	2	2.50	2.60	2.00	3.00	2	7	8	5
3	1	1	1	1	1	1	1	3.00	2.80	1.60	1.50	2	1	3	1
4	1	2	1	1	1	2	2	3.50	3.20	3.00	2.25	1	5	6	3
5	1	2	1	1	2	2	1	5.00	2.00	4.20	4.75	2	7	7	1
6	1	2	1	1	1	2	1	3.75	2.00	2.00	2.00	2	4	10	8
7	1	1	1	1	1	2	1	1.25	1.60	2.00	2.00	1	1	2	2
8	1	2	1	1	1	2	1	2.00	2.00	2.60	2.75	1	4	7	1
9	1	1	1	1	1	2	1	2.75	2.80	3.20	4.75	1	5	5	5
10	1	1	1	1	1	2	1	2.25	2.00	1.80	1.75	3	3	8	4
11	1	2	1	1	1	2	1	2.00	2.20	2.40	5.25	1	6	9	1
12	1	2	2	1	2	2	2	4.75	3.20	4.00	3.50	1	9	8	3
13	1	2	1	1	1	2	1	1.75	1.80	1.80	3.00	1	7	7	4
14	1	2	3	1	2	1	2	1.75	2.00	2.80	2.00	3	7	7	3
15	1	2	1	2	2	2	1	3.75	3.40	2.40	2.00	1	9	6	5

Knowledge, especially in cognitive science, enables human beings to classify objects. In this study, there is a finite set of objects, the 63 participants in this study, called U (the universe). The universe set (the set of participants) can be categorized by participant gender. Therefore, C = {X1, X2}, $X_1, X_2 \subseteq U$, $X_1, X_2 \neq \emptyset$, $X_1 \cap X_2 = \emptyset$, $X_1 \cup X_2 = U$ is an abstract knowledge about U, and X_1 and X_2 are two concepts or categories. If the set of participants is categorized based on multiple basic classifications (gender, age, Internet skill, etc.), then the family of classifications over U is called the knowledge base over U. Each of these classifications is called an equivalence relation. Each U/gender, U/age, or U/Internet skill means that the family of equivalence classes of each these equivalence relations is referred to as categories of "gender," "age," or Internet skill." A knowledge base is a relational system $K = (U, R), \neq \emptyset$, and R is a family of equivalence relations over U. In this study, R = {R_1 = S/C, R_2 = gender, R_3 = age...R_{11} = involvement, R_{12} = CVPA}.

If $P \subseteq R, P \neq \emptyset$, then $\cap P$ also is an equivalence relation, which is called IND(P) (indiscernibility relation over P). For example, P = {age, gender, Internet skill} ⊆ R is an equivalence relation, which is called the indiscernibility relation over P. Therefore, U/IND (P) or in the simple form U/P (the family of all equivalence classes of the equivalence relation IND (P)) indicates knowledge associated with the family of equivalence relation P (P-basic knowledge about U in K).

In summary then, U = {x_1,x_2,…x_{62}, x_{63}} is the universe set of participants (objects). Suppose that these people have different characteristics, such as age (under 25, over 25), gender (male, female), etc. Also, R = {R_1 = S/C, R_2 = gender, R_3 = age..., R_{11} = involvement, R_{12} = CVPA} is the family of equivalence relations over U. Therefore, $K = (U, R)$ is the knowledge base. Any combination of R constitutes different equivalence relations. For example, if K = (U, {R_1 = S/C, R_2 = gender, R_3 = age}), then these three equivalence relations have the following equivalence classes:

U/R_1 = {{people who are students}, {people who are non students}}

U/R_2 = {{men}, {women}}

U/R_3 = {{under or equal to 25 years old}, {above 25 years old}}

These are the elementary categories in the knowledge base. For instance, sets {people who are students}∩{men} = {men students}, {people who are non-students}∩{women} = {women non-students} are {R_1 = S/C, R_2 = gender} basic categories.

Any intersection of elementary categories constitutes basic categories that sometimes might not be available in the knowledge base.

4.1.2 Discretization

Since attribute (equivalence relation) values of "skill," "utility," "hedonic," and "involvement" are continuous, they must undergo a process called discretization to reduce continuous measures into fewer categories (in this study, two categories). The method of discretization for these sets of data is to simply round up the values and then generalize the categories by dividing them into two intervals in which the first interval includes any values between one and three (including three) and the rest of the values (four to six) are in the second interval.

4.1.3 Generalization Versus Specification

Knowledge base K = (U, P) is finer than knowledge base K' = (U, Q) (or knowledge Q is coarser than knowledge P), if IND (P) $\subset$IND (Q). In such a situation, Q is a generalization of P, or P is a specialization of Q. In our study, there are two types of specializations. The first is a specialization based on different equivalence relations. For instance, in our information system, each object can be characterized by any combination of R = {R_1 = S/C, R_2 = gender, R_3 = age... R_{11} = involvement, R_{12} = CVPA}. For example, objects can be classified by Q = {R_1 = S/C, R_2 = gender} or P = {R_1 = S/C, R_2 = gender, R_3 = age, R_{11} = involvement, R = CVPA} ((Chaparro, 2009). In this case, since the family of all equivalence classes of equivalence relations IND(P) is the subset of the family of all equivalence classes of equivalence relations IND(Q), then Q is the generalization of P (Q is coarser than P), and P is the specialization of Q (P is finer than Q). This kind of generalization is used in Chapter 5 (results), where objects are combined into groups, ignoring some insignificant differences.

Another type of generalization that is used in this study is the generalization (or combination) of subcategories into a larger category within each equivalence relation. For example, for the “age” equivalence relation, participants can be categorized in below 25, 25–40, 40–50, 50–60, and over 60 equivalence classes. Forming this number of subcategories is not always an advantage. In this case, it reduces the applicability and usability of the classifications and increases the complexity of generated rules. Therefore, in this kind of situation in order to obtain coarser knowledge, generalization is recommended. In this study, this kind of generalization was implemented for “age,” “frequency of use,” “CVPA,” and all Kansei evaluation data. Since the corresponding ratings were too fine to obtain ultimate usable coarser categorizations, it was decided to combine different classes (for example ten levels in Kansei evaluations) into two categories.

The generalization (combination) method in this study involves sorting the ratings and considering ratings below the midpoint as the first category and above the midpoint as the second category. For example, for the Kansei evaluation, ratings up to and including five are located in the first category, and ratings 6–10 are located in the second category. Although there are many generalization methods, this method was used since it was easy to apply and it combined fine categories into coarser classes based on rating scores, which generate more logical decision classes. This generalization helped to develop more efficient design rules in this study.

4.1.4 Incompleteness

Sometimes there are missing data in both condition attribute values and decision (Kansei) attribute values. RSES 2 (rough set theory software) will take care of this problem. This software uses three methods to estimate the missing values: remove objects with missing values, complete with most common or mean value, or complete with most common or mean value with respect to decision class. Since the set of data is not large, the first approach, which removes some of the valuable data, was not used. In this study, the third approach was used to fill the missing values with the mean over all attribute values in the decision class, since it generates more logical values for missing data compared to the second method.

4.1.5 Internal Scale

Another transformation function that was needed and used in this study was adjusting the evaluation of all websites that was given by each person with respect to each Kansei. For example, for participant one, his/her evaluation ratings with respect to one specific Kansei such as “interesting” for all websites were considered. Then, the middle point of the maximum rating score was used as the divider to separate the “low” category from the “high” category. If the maximum score given by a participant rating websites one to 24 with respect to “interesting” was six, then any rating score equal to or below three (one, two, and three) was considered a “low” category and above three (four, five, and six) was considered a “high” category.

In this method, the rating scores were adjusted according to users’ internal measures.

4.1.6 Randomness

To determine if the developed approach was able to capture the pattern in the data, the Kansei data evaluations were generated randomly. By comparing the information extracted from the random data and the original data using the developed approach, it can be determined if the approach can generate significantly useful information. If the results from random data are as good as results from original data, then this approach would be useless. This issue is discussed in section 5.7 using real data.

4.1.7 Website Data

The second stage of the proposed approach requires constructing an information system table consisting of user ratings for up to 24 different websites. The number of rows is based on the number of websites upon which a cluster of participants agree. The columns consist of website features and the ratings of each website, as condition and decision attributes, respectively. In this step, six main design attributes along with their corresponding subcategories were defined (Chaparro, 2009). Table 4.2 depicts the list of these features. In addition, Table 4.3 partially shows the attribute values of all 24 website features.

TABLE 4.2 WEBSITE MAIN FEATURES, SUBFEATURES, AND THEIR VALUE LEVELS

Attribute	Sub-attributes	Values Levels		
		1	2	3
Font	Size	Small	Large	
	Style	Normal (Italic)	Not Normal	
	Color	Black	Other	
	Type	Standard	Not Standard	
	Contrast	Black on White (Background and Foreground)	Not	Mixture
Image	Size	Small	Medium	Large
	Quality (Resolution)	Low (DPI)	High	
	Content	Natural	Not Natural	
Overall Page Layout	Density	High	Low	
	Color Combination	Good (Graphical Designers)	Bad	
	Consistency	Consistent	Not Consistent	
Picture		No Picture	Single	Multiple
Line length		Short	Long	
Display	Color	Black	White/Gray	Colorful
	Background Graphic	Has	Does Not	

TABLE 4.3 ATTRIBUTE VALUES OF 24 WEBSITES (PARTIALLY)

Website	Font			Image			Overall Page Layout	
	Size	Color	Type	Size	Quality(Resolution)	Content	Density	Color Combination
	Small-Large	Black-Other	Standard-Not Standard	Small-Medium-Large	Low (DPI)-High	Natural-Not Natural	High-Low	Good (Graphical Designers)-Bad
34	2	2	1	3	2	1	2	1
28	1	2	1	3	2	1	2	1
29	1	2	1	1	2	2	1	2
30	1	2	1	2	2	1	1	1
32	1	1	1	2	2	2	1	2
33	2	2	1	2	2	1	1	1
26	1	2	1	3	2	1	2	1
35	1	2	1	3	2	2	2	1
38	2	1	1	3	2	2	1	1
41	2	1	1	1	2	2	2	1
46	1	2	1	3	2	1	2	1
50	1	1	1	3	2	1	2	1
1	1	2	1	2	1	1	1	2
5	1	1	1	2	1	2	1	2
8	1	2	2	1	1	1	1	2
9	2	2	2	1	1	1	1	2

4.2 Preliminaries and Notations

Suppose the set of the universe U = {x_1, x_2,…, x_{62}, x_{63}} and the family of equivalence relations R = {R_1 = S/C, R_2 = gender, R_3 = age, ..., R_{11} = involvement, R_{12} = CVPA, R_i}, I = 1,2,3 (equal to the number of the decision classes d = {interesting, attractive, imaginative}) are given. Note that 63 participants evaluated 24 websites with respect to each decision class (Kansei). Therefore, there are $24 \times 3 = 72$ information systems S = (U, Q, V, ρ), where U is the finite set of participants, Q is the finite set of attributes (12 condition attributes and one decision attribute), and V = $\bigcup_{q \in Q} V_q$, in which V_q is a domain of the attributes. In this problem, V_q = {1, 2} for all attributes after generalization and discretization, and ρ: U × Q →V is a total function such that ρ(x, q)∈Vq for every q ∈ Q, x ∈U, called an information function. For example, ρ(x1(participant one), q_1 (gender)) = 1 means that the gender of participant one (or the attribute's value of participant one) is male = 1. For each decision class, let ϒ = {Y_1, Y_2} be a partition of U in the decision class "interesting," Y_1 = {people who think a specific website is interesting}, and Y_2 = {people who think a specific website is not interesting}. The rough set theory notation for this specific set of data is expressed as follows (for more notations and definitions of rough set theory refer to glossary in Appendix A):

4.2.1 Universe

Here, $U \neq \emptyset$ is considered the finite set of 63 participants.

4.2.2 Category or Concept

Each subset of participants X ⊆ U constitutes a concept or a category in U. There are two types of categories or concepts. If participants are categorized according to one specific characteristic (say gender), then the two classes of male and female are the elementary categories or concepts in U. Basic categories are the combination of elementary categories, for example, people who are male and less than 25 years of age, and have a high level of using the Internet for entertainment purposes.

4.2.3 Abstract Knowledge

Any family of concepts or categories is called an abstract knowledge about U or, in short, knowledge about U. For example, the classification of participants in terms of their age can be abstract knowledge about participants, i.e., C = {X_1, X_2}, X_1 = {people who are young (under 25)}, X_2 = {people who are older (above 25) than the first group}.

4.2.4 Basic Classification or Partition

Classification of the participants in terms of any specific equivalence relation or classification, such as gender, age, or even a Kansei such as attractiveness (people who think a specific website is attractive), constitutes a basic partition. In other words, in this data C = {X_1, X_2}, $X_i \subseteq U$, $X_i \neq \emptyset$, $X_i \cap X_i = \emptyset$, for i#j, i, j = 1,2.

4.2.5 Knowledge Base over U

Any family of classifications over U (different basic classification, e.g., according to skill, Kansei [imaginative], CVPA, etc.) is the knowledge base over U. In other words, participants could be classified based on any combination of attributes (both condition and decision attributes, 12 users characteristics, and one Kansei [or even three Kansei]). For example, the following family of classifications is one instance of many knowledge bases over all participants:

- {People who are male, students, and their CVPA is high}
- {People who are male, students, and their CVPA is low}
- {People who are male, not students, and their CVPA is high}

- {People who are male, not students, and their CVPA is low}
- {People who are female, students, and their CVPA is high}
- {People who are female, students, and their CVPA is low}

4.2.6 Equivalence Relation and Equivalence Class

The term "equivalence relation," denoted by R, is often used instead of classification. Furthermore, U/R means the family of equivalence classes of R (or classification of U). In addition, any of the sets in the above knowledge base is an equivalence class, e.g., {people who are female, students, and their CVPA is high}.

4.2.7 Indiscernibility Relation over P

If $P \subseteq R$ and $P \neq \emptyset$, then an associated equivalence relation IND (P) can be defined as

$$\mathrm{IND}\ (\mathrm{P}) = \left\{ (\mathrm{x}, \mathrm{y}) \in \mathrm{U}^2 \middle| \forall \alpha \in \mathrm{P}, \alpha(\mathrm{x}) = \alpha(\mathrm{y} \right\} \tag{4.1}$$

For example, if P = {gender, S/C, CVPA}, then any of the sets in the above knowledge base is an indiscernibility relation over P.

4.2.8 U/IND (P) or U/P

U/IND (P) or U/P is the family of all equivalence classes of the equivalence relation IND (P) and is called P-basic knowledge or basic knowledge about U in K = (U,R) (relational system). For example if P = {gender, S/C, CVPA}, then U/P = the following:

- {People who are male, students, and their CVPA is high}
- {People who are male, students, and their CVPA is low}
- {People who are male, not students, and their CVPA is high}
- {People who are male, not students, and their CVPA is low}
- {People who are female, students, and their CVPA is high}

- {People who are female, students, and their CVPA is low}

4.2.9 Equivalence Classes of IND (P)

Equivalence classes of IND (P) are called basic categories (or concepts) of knowledge P. For example, if people are classified based on P_1 = {gender} or P_2 = {CVPA} separately, then equivalence classes of IND (P), P = {gender, CVPA}, are {male, CVPA high}, {male, CVPA low}, {female, CVPA high}, and {female, CVPA low}. Note that some classes may not be available in this specific knowledge base.

4.2.10 Information System

In the first stage of the proposed approach, the rows of information system table represent participants, and the columns show user characteristics; while in the second stage, the rows of the information system table stand for different products, and the columns correspond to the product features. The entries of the table are attribute values or descriptors. In both stages, the information system table is a decision table, since the set of attributes is divided into two subsets: condition attributes (customer characteristics or product features) and decision attributes (Kanseis evaluations).

4.2.11 Set Approximation

Set approximation includes the following:

- Lower approximation (denoted by $\underline{P}Y$), for any set such as equivalence classes of IND (P), or specific partition, can be defined as the union of all granules that are entirely included in the set.
- Upper approximation (denoted by $\overline{P}Y$) is the union of all granules that have a non-empty intersection with the set.

- Boundary region is the difference between the upper and lower approximations of the set.

For example, for partition “people who see the website as attractive,” if there are two indiscernible sets {male, CVPA high, attractive} and {female, CVPA high, attractive}, then the lower approximation of this particular partition includes the people who are male or female and believe that the website is attractive. Among those people, whether male or female, with high CVPA, some believe the website is attractive and some believe the website is unattractive. All those users belong to the boundary region. The upper approximation is the union of the lower approximation set with the boundary region set.

4.2.12 Categories of Vagueness in the Data

As mentioned previously, there are four basic classes of rough sets or vagueness:

- $\underline{P}Y \neq \emptyset \; and \; \overline{P}Y \neq U, iff \; Y \; is \; roughly \; P\_definable$
- $\underline{P}Y = \emptyset \; and \; \overline{P}Y \neq U, iff \; Y \; is \; internally \; P\_indifinable$
- $\underline{P}Y \neq \emptyset \; and \; \overline{P}Y = U, iff \; Y \; is \; externally \; P\_definable$
- $\underline{P}Y = \emptyset \; and \; \overline{P}Y = U, iff \; Y \; is \; totally \; P\_indefinable$

The last category of vagueness implies that any element of U cannot be decided, whether it belongs to Y or –Y, using P. This is exactly what existed in the data before their characteristics (condition attributes) were considered and before it was assumed that people who are surveyed are characterized in the same way.

4.2.13 Set Approximation Measures

In this section, the corresponding approximation measures for one set of data (among 72 sets) are demonstrated. These measures are used to find the most powerful classificatory attributes.

4.2.14 Vagueness and Uncertainty

Vagueness can be explained by approximations in the properties of a set. On the other hand, uncertainty, which is the properties of elements of a set, can be expressed by the rough membership function. In this data, the set of people who believe a particular website is attractive, imaginative, or interesting is vague. This vagueness creates uncertainty for the person who has similar characteristics with another person, but one of them belongs to the positive-feeling group and another to the negative-feeling group.

4.2.15 Positive Region

The equation $POS_P(У) = \bigcup_{Y_i \in У} \underline{P}\, Y_i$ is called the P-positive region of partition U/У (or U/I (У) or indiscernibility У) with respect to P and applies to the set of all objects of the universe U, which can be certainly classified as the member of each class of У = {Y_1, Y_2…Y_n} using knowledge P. For example, when the set (or the reduced set-reduct set) of condition attributes {gender, CVPA} is used to classify users, the positive region of partition У = {Y_1 = users who believe the website is attractive, Y_2 = users who believe the website is not attractive} is the set of all users who certainly can be classified as either Y_1 or Y_2.

4.2.16 Dependency of Attributes

An attributes set D depends totally upon attributes set C, denoted by $C \Rightarrow D$, if all values of attributes belonging to D are uniquely determined by values of attributes belonging to C. If only some values of D are determined by values of C, then there is partial dependency. Therefore, if D depends upon C in a degree $\mathrm{k}\ (0 \leq \mathrm{k} \leq 1)$, denoted by $\mathrm{C} \underset{\mathrm{k}}{\Rightarrow} \mathrm{D}$, if $\mathrm{k} = \gamma(\mathrm{C}, \mathrm{D})$, then k = 1 applies to the totally dependency of D on C, whereas $\mathrm{k} \leq 1$ denotes partial dependency of C on D. The coefficient k expresses the ratio of all elements of the universe, which can be properly classified to blocks of the partition U/D, employing attributes C.

Therefore, the dependency of attributes shows the consistency of the decision table. Also, if $C \underset{k}{\Rightarrow} D$, if k = 1, then $I\ (C) \subseteq I(D)$, which implies the fact that the partition generated by C is finer than the partition generated by D.

In the above example, if the values of D (here 1 = low attractiveness, 2 = high attractiveness) are uniquely determined by values of C = {gender (1 = male, 2 = female), CVPA (1 = high, 2 = low)}, then D totally depends upon C; otherwise, D partially depends upon C. It is obvious that if all users are in the positive region, then k = 1.

4.2.17 Reduction of Knowledge

This section attempts to determine if all available knowledge is necessary to define some basic categories in the knowledge base and, if not, which part of the knowledge is redundant and which part is essential. Specifically, in this approach, the focus is on eliminating redundant attributes and attribute values for multiple users and multiple products for one specific Kansei.

4.2.18 Elimination of Redundant Attributes

4.2.18.1 Reduct Set of Attributes

Subset *P`* of *P* is a reduct of *P* and shown by *Red (P)*, if *P`* is independent and *I (P`) = I (P).* The reduced set of attributes provides the same quality of sorting as the original set of attributes $\gamma_P(У) = \gamma_{P`}(У)$. In other words, the minimal subset of attributes that has the same power of classification of elements as the whole set of attributes is the reduct set. For example, if one of the reduct sets is considered as {gender, CVPA}, then this means that not all of user characteristics (available knowledge) are necessary to define some user categories in the knowledge base. Therefore, the rest of the user characteristics are redundant knowledge.

4.2.18.2 Core of Attributes

The core set of attributes includes the most important attributes in which none can be removed without affecting the classification power of attributes. The core is the intersection of all reducts and denoted by the following: Core (P) = $\bigcap Red(P)$

In this study, either one or multiple reduct sets can be identified for each information system table (IST). The core set is the intersection of reduct sets that could be available or not.

4.2.18.3 Indispensable and Dispensable Attribute

Let $P \subseteq Q$, and let α belong to P. Then α is dispensable in P, if I (P) = I (P-{ α}); otherwise, α is indispensible in P.

In this study, any user characteristic that does not belong to a reduct set can be considered as a dispensable characteristic. Characteristics that belong to a core set are indispensible characteristics.

4.2.19 Independent Set

Set P is independent, if all its attributes are indispensible.

4.2.20 Significance of Attributes

The significance of an attribute can be measured by the effect of removing it from an information system table and calculating it as follows:

$$\sigma_{(C,D)}(\alpha) = \frac{\gamma(C,D)-\gamma(C-\{a\},D)}{\gamma(C,D)} = 1 - \frac{\gamma(C-\{a\},D)}{\gamma(C,D)} \tag{4.2}$$

Obviously, $0 \leq \sigma_{(C,D)}(\alpha) \leq 1$, and the greater the number $(\sigma_{(C,D)}(\alpha))$, the more important the attribute α. In addition, the coefficient $\sigma_{(C,D)}(\alpha)$ can be considered an error that occurs when attribute α is dropped.

Obviously, the number of occurrences of user characteristics in reduct sets defines the significance of the characteristics, meaning that user characteristics that are in the core set are more significant attributes than the characteristics that are in the reduct set. Characteristics not belonging to any reduct sets are not significant.

4.2.21 Elimination of Redundant Attribute Values

4.2.21.1 Indispensable and Dispensable Values of Attributes

The value of attribute $\alpha \in B$ is dispensable for x (an object in universe), if $P(x) = P^{a}(x)$, where $P^{a} = P - \{\alpha\}$; otherwise, the value of attribute α is indispensable for x.

4.2.21.2 Orthogonal Set

If $\forall \ \alpha \in P, a(x)$ is dispensable (the value of α is dispensable) for x, then P is called the orthogonal for x.

4.2.21.3 Reduct Set of Values

The subset P`$\subseteq$ P is a value reduct of P for x, iff P` is orthogonal for x and P`(x) = P(x).

4.2.22 An Important Property

If P` is a reduct of P, then neither $\{a\} \Rightarrow \{b\}$, nor $\{b\} \Rightarrow \{a\}$ holds for every $a, b \in$ P`. It means that all attributes in a reduct are pair wise independent.

4.2.23 Indiscernibility Matrix

The indiscernibility matrix is denoted by M (P), and each entry of this matrix c_{ij} consists of all attributes that discern object x_i and x_j and are defined as follows: $c_{ij} = \{a \in C \mid a(x_i) \neq a(x_j)\}$ I, j = 1, 2…m, (m represents the number of attributes), as mentioned in Chapter 3. The reduct is the minimal subset of attributes that discerns all objects discernible by the whole set of attributes. This matrix was constructed in section 3.4 of Chapter 3.

4.2.24 Indiscernibility Function

An indiscernibility function f_A for information system A is a Boolean of m Boolean variables $a^*_{\dots}$ a^*_m (corresponding to the attributes $a_1 \dots a_m$) defined as follows: $(a^*_1 \dots a^*_m)$ $= \wedge \left\{ \vee\, c^*_{ij} \middle| 1 \leq j \leq i \leq n, c_{ij} \neq \emptyset \right\}$, where $c^*_{ij} = \left\{ a^* \middle| a \in c_{ij} \right\}$, and n is the number of objects.

In addition, this function was calculated in section 3.4 using equation (3.2).

4.2.25 Significant of the Set of Attributes

If $B \subset C$ (C and D denote the set of condition attributes and decision attributes), then $\varepsilon(C, D)(B)$. The significant coefficient of set of attributes or error of reduct approximation is calculated as

$$\varepsilon(C, D)(B) = \frac{\gamma(C,D) - \gamma(C-B,D)}{\gamma(C,D)} = 1 - \frac{\gamma(C-B,D)}{\gamma(C,D)} \tag{4.3}$$

If B is a reduct of C, then $\varepsilon(C, D)(B)$ = 1, which means that removing any reduct from a set of attributes makes it impossible to make certain decisions.

Any subset B of C is called an approximate reduct of C. The error of reduct approximation (denoted by $\varepsilon(C, D)(B)$) tells us exactly how the set of attributes B approximates the set of condition attributes C. This measure is calculated as

$$\varepsilon(C, D)(B) = \frac{\gamma(C,D) - \gamma(B,D)}{\gamma(C,D)} = 1 - \frac{\gamma(B,D)}{\gamma(C,D)} \tag{4.4}$$

For any reduct set of C, $\varepsilon(C, D)(B)$ = 0.

The next section introduces the performance measures of decision rules, which are used to define efficient rules.

4.2.26 Bayes' Theorem and Rough Set

Bayes' theorem is a mechanism that helps us learn from data. Let H = hypothesis, D = data, P (H) = probabilistic statement of belief about H before obtaining data D (what is known about H without knowing of the data-priori distribution of H), and P (H\D) = probabilistic statement of belief about H after obtaining data D (what is known about H given knowledge of the data-posterior distribution of H given D). Then the Bayes' theorem is

$$P(H|D) = P(D|H) \times P(H) \,/\, P(D) \tag{4.5}$$

Let S = (U, A) be a decision table. With every $B \subseteq A = C \cup D$, there is an associated set of formulas F or (B), which are built up from the attribute-value pair (a, v), where $a \in B$ and $v \in V_a$ by means of logical connectives $\wedge(and), \vee(or), \backsim (not)$ in the standard way.

For any $\phi \in F\ or\ (B)$ by$\|\phi\|_S$, the set of all objects $x \in U$ satisfying ϕ in S defined is denoted inductively as follows: $\|(a, v)\|_S$ = $\{x \in U\colon a(v) = x\}$, $\forall \in B\ and\ v \in V_a$.

4.2.27 Decision Rule and Decision-Rule Performance Measures

A decision rule in L(S) (decision language) is an expression $\phi \rightarrow \psi$, where $\phi \in \mathrm{F\ or\ (C)}, \psi \in \mathrm{F\ or\ (D)}$ (where C and D are condition and decision attributes, respectively). In this research, four performance measures, which are explained below, were used to compare rules and define the most efficient.

4.2.27.1 Support of Decision Rule

The following number is called the support of the rule $\phi \rightarrow \psi$ in S:

$$\mathrm{supp_S}(\phi, \psi) = |(\|\phi \wedge \psi\|_S)| \tag{4.6}$$

Also, without using decision language, by referring only to the decision table, the following is called the support of the decision rule $\mathrm{C} \rightarrow_{\mathrm{x}} \mathrm{D}$:

$$\mathrm{supp}_x(C, D) = |C(x) \cap D(x)| \tag{4.7}$$

where S = (U,C,D) is a decision table, and $x \in U$ determines a sequence $c_1(x), c_2(x), c_3(x), \dots, c_n(x), d_1(x), d_2(x), d_3(x), \dots, d_m(x)$, where $\{c_1, c_2, c_3, \dots, c_n\} = C$ and $\{d_1, d_2, d_3, \dots, d_n\} = D$.

4.2.27.2 Certainty Factor of Decision Rule (Confidence Coefficient in Data Mining)

With every decision rule, $\phi \to \psi$, there is an associated conditional probability as

$$\mathrm{cer}_S(\phi, \psi) = \pi_S(\psi|\phi) = PU(\|\psi\|_S | \|\phi\|_S) = \frac{|(\|\phi \wedge \psi\|_S)|}{|(\|\phi\|_S)|} \tag{4.7}$$

where $\|\phi\|_S \neq \emptyset$

If $\mathrm{cer}_S(\phi, \psi) = \pi_S(\psi|\phi) = 1$, then $\phi \to \psi$ is called a certain decision; otherwise, the decision rule is referred to an uncertain decision rule in S. Also, without using decision language:

$$\mathrm{cer}_x(C, D) = \frac{|C(x) \cap D(x)|}{|C(x)|} = \frac{\mathrm{supp}_S(C, D)}{|C(x)|} = \frac{\sigma_x(C, D)}{\pi(C(x))} \tag{4.8}$$

where $\pi(C(x)) = \frac{|C(x)|}{|U|}$.

4.2.27.3 Coverage Factor of Decision Rule

The coverage factor of the decision rule is

$$\mathrm{cov}_S(\phi, \psi) = \pi_S(\phi|\psi) = PU(\|\phi\|_S | \|\psi\|_S) = \frac{|(\|\phi \wedge \psi\|_S)|}{|(\|\psi\|_S)|} \tag{4.9}$$

where $\|\psi\|_S \neq \emptyset$. Also, without using decision language,

$$\mathrm{cov}_x(C, D) = \frac{|C(x) \cap D(x)|}{|D(x)|} = \frac{\mathrm{supp}_S(C, D)}{|D(x)|} = \frac{\sigma_x(C, D)}{\pi(D(x))} \tag{4.10}$$

where $\pi(D(x)) = \frac{|D(x)|}{|U|}$.

4.2.27.4 Strength of Decision Rule

The strength of the decision rule is

$$\sigma_S(\phi, \psi) = \frac{supp_S(\phi, \psi)}{|U|} = \pi_S(\psi|\phi).\ \Pi_S(\phi) \tag{4.11}$$

Without using decision language,

$$\sigma_x(C, D) = \frac{supp_S(C, D)}{|U|} \tag{4.12}$$

4.2.28 Efficient Rules

In this study, there are multiple users, multiple designs (realizations) of websites, and multiple Kanseis. In the first stage of the proposed approach, the reduct sets along with customer segmentation rules are generated for each Kansei. In the second stage, for each specific Kansei, the most influential product features and design rules are generated. If some of the ultimate design rules are the same for different Kansei, then applying these rules in website design can satisfy multiple Kanseis.

The purpose of this section is to find the most efficient user classification (the first stage) and design rules (the second stage). Therefore, some criteria are needed to define an efficient rule. We call a rule an efficient rule if it meets the following criteria:

- Covers as many objects (participants) as possible (supports of the rule).
- Has high certainty.
- Has high coverage.
- Contains as few attributes as possible (if it covers the same number of objects) (axiom two of axiomatic design [Wortman, Richardson, Gee, Williams, et al., (2007)].
- Has greater rule strength (if it meets the above two criteria).

These criteria are used as performance measures of generated rules for identifying efficient rules.

4.3 Software Illustration

Sixty-three users, described using 12 characteristics, evaluated 24 websites with respect to attractiveness. Table 4.4 illustrates user characteristics and their corresponding evaluations of some websites (websites 26, 28, and 34) with respect to attractiveness.

TABLE 4.4 PARTIAL ILLUSTRATIONS OF ALL DATA

Participants	Student/ Community	Gender	Age	Frequency of use	Participants' entertainment	Participants' work	Participants' communication	Skill	Utility	Hedonic	Involvement	CVPA	Att26	Att28	Att34
1	1	1	1	1	2	2	1	2	1	2	2	2	1	1	1
2	1	2	1	2	2	2	2	1	1	1	1	1	1	1	1
3	1	1	1	1	1	1	1	1	1	1	1	2	1	2	1
4	1	2	1	1	1	2	2	2	1	1	1	1	1	2	1
5	1	2	1	1	2	2	1	2	1	2	2	2	1	1	1
6	1	2	1	1	1	2	1	2	1	1	1	1	1	1	1
7	1	1	1	1	1	2	1	1	1	1	1	1	1	1	2
8	1	2	1	1	1	2	1	1	1	1	1	1	1	1	1
9	1	1	1	1	1	2	1	1	1	1	2	1	1	1	1
10	1	1	1	1	1	2	1	1	1	1	1	2	2	1	1
11	1	2	1	1	1	2	1	1	1	1	2	1	1	1	1
12	1	2	1	1	2	2	2	2	1	2	2	1	1	1	1
13	1	2	1	1	1	2	1	1	1	1	1	1	1	1	1
14	1	2	2	1	2	1	2	1	1	1	1	2	1	1	1
15	1	2	1	2	2	2	1	2	1	1	1	1	1	1	1
…	…	…	…	…	…	…	…	…	…	…	…	…	…	…	…

CVPA (centrality of visual product aesthetics), Att26 (attraction level of website 26), Att28 (attraction level of website 28), Att34 (attraction level of website 34)

In addition, Figure 4.1 shows a schematic of the rough set project, including all data, reduct sets, and rules pertinent to one specific Kansei (here, attractiveness), and 24 websites generated by RSES 2.2.2. As can be seen in Figure 4.1, two center squares (called Att and AttR) represent user characteristics and their evaluation data, respectively (one Kansei, attractiveness, and 24 websites) before and after completion of missing data. Also, each path includes three icons: one for the data of each website, one for reduct sets, and one for the rules.

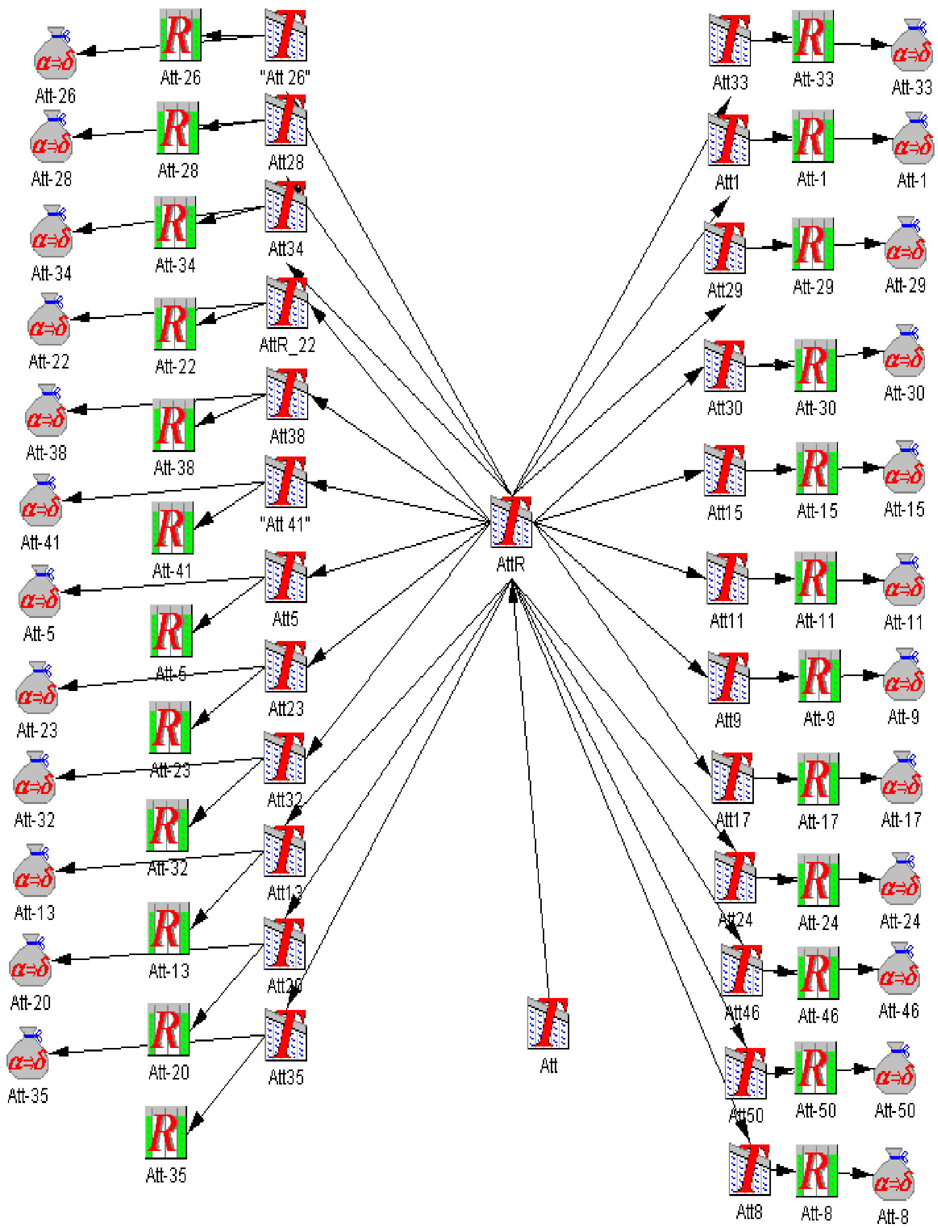


Figure 4.1. Schematic of rough set project: Data, reduct, and rule sets for attractiveness and 24 websites generated by RSES 2.2.2.

Figure 4.2 depicts a partial data set (in RSES format) from which reduct sets and rule sets were extracted. The domain of each attribute (user characteristics) is {1, 2}. This project has 72 data sets of this kind.

Table: "Att 26"

63 / 13	STUC...	Gender	AGE	FREQ_...	PENTER	PWORK	PCOMM	Skill	Utility	Hedonic	Involv...	CVTER	attrac26
0:1	1	1	1	1	2	2	1	2	1	2	2	2	1
0:2	1	2	1	2	2	2	2	1	1	1	1	1	1
0:3	1	1	1	1	1	1	1	1	1	1	1	2	1
0:4	1	2	1	1	1	2	2	2	1	1	1	1	1
0:5	1	2	1	1	2	2	1	2	1	2	2	2	1
0:6	1	2	1	1	1	2	1	2	1	1	1	1	1
0:7	1	1	1	1	1	2	1	1	1	1	1	1	1
0:8	1	2	1	1	1	2	1	1	1	1	1	1	1
0:9	1	1	1	1	1	2	1	1	1	1	2	1	1
0:10	1	1	1	1	1	2	1	1	1	1	1	2	2
0:11	1	2	1	1	1	2	1	1	1	1	2	1	1
0:12	1	2	1	1	2	2	2	2	1	2	2	1	1
0:13	1	2	1	1	1	2	1	1	1	1	1	1	1
0:14	1	2	2	1	2	1	2	1	1	1	1	2	1
0:15	1	2	1	2	2	2	1	2	1	1	1	1	1
0:16	1	2	1	2	2	2	2	1	1	2	2	1	2
0:17	1	2	1	1	1	1	1	1	1	1	1	2	2
0:18	1	2	1	1	1	1	2	1	1	1	1	1	2
0:19	1	2	2	1	2	2	2	2	1	1	2	2	1
0:20	1	2	1	2	1	2	2	2	2	1	2	1	1
0:21	1	2	1	2	2	2	2	1	1	1	2	1	1
0:22	1	2	1	1	2	2	1	1	1	1	1	1	1
0:23	1	2	1	1	1	2	1	1	1	1	1	2	2
0:24	1	1	1	2	1	2	2	1	1	1	2	2	1
0:25	1	2	1	1	2	1	1	1	1	1	1	2	1
0:26	1	1	1	1	2	1	2	1	1	1	1	2	1
0:27	1	2	1	2	1	1	2	1	1	2	1	2	1
0:28	1	1	1	1	1	2	1	1	1	1	2	1	1
0:29	1	1	1	1	1	1	2	2	1	1	2	2	1
0:30	1	2	1	1	1	1	2	1	1	1	1	2	1
0:31	1	1	1	2	1	2	2	1	2	1	2	1	2
0:32	2	1	1	1	1	2	1	1	1	1	2	2	1
0:33	2	2	2	1	2	2	2	1	1	2	2	1	2
0:34	2	1	1	1	1	2	2	1	1	1	2	2	2
0:35	2	2	1	1	1	1	2	1	1	1	2	2	1
0:36	2	2	2	1	1	1	1	1	1	2	1	2	1
0:37	2	2	1	2	1	2	1	1	1	1	1	2	1
0:38	2	2	2	1	2	1	2	2	1	2	2	2	1
0:39	2	2	1	2	1	2	2	2	1	1	1	1	1
0:40	2	1	1	1	1	2	1	1	1	1	2	2	1
0:41	2	1	1	1	2	1	2	1	1	1	1	2	1
0:42	2	1	1	1	1	1	1	2	1	1	1	1	1
0:43	2	1	2	1	2	1	2	2	1	2	2	1	1
0:44	2	2	2	1	2	2	2	2	1	1	1	1	1

Figure 4.2 Partial data set of website 26 evaluation with respect to attractiveness.

Figure 4.3 shows the statistics of attributes from an information system table that includes customer evaluation of website 26 with respect to "attractiveness." As shown, the number of attributes is 13, which includes 12 condition attributes along with one decision attribute.

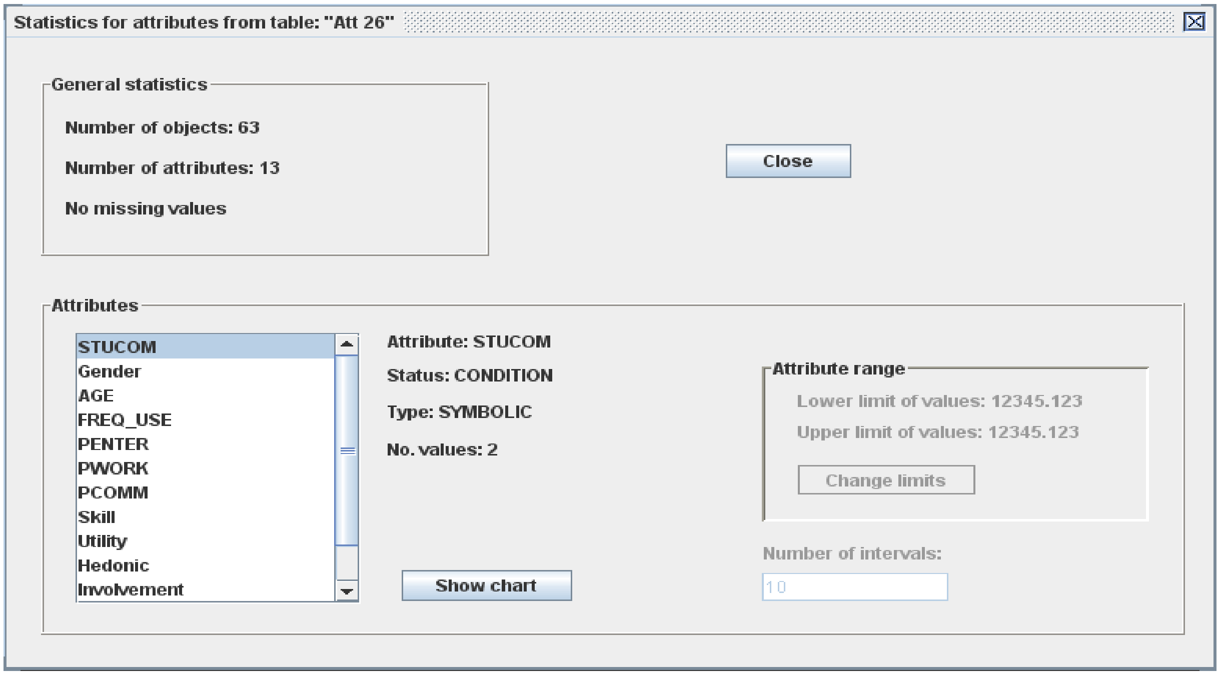


Figure 4.3 Statistics for attributes of website 26—attractivness.

Table 4.5 shows statistics for the reduct sets of the information system table for website 26—attractiveness. There are two reduct sets for this IST (IST 26), each one including nine attributes. The core set of these reduct sets includes eight attributes. Table 4.5 illustrates the number of occurrences of each attribute in reducts, and as can be seen, the core set includes attributes with a higher percentage of occurrences in reducts. In other words, elements of the core set have shown themselves more frequently in reduct sets than elements not belonging to the core set. These elements imply the essential part of the knowledge about the users and cannot be eliminated without reducing the ability to classify users to different categories.

TABLE 4.5 STATISTICS FOR REDUCT SET OF WEBSITE 26—ATTRACTIVENESS

Attribute	Occurrence of Attributes in Reducts										
	SC	G	A	FU	PE	PW	PC	S	U	I	CVPA
Count	2	2	2	1	2	2	2	2	2	2	2
Percent	11.1	11.1	11.1	5.6	11.1	11.1	11.1	11.1	11.1	11.1	11.1
Core	Yes	Yes	Yes	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Reduct No.	Length of Reducts			Size of Core
	Max	Min	Mean	
2	9	9	9	8

Table 4.6 shows the content of each reduct set and the corresponding positive region. In other words, for this specific set of data, the reduct sets explain the knowledge about data that can be reduced by eliminating the three unnecessary user characteristics. These three redundant characteristics are either {skill, age, utility} or {age, frequency of use, utility}.

TABLE 4.6 REDUCT SETS OF WEBSITE 26—ATTRACTIVENESS

Reduct Set	Size	Positive Region	Reducts
1	9	0.968	{SC, G, FU, PE, PW, PC, H, I, CVPA}
2	9	0.968	{SC, G , PE, PW, PC, S, H, I, CVPA}

Therefore, either of the following sets of user characteristics along with website 26 features influences users' Kansei "attractiveness." These two reduct sets, shown in Table 4.6 can be translated as follows:

C (condition attributes) = {SC (student/community), G (gender), A (age), FU (frequency of use), PE (participant's entertainment), PW (participant's work), PC (participant's communication), S (skill), U (utility), H (hedonic), I (involvement), CVPA (centrality of visual product aesthetics}

$\Upsilon = \{Y_1, Y_2\}$

where Y_1 is the set of users that evaluate website 26 as attractive, and Y_2 is the set of users that evaluate the website as unattractive.

$RED^1_{\Upsilon}(\mathrm{C}) =$ {STUCOM, gender, PENTER, PWORK, PCOMM, involvement, CVTR, frequency of use, hedonic}

$RED^2_{\Upsilon}(\mathrm{C}) =$ {STUCOM, gender, PENTER, PWORK, PCOM, skill, involvement, CVTR, hedonic}

In addition, the core set of the above reduct sets can be obtained by finding the intersection of the reduct sets:

$$\mathrm{COREY(C)} = \mathrm{RED}^1_Y(C) \cap \mathrm{RED}^2_Y(C) = \{\mathrm{STUCOM, gender, PENTER, PWORK, PCOMM, involvement, CVTR, hedonic}\}$$

The reduct sets have the same positive region (0.968) (Table 4.6).

Table 4.7 shows the statistics of the rules, which are generated based on the data of IST of website 26—attractiveness.

TABLE 4.7 STATISTICS OF RULES BASED ON DATA OF IST OF WEBSITE 26—ATTRACTIVENESS

Support of Rules			Length of Rules Premises			Distribution of Rules Among Decision Class		No. of Rules
Max	Min	Mean	Max	Min	Mean	Decision Class	Count	
19	1	3.5	6	1	4	1	121	233
						2	130	

As Table 4.7 shows, 233 rules are generated, and the number of matches or supporters of the rules are between one and 19 persons. Later, we call this number “support of the decision rule,” which shows the number of people who are described by the particular rule. In addition, the “length of the rule premises” indicates the minimum and maximum number of condition attributes that are in the rule. This number is called the size of the rule. As shown in Table 4.7, some groups of users can be described by one condition attribute, while other groups can be defined by six condition attributes (length of rule premises), and it is known that the fewer attributes involved, the better the applicability of the rule. Furthermore, Table 4.7 illustrates the distribution of rules within decision classes and indicates the number of the rules for each decision class. There are 120 rules that correspond to attractiveness (decision class one), while 130 rules support unattractiveness (decision class two). The reason that the total number of rules is less than the sum of the number of rules for both classes is that some condition attributes describe both those people who believe the website is attractive and those people who evaluate the website as unattractive. This is the result of the imperfectness of the positive region.

Figure 4.4 shows a histogram of the size of the rules, most of which have four condition attributes.

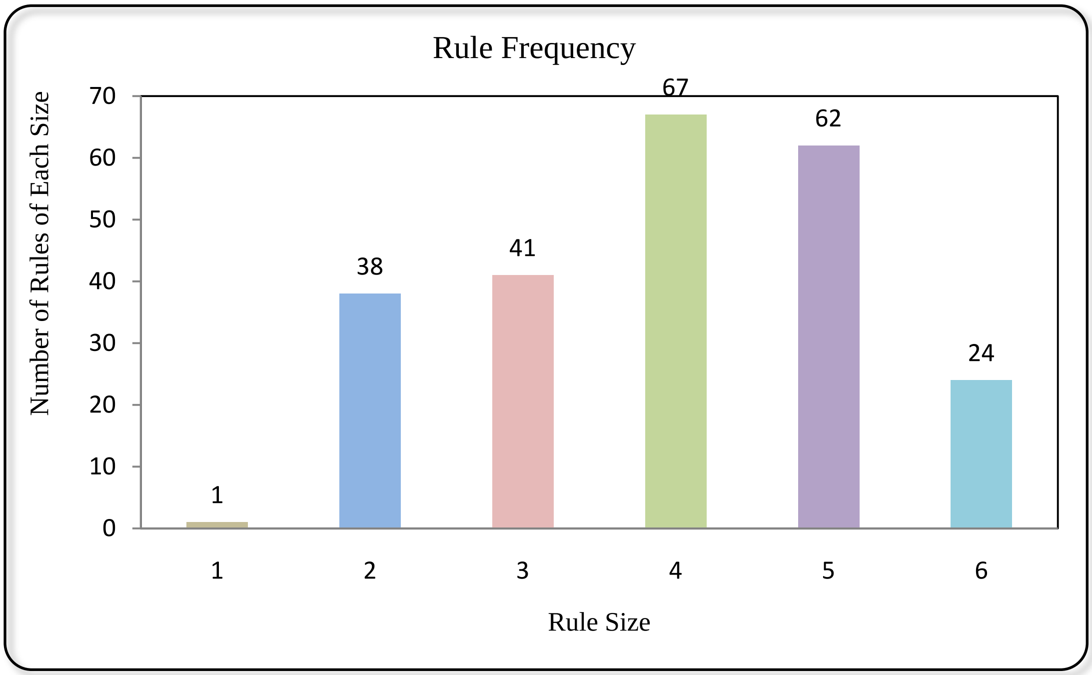


Figure 4.4 Size of rules of IST 26—attractive.

Figure 4.5 shows part of the table that includes all rules generated using IST of website 26—attractiveness. Here, each if-then rule is shown in detail. For example, rule no. 2, which is supported by 17 people, is as follows: if stucom = 2 and Pcomm = 1→ attractiveness = 1). This rule can be interpreted as "people who are not students and do not use the Internet for communication purposes often believe that website 26 is attractive."

Rule set: "Att 26"

(1-233)	Match	Decision rules
1	19	(Skill=2)=>(attrac26={1[19]})
2	17	(STUCOM=2)&(PCOMM=1)=>(attrac26={1[17]})
3	17	(STUCOM=2)&(Involvement=1)=>(attrac26={1[17]})
4	16	(PWORK=2)&(Utility=1)&(Hedonic=1)&("CVTER "=1)=>(attrac26={1[16]})
5	16	(PCOMM=1)&("CVTER "=1)=>(attrac26={1[16]})
6	15	(PENTER=2)&(Hedonic=1)=>(attrac26={1[15]})
7	15	(Gender=2)&(PWORK=2)&(Hedonic=1)&("CVTER "=1)=>(attrac26={1[15]})
8	12	(PENTER=2)&("CVTER "=2)=>(attrac26={1[12]})
9	12	(PCOMM=1)&(Involvement=2)=>(attrac26={1[12]})
10	12	(PENTER=2)&(Involvement=1)=>(attrac26={1[12]})
11	12	(FREQ_USE=1)&(PWORK=2)&(Hedonic=1)&("CVTER "=1)=>(attrac26={1[12]})
12	11	(Gender=2)&(PWORK=2)&(PCOMM=2)&(Hedonic=1)=>(attrac26={1[11]})
13	11	(PWORK=2)&(Involvement=1)&("CVTER "=1)=>(attrac26={1[11]})
14	11	(PENTER=1)&(PWORK=2)&(Utility=1)&("CVTER "=1)=>(attrac26={1[11]})
15	11	(STUCOM=2)&(Hedonic=1)&("CVTER "=1)=>(attrac26={1[11]})
16	10	(AGE=1)&(FREQ_USE=1)&(PWORK=2)&("CVTER "=1)=>(attrac26={1[10]})
17	10	(STUCOM=1)&(FREQ_USE=1)&(PWORK=2)&("CVTER "=1)=>(attrac26={1[10]})
18	9	(Gender=2)&(FREQ_USE=2)&(Hedonic=1)=>(attrac26={1[9]})
19	9	(FREQ_USE=2)&(Utility=1)&(Hedonic=1)=>(attrac26={1[9]})
20	9	(FREQ_USE=1)&(PENTER=1)&(PWORK=2)&("CVTER "=1)=>(attrac26={1[9]})
21	9	(Gender=2)&(PENTER=1)&(PWORK=2)&("CVTER "=1)=>(attrac26={1[9]})
22	9	(AGE=2)&(Hedonic=1)=>(attrac26={1[9]})
23	9	(AGE=2)&("CVTER "=2)=>(attrac26={1[9]})
24	9	(STUCOM=2)&(PENTER=1)&("CVTER "=1)=>(attrac26={1[9]})
25	8	(STUCOM=1)&(FREQ_USE=1)&(PENTER=2)=>(attrac26={1[8]})
26	8	(PENTER=2)&(PCOMM=1)=>(attrac26={1[8]})
27	8	(STUCOM=1)&(FREQ_USE=1)&(Involvement=2)=>(attrac26={1[8]})
28	8	(Gender=1)&(PWORK=1)=>(attrac26={1[8]})
29	8	(AGE=2)&(Involvement=1)=>(attrac26={1[8]})
30	8	(PENTER=2)&(PWORK=1)=>(attrac26={1[8]})
31	8	(STUCOM=2)&(Gender=2)&(PWORK=2)&(Hedonic=1)=>(attrac26={1[8]})
32	8	(STUCOM=2)&(AGE=1)&("CVTER "=1)=>(attrac26={1[8]})
33	8	(STUCOM=2)&(PWORK=1)&("CVTER "=1)=>(attrac26={1[8]})
34	7	(AGE=1)&(FREQ_USE=1)&(PENTER=2)=>(attrac26={1[7]})
35	7	(Hedonic=2)&("CVTER "=2)=>(attrac26={1[7]})
36	7	(FREQ_USE=2)&(Involvement=1)=>(attrac26={1[7]})
37	7	(STUCOM=1)&(Utility=1)&(Hedonic=1)&(Involvement=2)=>(attrac26={1[7]})
38	7	(PENTER=1)&(Utility=1)&(Involvement=2)&("CVTER "=1)=>(attrac26={1[7]})
39	7	(Utility=1)&(Hedonic=1)&(Involvement=2)&("CVTER "=1)=>(attrac26={1[7]})
40	7	(Gender=2)&(PWORK=2)&(Hedonic=1)&(Involvement=2)=>(attrac26={1[7]})
41	7	(Gender=2)&(Hedonic=1)&(Involvement=2)&("CVTER "=1)=>(attrac26={1[7]})
42	7	(AGE=2)&(PWORK=1)=>(attrac26={1[7]})
43	7	(STUCOM=1)&(PCOMM=2)&("CVTER "=2)=>(attrac26={1[7]})
44	7	(PCOMM=2)&(Involvement=1)&("CVTER "=2)=>(attrac26={1[7]})
45	6	(PCOMM=1)&(Hedonic=2)=>(attrac26={1[6]})
46	6	(PWORK=2)&(PCOMM=2)&(Involvement=1)=>(attrac26={1[6]})
47	6	(Gender=1)&(FREQ_USE=1)&("CVTER "=1)=>(attrac26={1[6]})
48	6	(Gender=1)&(Utility=1)&("CVTER "=1)=>(attrac26={1[6]})
49	6	(AGE=1)&(FREQ_USE=1)&(Involvement=2)&("CVTER "=1)=>(attrac26={1[6]})
50	6	(FREQ_USE=1)&(PENTER=1)&(Involvement=2)&("CVTER "=1)=>(attrac26={1[6]})
51	6	(FREQ_USE=1)&(Hedonic=1)&(Involvement=2)&("CVTER "=1)=>(attrac26={1[6]})
52	6	(Gender=2)&(PENTER=1)&(Involvement=2)&("CVTER "=1)=>(attrac26={1[6]})
53	6	(Gender=2)&(FREQ_USE=2)&(PENTER=1)=>(attrac26={1[6]})
54	6	(FREQ_USE=2)&(PENTER=1)&(Utility=1)=>(attrac26={1[6]})
55	6	(PWORK=1)&(Hedonic=2)=>(attrac26={1[6]})
56	5	(STUCOM=1)&(Gender=1)&(Utility=1)&(Involvement=2)=>(attrac26={1[5]})
57	5	(STUCOM=1)&(Involvement=2)&("CVTER "=2)=>(attrac26={1[5]})

Figure 4.5. Rules of IST of website 26—attractiveness.

4.4 "Reduct-Rule Integrator" Software

Software was programmed in visual basic language to integrate all reduct sets and rules that are generated for all ISTs of all 24 websites and/or for different Kanseis. This software is called the “Reduct-Rule Integrator.” By putting all reduct sets generated from all ISTs together, their elements can be compared, and then it is possible to identify those that have the same members. Table 4.8 shows all reduct sets of ISTs of all websites for one specific Kansei, “attractiveness.”

TABLE 4.8 ALL REDUCT SETS OF ISTS FOR ALL WEBSITES FOR KANSEI “ATTRACTIVENESS”

Attractive Website	User Characteristics											
	SC	G	PE	PW	PC	S	I	CVPA	A	FU	H	U
Att-1	1	1	1	1	1	1	1	1				
Att-11	1	1	1	1	1	1	1	1	1	1		
Att-11	1	1	1	1	1	1	1	1			1	
Att-13	1	1	1	1	1	1	1	1			1	
Att-15	1	1	1	1	1	1	1	1				
Att-17	1	1		1	1		1	1	1	1		
Att-17	1	1	1		1	1	1	1	1	1		
Att-17	1	1	1	1	1	1	1	1	1			
Att-17	1	1	1	1		1	1	1	1			1
Att-17	1	1	1	1		1	1	1	1		1	
Att-17	1	1	1	1	1		1	1		1	1	
Att-17	1	1	1		1	1	1	1		1	1	
Att-17	1	1	1	1	1	1	1	1			1	
Att-20	1	1	1	1	1	1	1	1		1	1	
Att-22	1	1	1	1	1	1	1	1			1	
Att-23	1	1	1	1	1		1	1				
Att-23	1	1		1	1		1	1	1		1	
Att-24	1	1	1	1			1	1			1	
Att-24	1	1		1	1		1	1			1	
Att-26	1	1	1	1	1		1	1		1	1	
Att-26	1	1	1	1	1	1	1	1			1	
Att-28	1	1	1	1	1	1		1		1		
Att-28	1	1	1	1	1	1		1	1			1
Att-28	1	1	1	1	1	1	1	1				
…	…	…	…	…	…	…	…	…	…	…	…	…

For example, row four of Table 4.8 indicates that one of the reduct sets of IST corresponding to website no. 13 includes nine attributes: SC (student/community), G (gender), PE (participant’s entertainment), PW (participant’s work), PC (participant’s communication), S (skill), H (hedonic), I (involvement), and CVPA (centrality of visual product aesthetics).

In addition, Table 4.9 shows those websites that have the same reduct set. For example in group 1, there are seven websites. This means that people associated with the same characteristics evaluate websites 11, 13, 17, 22, 26, 30, and 32, with respect to attractiveness.

TABLE 4.9 GROUP OF WEBSITES WITH SAME REDUCT SET

Reduct	SC	G	PE	PW	PC	S	I	CVPA	A	FU	H	U
Group 1												
Att-11	1	1	1	1	1	1	1	1			1	
Att-13	1	1	1	1	1	1	1	1			1	
Att-17	1	1	1	1	1	1	1	1			1	
Att-22	1	1	1	1	1	1	1	1			1	
Att-26	1	1	1	1	1	1	1	1			1	
Att-30	1	1	1	1	1	1	1	1			1	
Att-32	1	1	1	1	1	1	1	1			1	
Group 2												
Att-1	1	1	1	1	1	1	1	1				
Att-15	1	1	1	1	1	1	1	1				
Att-28	1	1	1	1	1	1	1	1				
Att-50	1	1	1	1	1	1	1	1				
Group 3												
Att-20	1	1	1	1	1	1	1	1		1	1	
Att-29	1	1	1	1	1	1	1	1		1	1	
Att-5	1	1	1	1	1	1	1	1		1	1	

Using the “Reduct-Rule Integrator,” it is possible to put all rules of all ISTs of all websites together, as shown in Table 4.10, and then select those that have a larger match (which is used in the proposed approach).

TABLE 4.10 RULES OF ISTS

Kansei	Website No.	Strength	S1	S2	Certainty1	Certainty2	Coverage1	Coverage2	No. of Rules	SC	G	A	FU	PE	PW	PC	S	U	H	I	CVPA
Att	1	0.22		14	0	1	0	0.28	2							2				1	
Att	1	0.21		13	0	1	0	0.26	4		2	1		1							1
Att	1	0.19		12	0	1	0	0.24	2							1				2	
Att	1	0.19		12	0	1	0	0.24	3	2	2	1									
Att	1	0.17		11	0	1	0	0.22	4		2	1	1								1
Att	1	0.17		11	0	1	0	0.22	4		2		1				1				1
Att	1	0.17		11	0	1	0	0.22	4	2	2			1	1						
Att	1	0.17		11	0	1	0	0.22	4	2	2			1			1				
Att	1	0.16		10	0	1	0	0.2	4		2	1		1	2						
Att	1	0.16		10	0	1	0	0.2	4		2	1		1		2					
Att	1	0.16		10	0	1	0	0.2	3						1	2	1				
Att	1	0.16		10	0	1	0	0.2	3						1	2			1		
Att	1	0.14		9	0	1	0	0.18	4		2	1	1		2						
Att	1	0.14		9	0	1	0	0.18	4		2		1							1	1
Att	1	0.14		9	0	1	0	0.18	4		2		1		2	1					
Att	1	0.14		9	0	1	0	0.18	4		2		1			1					1
Att	1	0.14		9	0	1	0	0.18	4		2					1	1				1
Att	1	0.14		9	0	1	0	0.18	3			1			1	2					
Att	1	0.14		9	0	1	0	0.18	4	2	2			1		1					

In Table 4.10, S_1 and S_2 denote the number of supports for each decision class for the specific rule. The number of people who believe that the website is attractive is denoted by S_1, and the number of people who believe that the website is unattractive is denoted S_2. For example, row 2 of Table 4.10 shows that 14 people with PCOMM = 2 and involvement = 1 view website no. 1 as unattractive. In addition, the strength of the rule is calculated by the number of supports of the rule divided by the total number of people, as shown by equation (4.11). Also the certainty and the coverage of the rule for each decision class are calculated using equations (4.7) and (4.9).

Table 4.11 shows common rules that are defined across different websites' rules.

TABLE 4.11 COMMON RULES OF DIFFERENT WEBSITES

Group	Website No.	Match	Strength	1	2	Certainty1	Certainty2	Coverage1	Coverage2	No. of Rules	SC	G	A	FU	PE	PW	PC	S	U	H	I	CVPA
Group 1																						
Att	11	6	0.1		6	0	1	0	0.13	2						1				2		
Att	17	6	0.1		6	0	1	0	0.11	2						1				2		
Att	23	6	0.1		6	0	1	0	0.12	2						1				2		
Att	26	6	0.1	6		1	0	0.11	0	2						1				2		
Att	30	6	0.1	6		1	0	0.12	0	2						1				2		
Att	33	6	0.1	6		1	0	0.11	0	2						1				2		
Att	35	6	0.1	6		1	0	0.11	0	2						1				2		
Att	38	6	0.1	6		1	0	0.11	0	2						1				2		
Att	41	6	0.1	6		1	0	0.11	0	2						1				2		
Att	46	6	0.1	6		1	0	0.11	0	2						1				2		
Att	8	6	0.1		6	0	1	0	0.12	2						1				2		
Group 2																						
Att	11	6	0.1		6	0	1	0	0.13	2							2			2		
Att	13	6	0.1		6	0	1	0	0.15	2							2			2		
Att	17	6	0.1		6	0	1	0	0.11	2							2			2		
Att	24	6	0.1		6	0	1	0	0.1	2							2			2		
Att	30	6	0.1	6		1	0	0.12	0	2							2			2		
Att	33	6	0.1	6		1	0	0.11	0	2							2			2		
Att	38	6	0.1	6		1	0	0.11	0	2							2			2		
Att	41	6	0.1	6		1	0	0.11	0	2							2			2		
Att	46	6	0.1	6		1	0	0.11	0	2							2			2		
Att	8	6	0.1		6	0	1	0	0.12	2							2			2		
Group 3																						
Att	11	9	0.14		9	0	1	0	0.19	2						1					2	
Att	15	9	0.14		9	0	1	0	0.18	2						1					2	
Att	17	9	0.14		9	0	1	0	0.17	2						1					2	
Att	33	9	0.14	9		1	0	0.17	0	2						1					2	
Att	38	9	0.14	9		1	0	0.17	0	2						1					2	
Att	46	9	0.14	9		1	0	0.17	0	2						1					2	
Att	50	9	0.14	9		1	0	0.17	0	2						1					2	
Att	9	9	0.14		9	0	1	0	0.15	2						1					2	

For example, group 1 demonstrates 11 similar rules for websites 11, 17, 23, 26, 30, 33, 35, 38, 41, 46, and 8. This rule implies the fact that six users who do not use the Internet for work purposes very often and with low involvement view websites 26, 30, 33, 35, 38, 41, and 46 as attractive and websites 11, 17, 23, and 8 as unattractive.

In this chapter, the possibility of applying the proposed approach in a real-world scenario was verified. In this study, an attempt was made to define a user feelings-oriented webpage development project in which corresponding data were obtained from the software usability research laboratory in the Department of Psychology at Wichita State University. The relevant notions of rough set theory along with their counterpart concepts in this problem were described. Also, the rough set theory software (RSES 2) with its relevant outputs, which will be used in the proposed approach, was demonstrated. Now the proposed approach will be applied, and customer grouping and design rules will be extracted. These rules along with some other results are described and discussed in Chapter 5.

CHAPTER 5

RESULTS

This chapter discusses the output of each stage of the developed method, which henceforth will be called rough set-based Kansei Engineering. The four major outputs of each stage are discussed below.

Stage I:

- Identify influential user characteristics for one specific Kansei.
- Identify influential user characteristics affecting multiple Kanseis.
- Develop customer segmentation rules for one specific Kansei.
- Develop customer segmentation rules for multiple Kanseis.

Stage II:

- Identify influential product features for one Kansei of a specific cluster(s) of customer groups.
- Identify influential product features for multiple Kanseis of a cluster(s) of customer groups.
- Generate design rules for one specific Kansei for one specific cluster of customer groups.
- Generate design rules for multiple Kanseis for one specific cluster of customer groups.

RSBKE can also determine the group of users who have been identified through a specific set of product feature values. In the remainder of the chapter, the results of implementing RSBKE are compared with three other situations. The first comparison involves random data to determine if the pattern found in real-world data is stronger than that in random data. The second comparison is with the results of the statistical approach performed in the Phillips (2007) study. The third comparison is between RSBKE, based on extracting consistent discernible classes of users (common rules strategy), and the approach where classes of users are extracted based on rules with high coverage (max strength strategy). This comparison will examine and determine which approach provides more useful and reliable information for designers.

The chapter ends with other useful results of the implementation, results that can identify users (here called "lead" users) associated with the same set of characteristics for different judgment situations. Furthermore, the approach can enhance the Kano model by providing information regarding interaction between customer satisfaction elements when there are multiple expectations.

5.1 Stage I

5.1.1 Identify Influential User Characteristics for One Specific Kansei

For each website, many reducts were generated. Each reduct contained a set of user characteristics that impacted the perception of users for that specific website with respect to a specific Kansei. Each group of websites (group 1, group 2, etc.) in Table 5.1 shows the necessary (identified by "1") and redundant (identified by blank) knowledge or characteristics of users that influenced user perception regarding that specific set of websites. For example, Group 2 in 5.1 contains common reduct sets for websites 11, 13, 17, 22, 26, 30, and 32. In fact, each group in this table represents the set of characteristics for users that are consistent in their evaluation of a specific set of websites with respect to "attractiveness." In other words, group 1 in 5.1 depicts the essential part of the available knowledge about users {STUCOM, gender, PENTER, PWORK, PCOMM, skill, involvement, CVPA, hedonic}, which is associated with their "attractiveness" perception of websites 11, 13, 17, 22, 26, 30, and 32. This implies that knowledge {AGE, FREQ_USE, hedonic} is redundant knowledge for all users who evaluate the above websites.

TABLE 5.1 COMMON REDUCT SETS—ATTRACTIVENESS

Reduct	User characteristics											
	STUCOM	Gender	PENTER	PWORK	PCOMM	Skill	Involvement	CVPA	Age	FREQ_USE	Hedonic	Utility
Group 1												
Att-11	1	1	1	1	1	1	1	1			1	
Att-13	1	1	1	1	1	1	1	1			1	
Att-17	1	1	1	1	1	1	1	1			1	
Att-22	1	1	1	1	1	1	1	1			1	
Att-26	1	1	1	1	1	1	1	1			1	
Att-30	1	1	1	1	1	1	1	1			1	
Att-32	1	1	1	1	1	1	1	1			1	
Group 2												
Att-1	1	1	1	1	1	1	1	1				
Att-15	1	1	1	1	1	1	1	1				
Att-28	1	1	1	1	1	1	1	1				
Att-50	1	1	1	1	1	1	1	1				
Group 3												
Att-20	1	1	1	1	1	1	1	1		1	1	
Att-29	1	1	1	1	1	1	1	1		1	1	
Att-5	1	1	1	1	1	1	1	1		1	1	
Group 4												
Att-17	1	1	1	1	1		1	1		1	1	
Att-26	1	1	1	1	1		1	1		1	1	

It is interesting that while the set of user characteristics {STUCOM, gender, PENTER, PWORK, PCOMM, skill, involvement, CVPA, hedonic} is associated with users' evaluation of websites 11, 13, 17, 22, 26, 30, and 32 with respect to "attractiveness," the same set of characteristics, except hedonic, affects users' evaluation of websites 1, 15, 28, and 50. It should be noted that there may be single or multiple reduct set(s) for each website. The same reduct grouping for other Kanseis as "interesting" and "imaginative" are found in Appendix B (Table B.1 and Table B.2).

5.1.2 User Characteristics Associated with Multiple Kanseis

Table 5.2 illustrates the set of characteristics that impacted (either positively or negatively) the following user perceptions in assessing websites with respect to all Kanseis: "attractiveness," "interesting," and "imaginative." For example, group 1 in Table 5.2 shows the set of user characteristics {STUCOM, gender, PENTER, PWORK, PCOMM, skill, involvement, CVPA, hedonic} that influence users' assessments of websites 11, 13, 17, 22, 26, 30, and 32 with respect to "attractiveness"; assessments of websites 13, 28, 30, 32, and 5 with respect to "imaginative"; and assessments of websites 20, 22, 30, 32, and 41 with respect to "interesting."

TABLE 5.2 COMMON REDUCT SETS—ALL KANSEIS

Reduct	User characteristics											
	STUCOM	Gender	PENTER	PWORK	PCOMM	Skill	Involvement	CVPA	Age	FREQ_USE	Hedonic	Utility
Group 1												
Att-11	1	1	1	1	1	1	1	1			1	
Att-13	1	1	1	1	1	1	1	1			1	
Att-17	1	1	1	1	1	1	1	1			1	
Att-22	1	1	1	1	1	1	1	1			1	
Att-26	1	1	1	1	1	1	1	1			1	
Att-30	1	1	1	1	1	1	1	1			1	
Att-32	1	1	1	1	1	1	1	1			1	
Im-13	1	1	1	1	1	1	1	1			1	
Im-28	1	1	1	1	1	1	1	1			1	
Im-30	1	1	1	1	1	1	1	1			1	
Im-32	1	1	1	1	1	1	1	1			1	
Int-20	1	1	1	1	1	1	1	1			1	
Int-22	1	1	1	1	1	1	1	1			1	
Int-30	1	1	1	1	1	1	1	1			1	
Int-32	1	1	1	1	1	1	1	1			1	
Int-41	1	1	1	1	1	1	1	1			1	
Group 2												
Att-20	1	1	1	1	1	1	1	1		1	1	
Att-29	1	1	1	1	1	1	1	1		1	1	
Att-5	1	1	1	1	1	1	1	1		1	1	
Im-20	1	1	1	1	1	1	1	1		1	1	
Im-29	1	1	1	1	1	1	1	1		1	1	
Im-46	1	1	1	1	1	1	1	1		1	1	
Int-1	1	1	1	1	1	1	1	1		1	1	
Int-15	1	1	1	1	1	1	1	1		1	1	
Int-23	1	1	1	1	1	1	1	1		1	1	
Int-8	1	1	1	1	1	1	1	1		1	1	
Group 3												
Att-17	1	1	1		1	1	1	1		1	1	
Im-30	1	1	1		1	1	1	1		1	1	
Im-33	1	1	1		1	1	1	1		1	1	
Im-5	1	1	1		1	1	1	1		1	1	
Im-8	1	1	1		1	1	1	1		1	1	
Int-29	1	1	1		1	1	1	1		1	1	
Int-30	1	1	1		1	1	1	1		1	1	
Int-5	1	1	1		1	1	1	1		1	1	

In Table 5.2, "1" indicates existence of a characteristic in the reduct set; otherwise, the space is blank. It should be noted that the reduct sets do not indicate the value of characteristics that identify a group of users.

Furthermore, Table 5.2 represents the set of websites that are assessed by users who have the same set of characteristics. For instance, group 2 in Table 5.2 indicates that users who have the same set of characteristics {STUCOM, gender, PENTER, PWORK, PCOMM, skill, involvement, CVPA, FREQ_USE, hedonic} assess websites 20, 29, and 5 with respect to "attractiveness"; websites 20, 29, and 46 with respect to "imaginative"; and websites 1, 15, 23, and 8 with respect to "interesting."

5.1.3 Customer Segmentation Rules for Specific Kansei

Table 5.3 shows common rules across different websites for "attractiveness." Similar tables for "interesting" and "imaginative" are found in Appendix B (Table B.3 and Table B.4.)

Using RSBKE to obtain the set of common rules, first, rules were generated for all websites for one Kansei. Then, using the "Reduct-Rule Integrator," rules with strength greater than 10%, were selected. Rules with the same condition and decision attributes were then grouped together.

Table 5.3 shows the match, strength, certainty, and coverage for each class of decisions (attractive, not attractive) and number of characteristics involved in each rule. For example, for group 1, the match, strength, certainty, and coverage of the rule corresponding to website 11 are 6, 0.1, 1, and 0.13, respectively. Note that here, corresponding certainty and coverage of attractiveness rule is 0. The corresponding formulas are found in section 4.2.30.3, equation (4.9).

Each group of rules relates the values of a specific set of user characteristics to classes of user assessment for a set of websites with respect to one Kansei. For example, group 1 in Table

5.3, shows that people who do not usually use the Internet for work purposes (low use of Internet for work, Pwork = 1) and also score low on hedonic (less focus on activities that are exploratory and entertainment-derived, hedonic = 2) view websites 26, 30, 33, 35, 38, 41, and 46 as attractive, and websites 11, 17, 23, and 8 as not attractive. Website features that are associated with these perceptions will be identified in the following sections.

TABLE 5.3 USER GROUPING RULES FOR KANSEI "ATTRACTIVENESS"

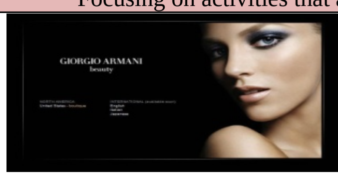
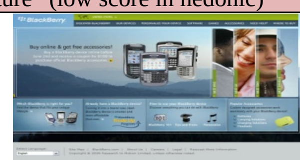
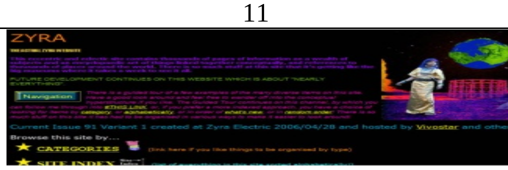
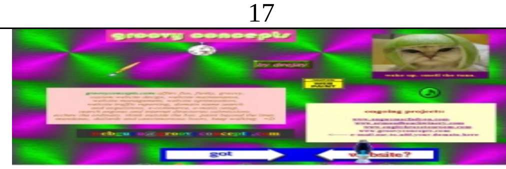
Group	Websites	Match	Strength	Attractive	Not Attractive	Certainty1	Certainty2	Coverage1	Coverage2	No. of Characteristics	User characteristics									
											STUCOM	Gender	Age	FREQ_USE	PENTER	PWORK	PCOMM	Skill	Utility	Hedonic
Group 1 (AttG1)																				
Att	11	6	0.1		6	0	1	0	0.13	2						1				2
Att	17	6	0.1		6	0	1	0	0.11	2						1				2
Att	23	6	0.1		6	0	1	0	0.12	2						1				2
Att	26	6	0.1	6		1	0	0.11	0	2						1				2
Att	30	6	0.1	6		1	0	0.12	0	2						1				2
Att	33	6	0.1	6		1	0	0.11	0	2						1				2
Att	35	6	0.1	6		1	0	0.11	0	2						1				2
Att	38	6	0.1	6		1	0	0.11	0	2						1				2
Att	41	6	0.1	6		1	0	0.11	0	2						1				2
Att	46	6	0.1	6		1	0	0.11	0	2						1				2
Att	8	6	0.1		6	0	1	0	0.12	2						1				2
Group 2 (AttG2)																				
Att	11	6	0.1		6	0	1	0	0.13	2							2			2
Att	13	6	0.1		6	0	1	0	0.15	2							2			2
Att	17	6	0.1		6	0	1	0	0.11	2							2			2
Att	24	6	0.1		6	0	1	0	0.1	2							2			2
Att	30	6	0.1	6		1	0	0.12	0	2							2			2
Att	33	6	0.1	6		1	0	0.11	0	2							2			2
Att	38	6	0.1	6		1	0	0.11	0	2							2			2
Att	41	6	0.1	6		1	0	0.11	0	2							2			2
Att	46	6	0.1	6		1	0	0.11	0	2							2			2
Att	8	6	0.1		6	0	1	0	0.12	2							2			2

Also, users with the same hedonic score as group 1 (hedonic = 2) who use the Internet for communication more frequently (PCOMM = 2), identified as group 2, view websites 11, 13, 17, 24, and 8 as not attractive, and websites 30, 33, 38, 41, and 46 as attractive. Many other groups can be identified, but these are not shown in Table 5.3.

Another analysis of the information in, Table 5.3 identifies the different sets of websites formed by different sets of user characteristics. Changing the user characteristics in various combinations led to a different formation of website groups and a different set of website features obtained for stage II. For example, the only difference between groups 1 and 2 is the type of computer usage (for communication or work). This is enough of a difference to form two different sets of websites, which are in turn, the basis for obtaining a different set of influential website features.

Table 5.4 shows "if-then rules" or, in other words, the relationship between the characteristics "low use of Internet for work" and "focusing on activities that are exploratory and entertainment-derived in nature." This table includes images of the websites.

TABLE 5.4 RELATIONSHIP BETWEEN "LOW USES OF INTERNET FOR WORK" AND "HEDONIC" AND USER WEBSITE ASSESSMENTS (IF-THEN RULES OF GROUP 1 OF TABLE 5.3)

<table>
<tr><th rowspan="2">Kansei</th><th colspan="6">User Characteristics</th></tr>
<tr><th colspan="6">"Low use of Internet for work"
"Focusing on activities that are exploratory and entertainment-derived in nature" (low score in hedonic)</th></tr>
<tr><td rowspan="2">Attractive</td><td colspan="2">
26</td><td colspan="2">
30</td><td colspan="2">
33</td></tr>
<tr><td colspan="2">
38</td><td colspan="2">
41</td><td colspan="2">
46</td></tr>
<tr><td rowspan="2">Not Attractive</td><td colspan="3">
11</td><td colspan="3">
17</td></tr>
<tr><td colspan="3">
23</td><td colspan="3">
8</td></tr>
</table>

5.1.4 Market Segmentation Rules for Multiple Kanseis

Customer grouping rules can be extracted for all Kanseis. As in the previous section, group 1 in Table 5.5 indicates the characteristics of people who perceive websites 11, 17, 23, 26, 30, 33, 35, 38, 41,46, and 8 as attractive; websites 11, 13, 17, 24, 34, 35, 38, 46, and 9 as imaginative; and websites 11, 23, 30, and 35 as interesting.

TABLE 5.5 USERS' GROUPING RULES FOR ALL KANSEIS

Group	Website No.	Match	Strength	Fulfill Kansei	Not-Fulfill Kansei	Certainty1	Certainty2	Coverage1	Coverage2	No. of Characteristics	User characteristics	
											PWORK	Hedonic
Att	11	6	0.1		6	0	1	0	0.13	2	1	2
Att	17	6	0.1		6	0	1	0	0.11	2	1	2
Att	23	6	0.1		6	0	1	0	0.12	2	1	2
Att	26	6	0.1	6		1	0	0.11	0	2	1	2
Att	30	6	0.1	6		1	0	0.12	0	2	1	2
Att	33	6	0.1	6		1	0	0.11	0	2	1	2
Att	35	6	0.1	6		1	0	0.11	0	2	1	2
Att	38	6	0.1	6		1	0	0.11	0	2	1	2
Att	41	6	0.1	6		1	0	0.11	0	2	1	2
Att	46	6	0.1	6		1	0	0.11	0	2	1	2
Att	8	6	0.1		6	0	1	0	0.12	2	1	2
Im	11	6	0.1		6	0	1	0	0.12	2	1	2
Im	13	6	0.1		6	0	1	0	0.13	2	1	2
Im	17	6	0.1		6	0	1	0	0.11	2	1	2
Im	24	6	0.1		6	0	1	0	0.11	2	1	2
Im	34	6	0.1	6		1	0	0.11	0	2	1	2
Im	35	6	0.1	6		1	0	0.11	0	2	1	2
Im	38	6	0.1	6		1	0	0.11	0	2	1	2
Im	46	6	0.1	6		1	0	0.13	0	2	1	2
Im	9	6	0.1		6	0	1	0	0.13	2	1	2
Int	11	6	0.1		6	0	1	0	0.12	2	1	2
Int	23	6	0.1		6	0	1	0	0.14	2	1	2
Int	30	6	0.1	6		1	0	0.14	0	2	1	2
Int	35	6	0.1	6		1	0	0.11	0	2	1	2

Group 1 In Table 5.5 is defined by a rule that shows a relationship between low frequency of Internet use for work purposes (characteristic number 6), low levels of activities that are more exploratory and entertainment-derived in nature (characteristic number 10), and their perception (attractiveness, imaginative, and interesting) of some websites. This implies that

these two characteristics relate to multiple perceptions. Many other such groups have been identified but are not shown in Table 5.5.

5.2 Stage II

This stage develops the relationship between website features and the perception of a specific group of users defined by their characteristics. At the end of this section, those website features that influence perceptions of clusters of user groups will be identified. The rules generated in this section clarify the effect of each website's technical features on user assessments.

5.2.1 Influential Product Features for One Kansei of User Cluster(s)

Table 5.6 shows the pattern of website design features that impact the perception (attractiveness) of different groups of users. Product features associated with interesting and imaginative are found in Appendix B (Table B.5 and Table B.6).

This set of different groups of users is called AttG1…AttG5, specified in Table 5.6 as clusters of groups of users. For example, cluster 1 (also called cluster A 1, 2, 3, 4, 5) contains AttG1…AttG5, each defined by different characteristics (Table 5.7)

Table 5.6 also identifies design rules for different clusters. For example "Image Quality Resolution" is one of the rules for cluster 1 and affects the perception of attractiveness of users in the cluster via specific sets of websites (identified in Table 5.8 to Table 5.11). In addition "font type" and "line length" both affect the judgment of the same cluster of groups of users with respect to attractiveness. In other words, there is a relationship between "line length" and "font type" of the same set of websites and user characteristics of Table 5.7.

Referring to Table 5.3, AttG1 to AttG5 includes the sets of websites in which each set is evaluated by the same user group as either attractive or not attractive. Table 5.7 shows the characteristics of user groups AttG1 to AttG5, or cluster 1.

TABLE 5.6 COMMON REDUCT SET OF WEB FEATURES—DIFFERENT CLUSTERS OF USER GROUPS——ATTRACTIVENESS

Reduct	Web features												
	Font Color	Image Size	Font Size	Line Length	Picture	Display Color	Font Type	Image Content	Page Layout Color Combination	Font Contrast	Display Background Graphic	Page Layout Density	Image Quality Resolution
Cluster 1													
AttG1				1			1						
AttG2				1			1						
AttG3				1			1						
AttG4				1			1						
AttG5				1			1						
Cluster 1													
AttG1													1
AttG2													1
AttG3													1
AttG4													1
AttG5													1
Cluster 3													
AttG1	1	1											
AttG2	1	1											
AttG5	1	1											
Cluster 4													
AttG1		1	1	1									
AttG2		1	1	1									
AttG5		1	1	1									
Cluster 5													
AttG1			1	1							1		
AttG2			1	1							1		
AttG5			1	1							1		
Cluster 6													
AttG1	1			1							1		
AttG2	1			1							1		
AttG5	1			1							1		

TABLE 5.7 CHARACTERISTICS OF EACH GROUP FORMED IN STAGE I (TABLE 5.3—ATTRACTIVENESS) TO BE USED IN STAGE II

Group	User Characteristics							Corresponding Websites
	STUCOM	Age	PWORK	PCOMM	Hedonic	Involvement	CVPA	
AttG1			1 = Low		2 = Low			11, 17, 23, 26, 30, 33, 35, 38, 41, 46, 8
AttG2				2 = High	2 = Low			11, 13, 17, 24, 30, 33, 38, 41, 46, 8
AttG3			1 = Low			2 = Low		11, 15, 17, 33, 38, 46, 50, 9
AttG4	1 = Student		2 = High				2 = High	15, 17, 23, 24, 29, 35, 41, 8
AttG5		2 = Above 40	1 = Low					24, 26, 38, 41, 46, 8, 9

Table 5.8 to Table 5.11 present the detailed information summarized in Table 5.3 and show the characteristics of selected groups of users, their corresponding website evaluations, and website images (corresponding information for AttG1 is already shown in Table 5.4).

TABLE 5.8 RELATIONSHIPS BETWEEN GROUPS OF CHARACTERISTICS AND USER EVALUATIONS—ATTG2

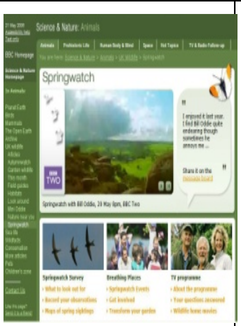
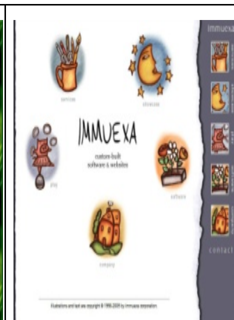
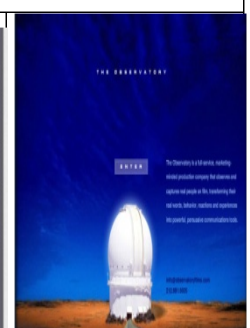
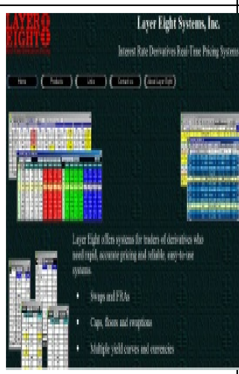
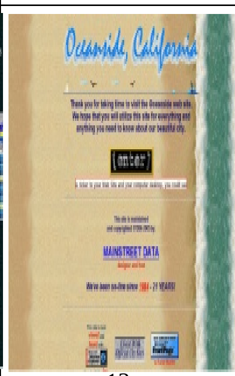
Kansei	User Characteristic				
	High use of Internet for communication purposes (high in communication score) Users focusing less on exploratory and entertainment-derived activities (low score in hedonic)				
Attractive	 30	 33	 38	 41	 46
Not Attractive	 11	 13	 17	 24	 8

TABLE 5.9 RELATIONSHIPS BETWEEN GROUPS OF CHARACTERISTICS AND USER EVALUATIONS—ATTG3

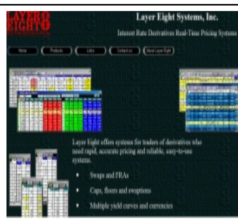
Kansei	User Characteristic			
	Low use of Internet for work (P-work = 1) Users immersed while completing tasks on Internet (low score in involvement = 2)			
Attractive	 33	 38	 46	 50
Not Attractive	 11	 15	 17	 9

TABLE 5.10 RELATIONSHIPS BETWEEN GROUPS OF CHARACTERISTICS AND USER EVALUATIONS—ATTG4

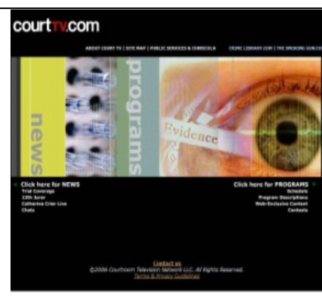
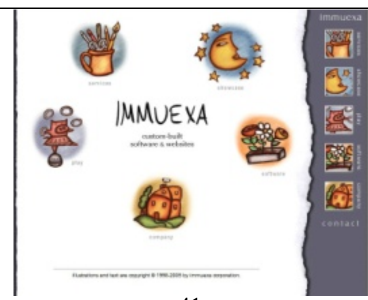
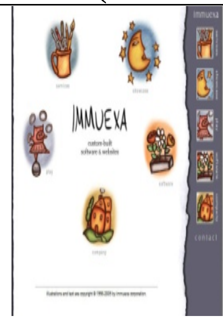
Kansei	User Characteristic				
	High use of Internet for work (P-work = 2) High, aesthetic importance to users in their purchasing decisions (CVTR = 2)				
Attractive	 29	 35	 41		
Not Attractive	 15	 17	 23	 24	 8

TABLE 5.11 RELATIONSHIPS BETWEEN GROUPS OF CHARACTERISTICS AND USER EVALUATIONS—ATTG5

<table>
<tr><td rowspan="2">Kansei</td><td colspan="4">User Characteristic</td></tr>
<tr><td colspan="4">Age over 40 (age = 2)
Low use of Internet for work (P-work = 1)</td></tr>
<tr><td>Attractive</td>
<td>

26</td>
<td>

38</td>
<td>

41</td>
<td>
46</td></tr>
<tr><td>Not Attractive</td>
<td>

24</td>
<td>

8</td>
<td colspan="2">

9</td></tr>
</table>

5.2.2 Influential Product Features for Multiple Kansei for Cluster(s) of Customer Groups

Table 5.12 shows the common reduct sets of website features that influence perceptions (attractiveness, interesting, and imaginative) of groups of users characteristics (for example, AttG1 represents the group of characteristics “P-work = 1” and “hedonic = 2”). Table 5.12 also shows that “image quality resolution” affects the attractiveness perception of user characteristic groups AttG1, AttG2, AttG3, AttG4, and AttG5; imaginativeness groups ImG1, ImG3, ImG4, and ImG5; and interest group IntG1. Moreover, “font type” and “line length” both influence AttG1 to AttG5, ImG2, ImG4, and ImG5.

TABLE 5.12 COMMON REDUCT SETS OF WEB FEATURES FOR CLUSTERS OF GROUP(S) FOR ALL KANSEIS

Group of Characteristics	Web features												
	Font Color	Image Size	Font Size	Line Length	Picture	Display Color	Font Type	Image Content	Page Layout Color Combination	Font Contrast	Display Background Graphic	Page Layout Density	Image Quality Resolution
Cluster 1													
AttG1													1
AttG2													1
AttG3													1
AttG4													1
AttG5													1
ImG1													1
ImG3													1
ImG4													1
ImG5													1
IntG1													1
Cluster 2													
AttG1				1			1						
AttG2				1			1						
AttG3				1			1						
AttG4				1			1						
AttG5				1			1						
ImG2				1			1						
ImG4				1			1						
ImG5				1			1						
Cluster 3													
AttG1	1	1											
AttG2	1	1											
AttG5	1	1											
IntG1	1	1											
Cluster 4													
AttG1								1			1	1	
AttG2								1			1	1	
AttG5								1			1	1	
ImG1								1			1	1	
Cluster 5													
AttG3		1											
ImG1		1											
ImG2		1											
ImG5		1											

5.2.3 Design Rules for One Kansei for One Cluster of Groups

After finding the most influential website features, the question becomes which values of these features have a positive or negative impression on user perception. The design rules that are generated in this stage answer this question. Table 5.13 shows design rules for “attractiveness.” For “interesting” and “imaginative” see Appendix B (Table B.7 and Table B.8). There is one rule for each cluster (group of user characteristics). For example, cluster 1 indicates that if image quality resolution is high (“image quality resolution = 2”), then people who are defined in AttG1, AttG2, AttG3, AttG4, and AttG5 (refer to Table 5.7) view the listed websites as attractive. For the same group of characteristics and websites included in AttG1, AttG2, AttG3, AttG4, and AttG5, there is another rule that relates the value 1 of “font type” (standard font type) and “line length” (short line length) to the impression of the group of user characteristics (defined in Table 5.7). Therefore, as a designer, either the first or the second rule or both can be adopted to improve the attractiveness of a website for the specified user groups. It is interesting that the system generates another rule (for cluster 3 in Table 5.13) when paired with the first rule and implies that if the “image quality resolution” is low, then those websites are attractive for that set of user characteristics. Table 5.13 is summarized in Table 5.14.

In summary, high image quality resolution, standard font type, and short line length are associated with websites specified in Table 5.8 to Table 5.11, as being more attractive for users who possess the characteristics specified in Table 5.7.

TABLE 5.13 DESIGN RULES FOR DIFFERENT USERS GROUPS AND FOR ATTRACTIVENESS

Cluster (C)	Match	Strength	Attractive	Not Attractive	Certainty1	Certainty2	Coverage1	Coverage2	No. of Attributes	Web features												
										Font Size	Font Color	Font Type	Font Contrast	Image Size	Image Quality Resolution	Image Content	Page Layout Density	Page Layout Color Combination	Picture	Line Length	Display Color	Display Background Graphic
C1																						
AttG1	7	0.64	7		1	0	1	0	1						2							
AttG2	5	0.5	5		1	0	1	0	1						2							
AttG3	4	0.5	4		1	0	1	0	1						2							
AttG4	3	0.38	3		1	0	1	0	1						2							
AttG5	4	0.57	4		1	0	1	0	1						2							
C2																						
AttG1	7	0.64	7		1	0	1	0	2			1								1		
AttG2	5	0.5	5		1	0	1	0	2			1								1		
AttG3	4	0.5	4		1	0	1	0	2			1								1		
AttG4	3	0.38	3		1	0	1	0	2			1								1		
AttG5	4	0.57	4		1	0	1	0	2			1								1		
C3																						
AttG1	4	0.36		4	0	1	0	1	1						1							
AttG2	5	0.5		5	0	1	0	1	1						1							
AttG3	4	0.5		4	0	1	0	1	1						1							
AttG4	5	0.63		5	0	1	0	1	1						1							
AttG5	3	0.43		3	0	1	0	1	1						1							
C4																						
AttG1	3	0.27		3	0	1	0	0.75	1											2		
AttG2	4	0.4		4	0	1	0	0.8	1											2		
AttG3	3	0.38		3	0	1	0	0.75	1											2		
AttG4	3	0.38		3	0	1	0	0.6	1											2		
AttG5	2	0.29		2	0	1	0	0.67	1											2		
C5																						
AttG1	4	0.36	4		1	0	0.57	0	2			1				1						
AttG2	3	0.3	3		1	0	0.6	0	2			1				1						
AttG3	3	0.38	3		1	0	0.75	0	2			1				1						
AttG5	2	0.29	2		1	0	0.5	0	2			1				1						
C6																						
AttG1	4	0.36	4		1	0	0.57	0	2							1						2
AttG2	3	0.3	3		1	0	0.6	0	2							1						2
AttG3	3	0.38	3		1	0	0.75	0	2							1						2
AttG5	2	0.29	2		1	0	0.5	0	2							1						2

TABLE 5.14 SUMMARY OF TABLE 5.13—DESIGN RULES FOR DIFFERENT USERS GROUPS (ATTG1, ATTG2, ATTG3, ATTG4, ATTG5) AND FOR ATTRACTIVENESS

Rule	Font Size	Font Color	Font Type	Font Contrast	Image Size	Image Quality Resolution	Image Content	Page Layout Density	Page Layout Color Combination	Picture	Line Length	Display Color	Display Background Graphic
1. Attractive						High							
2. Attractive			Standard								Short		
3. Attractive						Low							
4. Attractive											Long		
5. Attractive			Standard				Natural						
6. Attractive							Natural						No Background Graphic

Design rules associated with the impression “interesting” for groups with characteristics specified in Table 5.15 are shown in Table 5.16. Table 5.15 describes two clusters of user groups. Cluster 1 represents users who employ the Internet mostly for work and are immersed while completing Internet tasks. For this group, all rules in Table 5.16 are applied. Also, all of these rules are applied for group 2, except rule no. 2, in which case low “page layout density” is not interesting for group 2, while it is interesting for group 1.

TABLE 5.15 TWO SAMPLES OF GROUP CHARACTERISTICS WHERE “INTERESTING” PERCEPTION IS INFLUENCED BY TABLE 5.14’S RULES

Cluster	Website No.	STUCOM	PWORK	PCOMM	Hedonic	Involvement
Cluster 1 (IntG1)						
Int	11	Community	High			High
Int	17	Community	High			High
Int	22	Community	High			High
Int	24	Community	High			High
Int	41	Community	High			High
Int	46	Community	High			High
Int	50	Community	High			High
Cluster 2 (IntG2)						
Int	11			High	Low	
Int	13			High	Low	
Int	24			High	Low	
Int	29			High	Low	
Int	30			High	Low	
Int	5			High	Low	

TABLE 5.16 WEBSITE DESIGN RULES ASSOCIATED WITH INTERESTING WEBSITES FOR USERS CHARACTERIZED IN TABLE 5.15

Rule	Interesting	Font Size	Font Color	Font Type	Font Contrast	Image Size	Image Quality Resolution	Image Content	Page Layout Density	Page Layout Color Combination	Picture	Line length	Display Color	Display Background Graphic
1	Yes						High							
2	G1/Y, G2/N								Low					
3	Yes											Short		
4	Yes									Good			Colorful	
5	Yes							Natural						
6	Yes	Small								Good				
7	No											Long		
8	Yes		Black											
9	No										No Picture			
10	Yes										Multiple		Colorful	
11	No												Black	

It is interesting that low “page layout density” (rule no. 2 in Table 5.16) associated with websites 11, 17, 22, 24, 41, 46, and 50 is interesting for non-student users who often use the Internet for work and are highly immersed while completing tasks on the Internet (Cluster 1, Table 5.15). On the other hand, “page layout density” associated with websites 11, 13, 24, 29, 30, and 5 is not interesting for users who use the Internet mostly for communication and low focus on activities that are more exploratory and entertainment-derived in nature (Cluster 2, Table 5.15). In other words, the website with less content is more attractive for users who use the Internet for work with high involvement, than users who use the Internet for communication purposes and who are less interested in the exploratory and entertainment aspects of the Internet. It should be noted that the “match” of this rule for IntG2 is only 1 (website 13, Figure 5.1), which means that this rule applies for one website out of six. In other words, in only one website (website 13), if the “page layout density” is low, then the website is not attractive for group 2.

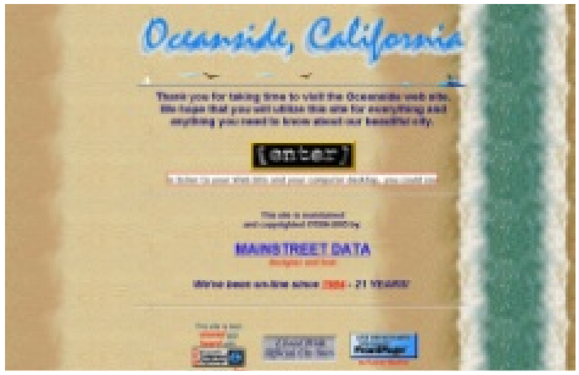


Figure 5.1. Website 13: If "page layout density" of this website is low, then website is not attractive for group.

Also, rules can be extracted for other groups, such as AttG1, AttG2, AttG3, and AttG5, as described in Table 5.17.

TABLE 5.17 WEBSITE DESIGN RULES ASSOCIATED WITH **"**ATTRACTIVE**"** FOR USER GROUPS ATTG1, ATTG2, ATTG3, AND ATTG5

Rule	Attractive	Font Size	Font Color	Font Type	Font Contrast	Image Size	Image Quality Resolution	Image Content	Page Layout Density	Page layout Color Combination	Picture	Line Length	Display Color	Display Background Graphic
1	Yes			Standard				Natural						
2	Yes							Natural						No Background Graphic
3	Yes					Large								
4	Yes								Low					
5	No									Bad				

Furthermore, some rules for AttG1 and AttG2 associated with "attractive" are shown in Table 5.18, and some associated with "not attractive" are shown in Table 5.19. AttG1 refers to the characteristics of users with low hedonic scores who seldom use the Internet for work. AttG2 refers to the characteristics of users with low hedonic scores and who mostly use the Internet for communication.

TABLE 5.18 WEBSITE DESIGN RULES ASSOCIATED WITH “ATTRACTIVE” FOR USER GROUPS ATTG1 AND ATTG2

Rule	Website Features
1	High Image Quality Resolution
2	Standard Font Type, Short Line Length
3	Standard Font Type, Natural Image Content
4	Natural Image Content No Display Background Graphic
5	Large Image Size
6 (only for AttG1)	Low Page Layout Density
7	Good Page Layout Color Combination Short Line Length
8	Short Line Length No Display Background Graphic
9	Good Page Layout Color Combination Colorful Display Color
10	Not Font Contrast, Short Line Length
11	Natural Image Content Good Page Layout Color Combination
12	Standard Font Type Multiple Picture Colorful Display Color
13	Other (Not Black)Font Color, Small Image Size
14	Small Image Size, High Page Layout Density
15	Not Font Contrast, Natural Image Content
16	Large Font Size, Short Line Length
17	Not Natural Image Content Short Line Length
18	Small Image Size Good Page Layout Color Combination
19	Multiple Picture Colorful Display Color No Display Background Graphic
20	Large Font Size Multiple Picture Colorful Display Color
21 (only for AttG1)	Small Font Size, Standard Font Type
22 (only for AttG2)	Medium Image Size
23 (only for AttG2)	Not Natural Image Content Multiple Picture Colorful Display Color
24 (only for AttG2)	Low Page Layout Density Good Page Layout Color Combination
25 (only for AttG2)	Low Page Layout Density Short Line Length
26 (only for AttG2)	Low Page Layout Density No Display Background Graphic

TABLE 5.19 WEBSITE DESIGN RULES ASSOCIATED WITH “NOT ATTRACTIVE” FOR USER GROUPS ATTG1 AND ATTG2

Rule	Website Features
1	Low Image Quality Resolution
2	Long Line Length
3	Bad Page Layout Color Combination
4	Font Color Other(Not Black) Small Image Size
5	Small Image Size High Page Layout Density
6	No Font Contrast Small Image Size
7	Not Natural Image Content High Page Layout Density No Display Background Graphic
8 (only for AttG1)	Font Color Other (Not Black) Not Natural Image Content High Page Layout Density
9 (only for AttG2)	Mixture Font Contrast Small Image Size
10 (only for AttG2)	No Font Contrast Not Natural Image Content No Display Background Graphic

TABLE 5.20 CHARACTERISTICS OF USER GROUPS ATTG1 AND ATTG2

Group	STUCOM	Age	PWORK	PCOMM	Hedonic	Involvement	CVPA	Corresponding Websites
AttG1			1 = Low		2 = Low			11, 17, 23, 26, 30, 33, 35, 38, 41, 46, 8
AttG2				2 = High	2 = Low			11, 13, 17, 24, 30, 33, 38, 41, 46, 8

5.2.4 Design Rules for Multiple Kanseis of One Group of Users

Table 5.21 shows some of the common rules for all Kanseis and different groups of user characteristics along with corresponding performance measures. The complete list of rules is found in Appendix B (Table B.9). These rules are associated with “attractive,” “interesting,” and “imaginative” to describe the website for different sets of user characteristics.

TABLE 5.21 DESIGN RULES FOR DIFFERENT USER GROUPS AND FOR MULTIPLE KANSEIS

Cluster (C)	Match	Strength	Attractive	Not Attractive	Certainty1	Certainty2	Coverage1	Coverage2	No. of Attributes	Web features										
										Font Size	Font Color	Font Type	Font Contrast	Image Size	Image Quality Resolution	Image Content	Page Layout Density	Page layout Color Combination	Picture	Line Length
Cluster 1																				
AttG1	7	0.64	7		1	0	1	0	1						2					
AttG2	5	0.5	5		1	0	1	0	1						2					
AttG3	4	0.5	4		1	0	1	0	1						2					
AttG4	3	0.38	3		1	0	1	0	1						2					
AttG5	4	0.57	4		1	0	1	0	1						2					
ImG1	4	0.44	4		1	0	1	0	1						2					
ImG2	4	0.5	4		1	0	0.8	0	1						2					
ImG3	3	0.43	3		1	0	1	0	1						2					
ImG4	4	0.57	4		1	0	1	0	1						2					
ImG5	3	0.5	3		1	0	1	0	1						2					
IntG1	4	0.57	4		1	0	1	0	1						2					
IntG2	2	0.33	2		1	0	0.67	0	1						2					
Cluster 2																				
AttG1	3	0.27		3	0	1	0	0.75	1											2
AttG2	4	0.4		4	0	1	0	0.8	1											2
AttG3	3	0.38		3	0	1	0	0.75	1											2
AttG4	3	0.38		3	0	1	0	0.6	1											2
AttG5	2	0.29		2	0	1	0	0.67	1											2
ImG1	5	0.56		5	0	1	0	1	1											2
ImG2	2	0.25		2	0	1	0	0.67	1											2
ImG3	2	0.29		2	0	1	0	0.5	1											2
ImG4	2	0.29		2	0	1	0	0.67	1											2
ImG5	2	0.33		2	0	1	0	0.67	1											2
IntG1	3	0.43		3	0	1	0	1	1											2
IntG2	3	0.5		3	0	1	0	1	1											2
Cluster 3																				
AttG1	4	0.36		4	0	1	0	1	1						1					
AttG2	5	0.5		5	0	1	0	1	1						1					
AttG3	4	0.5		4	0	1	0	1	1						1					
AttG4	5	0.63		5	0	1	0	1	1						1					
AttG5	3	0.43		3	0	1	0	1	1						1					
ImG1	5	0.56		5	0	1	0	1	1						1					
ImG3	4	0.57		4	0	1	0	1	1						1					
ImG4	3	0.43		3	0	1	0	1	1						1					
ImG5	3	0.5		3	0	1	0	1	1						1					
IntG1	3	0.43		3	0	1	0	1	1						1					

Table 5.22 illustrates only corresponding web features of these rules.

TABLE 5.22 DESIGN RULES FOR DIFFERENT CLUSTERS OF USER CHARACTERISTICS AND FOR MULTIPLE KANSEIS

Cluster	Kansei	Web features											
		Font Size	Font Color	Font Type	Font Contrast	Image Size	Image Quality Resolution	Image Content	Page Layout Density	Page layout Color Combination	Picture	Line length	Display Color
Cluster 1													
AttG1	Attractive						High						
AttG2	Attractive						High						
AttG3	Attractive						High						
AttG4	Attractive						High						
AttG5	Attractive						High						
ImG1	Imaginative						High						
ImG2	Imaginative						High						
ImG3	Imaginative						High						
ImG4	Imaginative						High						
ImG5	Imaginative						High						
IntG1	Interesting						High						
IntG2	Interesting						High						
Cluster 2													
AttG1	Not Attractive											Long	
AttG2	Not Attractive											Long	
AttG3	Not Attractive											Long	
AttG4	Not Attractive											Long	
AttG5	Not Attractive											Long	
ImG1	Not Imaginative											Long	
ImG2	Not Imaginative											Long	
ImG3	Not Imaginative											Long	
ImG4	Not Imaginative											Long	
ImG5	Not Imaginative											Long	
IntG1	Not Interesting											Long	
IntG2	Not Interesting											Long	
Cluster 3													
AttG1	Not Attractive						Low						
AttG2	Not Attractive						Low						
AttG3	Not Attractive						Low						
AttG4	Not Attractive						Low						
AttG5	Not Attractive						Low						
ImG1	Not Imaginative						Low						
ImG3	Not Imaginative						Low						
ImG4	Not Imaginative						Low						
ImG5	Not Imaginative						Low						
IntG1	Not Interesting						Low						

TABLE 5.22 (continued)

Cluster	Kansei	Web features											
		Font Size	Font Color	Font Type	Font Contrast	Image Size	Image Quality Resolution	Image Content	Page Layout Density	Page layout Color Combination	Picture	Line length	Display Color
Cluster 4													
AttG1	Attractive					Large							
AttG2	Attractive					Large							
AttG3	Attractive					Large							
AttG5	Attractive					Large							
ImG1	Imaginative					Large							
ImG2	Imaginative					Large							
ImG4	Imaginative					Large							
ImG5	Imaginative					Large							
IntG1	Interesting					Large							
Cluster5													
AttG1	Attractive								Low				
AttG3	Attractive								Low				
AttG4	Attractive								Low				
AttG5	Attractive								Low				
ImG2	Imaginative								Low				
ImG3	Imaginative								Low				
ImG4	Imaginative								Low				
ImG5	Imaginative								Low				
IntG1	Interesting								Low				
Cluster9													
AttG1	Not Attractive									Bad			
AttG2	Not Attractive									Bad			
AttG3	Not Attractive									Bad			
AttG5	Not Attractive									Bad			
ImG1	Not Imaginative									Bad			
ImG5	Not Imaginative									Bad			
IntG1	Not Interesting									Bad			
Cluster10													
AttG2	Not Attractive										No Pic		
AttG4	Not Attractive										No Pic		
ImG1	Not Imaginative										No Pic		
ImG2	Not Imaginative										No Pic		
IntG1	Not Interesting										No Pic		
IntG2	Not Interesting										No Pic		
Cluster11													
AttG3	Attractive										Single		
ImG1	Imaginative										Single		
ImG2	Imaginative										Single		
ImG4	Imaginative										Single		
ImG5	Imaginative										Single		
IntG1	Interesting										Single		

TABLE 5.22 (continued)

Cluster	Kansei	Web features											
		Font Size	Font Color	Font Type	Font Contrast	Image Size	Image Quality Resolution	Image Content	Page Layout Density	Page layout Color Combination	Picture	Line length	Display Color
Cluster14													
AttG1	Attractive									Good			Colorful
AttG2	Attractive									Good			Colorful
AttG3	Attractive									Good			Colorful
ImG1	Imaginative									Good			Colorful
IntG1	Interesting									Good			Colorful
Cluster15													
AttG2	Attractive		Black										
AttG3	Attractive		Black										
AttG5	Attractive		Black										
ImG2	Imaginative		Black										
IntG1	Interesting		Black										
Cluster17													
AttG1	Attractive	Small								Good			
AttG2	Attractive	Small								Good			
ImG1	Imaginative	Small								Good			
IntG1	Interesting	Small								Good			
Cluster22													
AttG2	Not Attractive		Other					Not Natural					
AttG3	Not Attractive		Other					Not Natural					
ImG2	Not Imaginative		Other					Not Natural					
IntG1	Not Interesting		Other					Not Natural					
Cluster26													
AttG5	Attractive				None								
ImG4	Imaginative				None								
ImG5	Imaginative				None								
IntG2	Interesting				None								
Cluster33													
AttG2	Attractive					Medium							
ImG2	Imaginative					Medium							
IntG2	Interesting					Medium							
Cluster36													
AttG4	Not Attractive	Large	Other										
ImG3	Not Imaginative	Large	Other										
IntG1	Not Interesting	Large	Other										

5.3 Rules Corresponding to Incremental Group Pattern

An interesting application is identifying what additional characteristics are required to address the perceptions of one group when a product has been designed based on the perceptions of another group. For example, if group 1 is "female with low score in hedonic," which implies that small font size is preferred, what would be required to include "females with high hedonic score?" This is the vertical movement from a single group to multiple groups in Table 5.23. Also, horizontal movement in Table 5.23 identifies the rules required to cover multiple Kanseis. Note that the user characteristics of G_1K_1 might be different from G_1K_2. Finally, movement along the diagonal of the matrix provides rules for multiple groups, multiple Kanseis. Table 5.23 can be completed with the results.

TABLE 5.23 GROUP PATTERN, MULTIPLE KANSEIS, MULTIPLE GROUPS

Customer Groups	Kanseis						
	K_1	K_2	K_3	K_1 K_2	K_1 K_3	K_2 K_3	K_1 K_2 K_3
G_1	Rules A, B, F						
G_2	Rules A, B						
G_3	Rules A, C						
G_4	Rules A,B						
G_5	Rules A, D						
G_1 G_2	Rules A,B						
G_1 G_3	Rule A						
G_1 G_4	Rules A,B						
…							
$\binom{5}{2}$							
G_1 G_2G_3							
G_1 G_2G_4							
$\binom{5}{3}$							
…							
G_1 G_2G_3 G_4							
G_1 G_2G_3 G_5							
$\binom{5}{4}$							
…							
G_1 G_2G_3 G_4G_5							

5.4 Identify Group of Users by a Specific Set of Product Features

RSBKE can also identify which groups of users are positively or negatively affected by a set of product features controlled by designers, as well as the particular specific design constraints—that is, what are the positive or negative effects of other design attributes? For example, assume that designers must use a small font for a webpage because of the amount of content. The issue here is what other feature combinations can be used to create a positive impact. RSBKE can identify not only the positive/negative combinations of features values, but it can also identify the specific groups of users influenced by that combination.

Table 5.24 shows the combination of features in the case where a small font size is required. As shown, small font size has made an “attractive” impression on user groups AttG1, AttG2, and AttG4. It is interesting that small font size, regardless of other values, has left an “attractive” impression on AttG3.

TABLE 5.24 GROUPS OF USERS POSITIVELY IMPRESSED BY “SMALL” FONT SIZE

Group	Match	Strength	Attractive	Certainty1	Certainty2	Coverage1	Coverage2	No. of Attributes	Product Features									
									Font Size	Font Type	Font Contrast	Image Content	Page Layout Density	Page Layout Color Combination	Picture	Line Length	Display Color	Display Back Graphic
AttG1	4	0.36	Yes	1	0	0.57	0	2	Small	Standard								
AttG1	4	0.36	Yes	1	0	0.57	0	2	Small					Good				
AttG1	2	0.18	Yes	1	0	0.29	0	3	Small						Multiple			2
AttG1	2	0.18	Yes	1	0	0.29	0	3	Small								Colorful	2
AttG1	1	0.09	Yes	1	0	0.14	0	3	Small		Not				Multiple			
AttG1	1	0.09	Yes	1	0	0.14	0	3	Small			Not Natural			Multiple			
AttG1	1	0.09	Yes	1	0	0.14	0	3	Small		Not						Colorful	
AttG2	2	0.2	Yes	1	0	0.4	0	2	Small					Good				
AttG2	1	0.1	Yes	1	0	0.2	0	3	Small	Standard					Multiple			
AttG2	1	0.1	Yes	1	0	0.2	0	3	Small						Multiple			2
AttG2	1	0.1	Yes	1	0	0.2	0	2	Small		Not							
AttG2	1	0.1	Yes	1	0	0.2	0	2	Small				Low					
AttG3	2	0.25	Yes	1	0	0.5	0	1	Small									
AttG4	2	0.25	Yes	1	0	0.67	0	3	Small			Not Natural			Multiple			
AttG4	2	0.25	Yes	1	0	0.67	0	3	Small						Multiple			2
AttG4	2	0.25	Yes	1	0	0.67	0	3	Small			Not Natural				Short		
AttG4	2	0.25	Yes	1	0	0.67	0	3	Small							Short		2
AttG4	1	0.13	Yes	1	0	0.33	0	3	Small	Standard	Not							

In addition, groups of users who were impressed by “large” font size (constant) along with other attributes can be identified, as shown in TABLE 5.25. It should be noted that “large” font size always produced a positive impression. Note that while large font size along with “natural” image content left a positive impression on groups AttG1 and AttgG2, large font size along with “not natural” image content produced a positive feeling in group AttG5. Groups AttG1, AttG2, and AttG5 are not considered as one unified cluster since two opposite features (“natural” and “not natural” image content) cannot be together.

Table 5.26 compares attractiveness groups AttG1 and AttgG2 with AttG5. The distinction between AttG5 and the other two groups is that group 5 has users over 40 years of age. In summary, for these users, the combination of “large” font and “not natural” image content makes the websites more attractive than those with “natural” image content.

TABLE 5.25 GROUPS OF USERS POSITIVELY IMPRESSED BY “LARGE” FONT SIZE

Group	Match	Strength	Attractiveness	Certainty1	Certainty2	Coverage1	Coverage2	No. of Attributes	Font Size	Font Type	Image Content	Page Layout Density	Picture	Line Length	Display Color	Display Back Graphic
AttG1	3	0.27	Yes	1	0	0.43	0	2	Large					Short		
AttG1	3	0.27	Yes	1	0	0.43	0	3	Large				Multiple		Colorful	
AttG1	1	0.09	Yes	1	0	0.14	0	2	Large		Natural					
AttG1	1	0.09	Yes	1	0	0.14	0	2	Large							Yes
AttG2	3	0.3	Yes	1	0	0.6	0	2	Large					Short		
AttG2	3	0.3	Yes	1	0	0.6	0	3	Large				Multiple		Colorful	
AttG2	1	0.1	Yes	1	0	0.2	0	2	Large		Natural					
AttG2	1	0.1	Yes	1	0	0.2	0	3	Large				Multiple			Yes
AttG2	1	0.1	Yes	1	0	0.2	0	3	Large			High				Yes
AttG5	2	0.29	Yes	1	0	0.5	0	2	Large	Standard						
AttG5	2	0.29	Yes	1	0	0.5	0	2	Large		Not Natural					
AttG5	2	0.29	Yes	1	0	0.5	0	2	Large					Short		
AttG5	1	0.14	Yes	1	0	0.25	0	2	Large							No

TABLE 5.26 CHARACTERISTICS OF ATTG1, ATTG2, AND ATTG5

Group	STUCOM	Age	PWORK	PCOMM	Hedonic	Involvement	CVPA	Corresponding Websites
AttG1			1 = Low		2 = Low			11, 17, 23, 26, 30, 33, 35, 38, 41, 46, 8
AttG2				2 = High	2 = Low			11, 13, 17, 24, 30, 33, 38, 41, 46, 8
AttG5		2 = Above 40	1 = Low					24, 26, 38, 41, 46, 8, 9

Discovering the combination of product features given a fixed feature value can be conducted for any feature. A specific level of a feature made a website attractive; the opposite value of that feature made the website not attractive. For example, high quality image resolution made the websites consistently more attractive, while low quality image resolution made the websites not attractive. However, not all features like quality image resolution had a consistent positive/negative effect. As shown in Table 5.24 and Table 5.25, both values of font size had a positive influence when combined with other features.

Other information useful to a designer is identifying the combination of features associated with the attractiveness of the majority of websites. For example, the rules indicated in Table 5.27 are those associated with the attractiveness of the majority of websites corresponding to attractiveness groups AttG1 and AttG2 (cluster 1-2). These websites are shown in Table 5.28. Also, specific websites corresponding to the specific rule can be identified using Table 5.28. For example, rule three of Table 5.27 indicates that black font type and short line length were associated with the attractiveness of websites 30, 33, 26, 35, 38, 41, and 46 and for users in AttG1. By applying these rules in the corresponding websites of AttG1 and AttG2, it can be expected that these website will become attractive to users in AttG1 and AttG2.

TABLE 5.27 ATTRACTIVE DESIGN RULES FOR MAJORITY OF WEBSITES

Cluster 1-2	Rule No.	Match	Strength	Attractive	Certainty1	Certainty2	Coverage1	Coverage2	No. of Attributes	Font Type	Font Contrast	Image Size	Image Quality Resolution	Image Content	Page Layout Density	Page layout Color Comb	Line length	Display Color	Display Back graphic
AttG1	1	7	0.64	Yes	1	0	1	0	1				High						
AttG1	2	7	0.64	Yes	1	0	1	0	2							Good	Short		
AttG1	3	7	0.64	Yes	1	0	1	0	2	Black							Short		
AttG1	4	6	0.55	Yes	1	0	0.86	0	2								Short		No
AttG1	5	5	0.45	Yes	1	0	0.71	0	2							Good		Colorful	
AttG1	6	5	0.45	Yes	1	0	0.71	0	2		2						Short		
AttG2	7	5	0.5	Yes	1	0	1	0	1				High						
AttG2	8	5	0.5	Yes	1	0	1	0	2							Good	Short		
AttG2	9	5	0.5	Yes	1	0	1	0	2	Black							Short		
AttG2	10	5	0.5	Yes	1	0	1	0	2							Good		Colorful	

TABLE 5.28 WEBSITES IN ATTG1, ATTG2…ATTG5

AttG1		AttG2		AttG3		AttG4		AttG5	
Site	Attractive	Site	Attractive	Site	Attractive	Site	Attractive	Site	Attractive
30	Yes	30	Yes	33	Yes	29	Yes	26	Yes
33	Yes	33	Yes	38	Yes	35	Yes	38	Yes
26	Yes	38	Yes	46	Yes	41	Yes	41	Yes
35	Yes	41	Yes	50	Yes	8	No	46	Yes
38	Yes	46	Yes	9	No	15	No	8	No
41	Yes	8	No	11	No	23	No	9	No
46	Yes	11	No	15	No	24	No	24	No
8	No	13	No	17	No	17	No		
11	No	24	No	33					
23	No	17	No						
17	No								

Also, Table 5.29 illustrates the design rules associated with those unattractive websites indicated in AttG1, AttG2, AttG3, and AttG4. These rules influence AttG1, AttG2, AttG3, and AttG4 (cluster 1-2-3-4). The corresponding websites are shown in Table 5.28: small image size and high page layout density (rule no. three) associated with websites 8, 11, 23, and 17.

TABLE 5.29 UNATTRACTIVE DESIGN RULES FOR MAJORITY OF WEBSITES

Cluster 1-2-3-4	Rule No.	Match	Strength	Attractive	Certainty1	Certainty2	Coverage1	Coverage2	No. of Attributes	Font Color	Font Type	Font Contrast	Image Size	Image Quality Resolution	Image Content	Page Layout Density	Page Layout Color Combination	Line Length
AttG1	1	4	0.36	No	0	1	0	1	2	Not Black			Small					
AttG1	2	4	0.36	No	0	1	0	1	1					Low				
AttG1	3	4	0.36	No	0	1	0	1	2				Small			High		
AttG2	4	5	0.5	No	0	1	0	1	2	Not Black			Small					
AttG2	6	5	0.5	No	0	1	0	1	1					Low				
AttG2	7	4	0.4	No	0	1	0	0.8	2				Small			High		
AttG2	8	4	0.4	No	0	1	0	0.8	1								Bad	
AttG2	9	4	0.4	No	0	1	0	0.8	1									Long
AttG2	10	4	0.4	No	0	1	0	0.8	2	Not Black					Not Natural			
AttG3	11	4	0.5	No	0	1	0	1	1				Small					
AttG3	12	4	0.5	No	0	1	0	1	1					Low				
AttG4	13	5	0.63	No	0	1	0	1	1					Low				

The same type of rules can be developed for the other Kansei, imaginative and interesting, and for multiple Kanseis.

Table 5.30 shows that high image quality resolution was associated with attractive, imaginative, interesting websites, for users included in user groups constituting the cluster A1…5, I1…5, In1.

TABLE 5.30 EXCLUSIVE IMPACT OF HIGH IMAGE QUALITY ON MULTIPLE KANSEIS

Clusters A1…5, I1…5, In1	Match	Strength	Attractive	Certainty1	Certainty2	Coverage1	Coverage2	No. of Attributes	Font Size	Font Color	Font Type	Font Contrast	Image Size	Image Quality Resolution
AttG1	7	0.64	Yes	1	0	1	0	1						High
AttG2	5	0.5	Yes	1	0	1	0	1						High
AttG3	4	0.5	Yes	1	0	1	0	1						High
AttG4	3	0.38	Yes	1	0	1	0	1						High
AttG5	4	0.57	Yes	1	0	1	0	1						High
ImG1	4	0.44	Yes	1	0	1	0	1						High
ImG2	4	0.5	Yes	1	0	0.8	0	1						High
ImG3	3	0.43	Yes	1	0	1	0	1						High
ImG4	4	0.57	Yes	1	0	1	0	1						High
ImG5	3	0.5	Yes	1	0	1	0	1						High
IntG1	4	0.57	Yes	1	0	1	0	1						High

The rules indicated in Table 5.31 provide an option for designers in the case where high image quality is not available. As shown, small font size along with other features can make specific websites (indicated in corresponding groups) look attractive, imaginative, and interesting.

TABLE 5.31 GOOD IMPRESSION WITH SMALL FONT SIZE WHEN HIGH QUALITY IMAGE OPTION IS NOT AVAILABLE

Clusters A1, I2,4,5, In1,2	Rule Number	Match	Strength	Positive Impression	Certainty1	Certainty2	Coverage1	Coverage2	No. of Rules	Font Size	Font Color	Font Type	Font Contrast	Page Layout Density	Page Layout Color Combination	Picture	Line Length
AttG1	1	4	0.36	Yes	1	0	0.57	0	2	Small		Standard					
AttG1	2	4	0.36	Yes	1	0	0.57	0	2	Small					Good		
ImG2	3	3	0.38	Yes	1	0	0.6	0	2	Small							Short
ImG4	4	4	0.57	Yes	1	0	1	0	2	Small							Short
ImG5	5	3	0.5	Yes	1	0	1	0	2	Small							Short
IntG1	6	3	0.43	Yes	1	0	0.75	0	2	Small			Not				
IntG1	7	3	0.43	Yes	1	0	0.75	0	2	Small					Good		
IntG2	8	3	0.5	Yes	1	0	1	0	2	Small						Multiple	

5.5 Random Data

Sets of random data for websites 26, 28, 34, 22, 41, 5, and 32 (arbitrary websites) were generated. These data were compared with the original data, which included user characteristics as independent variables and user evaluations as dependent variables. This comparison was done to examine whether RSBKE captures patterns, other than random, in the data. RCSBKE was applied to both sets of data. If the results were not different, then the RSBKE approach did not provide useful information. Rough set measures of the positive regions, numbers of rules, reducts, sizes of cores, and means of lengths of reducts for both data sets were compared.

First, the random data case was examined in which user characteristics were randomized. Results are shown under column “ONE-RND” of Table 5.32. As can be seen, there is no significant difference between “mean of length of reduct” and “positive region” of the case in which there is no structure and experimental data. This indicates that length of reduct or positive region should not be used to identify clusters of users. This deserves further investigation. The “number of reducts” and “number of rules” were greater for data with structure than for experimental data, while the “size of core” was reduced. It appears that once rough set theory relates the decisions of subjects (dependent variable) to data that has no structure, it produces many combinations of attributes from which decisions may originate. In this case, rough set theory cannot produce useful information. In other words, it cannot find the source of user decisions. Furthermore, the number of core attributes is limited, which means that the generated reduct sets are very scattered.

The same approach was used to generate data for the case that both independent (user characteristics) and dependent variables have no pattern. Results show that although the “number of reducts” and “number of rules” are much greater than the same statistics for the experimental data, they are less than the case where independent variables (user characteristics) are random but dependent variables (user decisions) are not random. This implies that there is a greater chance of relating the random dependent data to the random independent data.

TABLE 5.32 COMPARISON MEASURES OF RANDOM DATA AND ORIGINAL DATA

Website	Mean of Reduct Length			No. Reducts			Size of Core			Positive Region			No. Rules		
	ONE-RND	TWO-RND	NO-RND	ONE-RND	TWO-RND	NO-RND	ONE-RND	TWO-RND	NO-RND	ONE-RND	TWO-RND	NO-RND	ONE-RND	TWO-RND	NO-RND
26	7.6	8.7	9	56	37	2	0	0	8	1	0.968	0.968	1160	2740	223
28	7	8.9	8.3	89	25	3	0	2	7	1	0.968	0.968	616	2148	106
34	6.1	8.3	5	44	23	6	1	1	3	1	0.968	0.968	141	2276	32
22	8.2	8.1	9	25	43	1	2	0	9	1	0.968	0.873	2414	2030	730
41	7	8.6	7	40	25	5	1	1	6	1	1	1	742	2547	170
5	8.2	8.9	9.5	27	19	2	1	2	8	0.968	0.968	0.905	2426	2612	638
32	8.4	8.6	9	19	32	1	4	1	9	1	1	0.905	2298	2571	419

ONE-RND = random users characteristics
TWO-RND = both user characteristics and user decisions are random
NO-RND = original data

In summary, results show that the patterns of user characteristics in the original data influence user ratings of websites. In other words, RSBKE identifies user decisions that are better defined than that of random data. This method indicates that there is some intelligence in the assessment of users.

5.6 Statistical Approach

In this section, the results of the study by Phillips (2007) were used to collect the data for this work. As Phillips (2007) began to examine the aesthetic appeal of a website and the impact of that appeal on user satisfaction, she attempted to determine whether the individual differences affected the user ratings of its site appeal. The impacts of rating websites using adjectives such as simple, attractive, interesting, and imaginative were examined. In her study, two individual difference measures were used: Lida's Internet experience scales (hedonic, utility, involvement, and skill) (Lida, 2004) and centrality of visual product aesthetics (CVPA) (refer to section 4.1.1). Independent sample t-tests and one-way ANOVA were conducted. A summary of her findings are as follows:

- High agreement by all users as to which website has high appeal and which has low appeal.
- Significant correlation between IES subscales (hedonic, involvement, utility, and skill).
- No relation between IES and the adjective score.
- No significant correlation between hedonic and high appeal.
- Significant correlation between CVPA and low appeal.
- No relation between CVPA and high appeal adjective scores.
- No relation between IES and CVPA.

This statistical approach could not capture relationships between the elements of IES scale, CVPA (refer to section 4.1.1), and the adjective scores. The study showed that there was only a statistically significant correlation between CVPA and low-appeal website adjective scores, while the results obtained by RSBKE indicated that relationships exist between the elements of IES and website appeal adjectives. Moreover, rough set theory captured the relationship between the combinations of user characteristics and their perceptions compared to the study by Phillips (2007), which examined the relationships between individual differences and adjective scores one at the time, e.g., CVPA and low appeal, or hedonic and high appeal.

In addition, the statistical approaches used the average of the adjective scores across all pairs of adjectives for low-appeal and high-appeal websites. The statistics-based method did not provide any information about a relationship between individual user differences and a specific assessment such as attractiveness of a website. Furthermore, statistical methods cannot provide if-then rules, which indicate the relation between different levels of user characteristics and levels of aesthetic appeal. For example, RSBKE indicated that both “low hedonic” and “low work” influence the perception of some users to assess websites 11, 17, 23, and 8 as not attractive, and websites 26, 30, 33, 34, 38, 41, 46 as attractive. The last, but not least, advantage of RSBKE over the statistical approach in this study is that rough set theory did not require assumptions about the distribution of data while in ANOVA, and regression requires that the error terms be normally distributed.

In an attempt to assess the results from statistical analysis, the relationship between 12 user characteristics (selected for rough set theory analysis) and attractiveness for all 24 websites using the regression method was examined. Results indicated that except for websites 11, 13, 22, 24, 29, and 32, the p-values of ANOVA analysis were greater than 0.05, which means that there is not a statistically significant relationship between the variables. The corresponding p-values and adjusted R-square statistics for all 24 websites are shown in Table 5.33. For a majority of websites, regression could not capture a relationship with the independent variables. Detailed analyses are included in Appendix C.

TABLE 5.33 RESULTS OF REGRESSION ANALYSIS FOR ALL 24 WEBSITES

Website/Attractive	P-Value	Adjusted R-Square	Significant Relationship
1	0.299	0.23	No
5	0.453	0.20	No
8	0.623	-0.035	No
9	0.422	0.010	No
11	0.026	0.202	Yes
13	0.019	0.207	Yes
15	0.19	0.076	No
17	0.275	0.048	No
20	.0528	-0.014	No
22	0.010	0.236	Yes
23	0.131	0.102	No
24	0.024	0.198	Yes
26	0.907	-0.110	No
28	0.336	0.031	No
29	0	0.353	Yes
30	0.458	0.001	No
32	0.017	0.212	Yes
33	0.520	-0.013	No
34	0.822	-0.083	No
35	0.327	0.033	No
38	0.19	0.078	No
41	0.726	-0.058	No
46	0.616	-0.034	No
50	0.351	0.027	No

Even though regression was not able to capture the correlation relationship between the two sets of variables, the RSBKE approach indicated that there are relationships between the characteristics of users and their decisions regarding specific websites. In fact, RSBKE indicated precisely which users support which relationships. This is another advantage of this approach over the statistical method. Table 5.34 and Table 5.35 show some examples of these relationships along with the strength of the relationship.

TABLE 5.34 COMPARISON OF CHARACTERISTICS IDENTIFIED BY ROUGH SET THEORY AND THOSE IDENTIFIED AS SIGNIFICANT THROUGH REGRESSION

Website/Attractive	Rough Set															Regression (with Relation)		
	Attractive	Match	Strength	STUCOM	Gender	Age	FREQ_USE	PENTER	PWORK	PCOMM	Skill	Utility	Hedonic	Involvement	CVPA	P-Value	Adjusted R-Square	Adjusted R-Square-After Stepwise
11	No	13	0.21							2		1						
		Regression		S		S		S						-S		0.026	0.202	0.28
13	No	11	0.17			1		1	1	1								
		Regression						-S	-S			S				0.019	0.207	0.25
22	No	7	0.11			1	1	2										
		Regression				-S								S	S	0.010	0.236	0.26
24	No	19	0.3	2							1		1					
		Regression		S		-S	S		-S							0.024	0.198	0.22
29	Yes	7	0.11		2	1				1								
		Regression										S			-S	0	0.353	0.22
32	Yes	8	0.13		2	1							1	2				
		Regression				S										0.017	0.212	0.23
Ave		10.83	0.17													0.02	0.23	0.24

As shown here, rough set theory can address which user characteristics are connected to their perceptions. In addition, the strength of the connection expresses the proportion of users with the same specific characteristics who have a consistent perception of a specific website. For example, in spite of showing no relationship between user characteristics and attractiveness of website 1 (Table 5.35) using regression, rough set theory did find a relationship between user characteristics “participant communication and involvement” and “attractiveness” of the website. It is interesting that the average match and strength of rules corresponding to websites, for which regression could not find a relationship, are more than those websites that showed relationships using regression (match = 18.3 and strength = 0.30 for no relationship versus match = 10.83 and strength = 0.17 for with a relationship). This means that rough set theory performed better for the case in which the statistical approach could not find a pattern in the data. Another point is the average adjusted R-square for websites shown in Table 5.34 (showed relationship) as 24% (after stepwise regression), which is not a very strong relationship. Also, using a backward stepwise regression method, the significant user characteristics were identified, as shown in Table 5.34 as S (or -S for a negative relationship). What the rough set theory obtained as non-redundant characteristics was compared to what the regression proposed as significant characteristics. For websites, such as 13 and 32, some of the factors obtained by the two methods matched, while for other websites, each method identified different factors. For example, for website 13, rough set theory identified “participant’s communication” and “utility” as influential factors on user decisions, while regression offered “STUCOM,” “age,” “PENTER,” and “utility” as significant factors. This is an issue for future investigation.

TABLE 5.35 COMPARISON OF ROUGH SET THEORY AND REGRESSION—WEBSITES SHOWING NO RELATIONSHIP

Website/Attractive	Rough Set															Regression (No Relation)	
	User Characteristics															Statistics	
	Attractive	Match	Strength	STUCOM	Gender	Age	FREQ_USE	PENTER	PWORK	PCOMM	Skill	Utility	Hedonic	Involvement	CVPA	P-Value	Adjusted R-Square
1	No	14	0.22							2				1		0.299	0.23
5	Yes	7	0.11		2		2			2		1				0.453	0.20
8	No	11	0.17					2						2		0.623	-0.035
9	No	32	0.51												2	0.422	0.010
15	No	13	0.21	1										2		0.19	0.076
17	No	17	0.27						2	2						0.275	0.048
20	Yes	7	0.11				2		2							.0528	-0.014
23	No	15	0.24	1	2				2							0.131	0.102
26	Yes	17	0.27	2										1		0.907	-0.110
28	Yes	28	0.44											2		.336	0.031
30	Yes	15	0.24		2		1			1	1					0.458	0.001
33	Yes	24	0.38	1								1	1			0.520	-0.013
34	Yes	45	0.71													0.822	-0.083
35	Yes	15	0.24				1				1					0.327	0.033
38	Yes	19	0.30	2					1							0.19	0.078
41	Yes	18	0.29		2				2							0.726	-0.058
46	Yes	15	0.24	2	2						1	1				0.616	-0.034
50	Yes	17	0.27						2	2						0.351	0.027
Ave		18.3	0.3													0.4	0.07 (RMS)

5.7 Second Strategy (Maximum Strength Strategy)

This section examines another strategy for choosing user groups to discover whether it provided more efficient rules for identifying discernible classes of users (user cluster characteristics). This strategy was based on the maximum match or the strength of the rules that define the group of users for each website, in other words, identifying situations in which a majority of users for each website supports the rule. For example, Table 5.36 shows the user group rules for website 34. In this strategy, which we call the maximum strength strategy, from all rules, the rule corresponding to the maximum strength (rule no. 8, as highlighted) was selected. The same was done for the other 23 websites.

TABLE 5.36 MAXIMUM STRENGTH OF USER GROUP RULES OF WEBSITE 34—ATTRACTIVENESS

Kansei	Website No.	Match	Strength	Attractive	Not Attractive	Certainty1	Certainty2	Coverage1	Coverage2	Rule No.	STUCOM	Gender	Age	FREQ_USE	PENTER	PWORK	PCOMM	Skill	Utility	Hedonic	Involvement	CVPA
Att-1	34	12	0.19	12		1	0	0.2	0	1										2		
Att-2	34	13	0.21	13		1	0	0.21	0	1				2								
Att-3	34	19	0.3	19		1	0	0.31	0	1								2				
Att-4	34	23	0.37	23		1	0	0.38	0	1					2							
Att-5	34	15	0.24	15		1	0	0.25	0	1			2									
Att-6	34	32	0.51	32		1	0	0.52	0	1												2
Att-7	34	30	0.48	30		1	0	0.49	0	1							2					
Att-8	34	45	0.71	45		1	0	0.74	0	1		2										
Att-9	34	32	0.51	32		1	0	0.52	0	1	2											
Att-10	34	28	0.44	28		1	0	0.46	0	1						1						

For example, at least 45 users evaluated website 34 as an attractive website. Based on this strategy, the user then evaluated every website, as shown in Table 5.37. In addition, Table 5.38 depicts the corresponding user characteristics for each rule. The question is this: If the goal is to extract design rules for the majority of users described in Table 5.38, then can the system generate efficient rules?

TABLE 5.37 ALL WEBSITE USER EVALUATIONS BASED ON MAXIMUM RULE'S MATCH

Website No.	Attractive	Match	Strength	Website No.	Attractive	Match	Strength
1	No	14	0.22	26	Yes	17	0.27
5	Yes	7	0.11	28	Yes	28	0.44
8	No	11	0.17	29	Yes	7	0.11
9	No	32	0.51	30	Yes	15	0.24
11	No	13	0.21	32	Yes	8	0.13
13	No	11	0.17	33	Yes	24	0.38
15	No	13	0.21	34	Yes	45	0.71

TABLE 5.38 USER CHARACTERISTICS CORRESPONDING TO RULE WITH MAXIMUM MATCH—ATTRACTIVENESS

Kansei	Website No.	Match	Strength	Attractive	Not Attractive	Certainty1	Certainty2	Coverage1	Coverage2	No. of Rules	STUCOM	Gender	Age	FREQ_USE	PENTER	PWORK	PCOMM	Skill	Utility	Hedonic	Involvement	CVPA
Att	1	14	0.22		14	0	1	0	0.28	2							2				1	
Att	5	7	0.11	7		1	0	0.25	0	4		2		2			2		1			
Att	8	11	0.17		11	0	1	0	0.22	2					2						2	
Att	9	32	0.51		32	0	1	0	0.54	1												2
Att	11	13	0.21		13	0	1	0	0.27	3							2		1			1
Att	13	11	0.17		11	0	1	0	0.28	4			1		1	1	1					
Att	15	13	0.21		13	0	1	0	0.26	2	1										2	
Att	17	17	0.27		17	0	1	0	0.31	2						2	2					
Att	20	7	0.11	7		1	0	0.22	0	4				2		2				1		1
Att	22	7	0.11		7	0	1			3			1	1	2							
Att	23	15	0.24		15	0	1	0	0.31	3	1	2				2						
Att	24	19	0.3		19	0	1	0	0.33	3	2							1		1		
Att	26	17	0.27	17		1	0	0.31	0	2	2										1	
Att	28	28	0.44	28		1	0	0.48	0	1											2	
Att	29	7	0.11	7		1	0	0.18	0	2										2		2
Att	30	15	0.24	15		1	0	0.3	0	4		2		1			1	1				
Att	32	8	0.13	8		1	0	0.2	0	4		2	1							1	2	
Att	33	24	0.38	24		1	0	0.45	0	3	1								1	1		
Att	34	45	0.71	45		1	0	0.74	0	1		2										
Att	35	15	0.24	15		1	0	0.28	0	3				1				1				1
Att	38	19	0.3	19		1	0	0.35	0	2	2					1						
Att	41	18	0.29	18		1	0	0.32	0	3		2				2						1
Att	46	15	0.24	15		1	0	0.28	0	4	2	2						1	1			
Att	50	17	0.27	17		1	0	0.31	0	2						2	2					

The information in Table 5.37 was used as the basis to construct an information system table for all websites. Using RSES 2.2.2 and the “Reduct-Rule Integrator” program, the corresponding design rules were developed, as shown in Table 5.39. The strength, certainty, and coverage of the rules were calculated using equations in sections4.2.30.1, 4.2.30.2, 4.2.30.3, and 4.2.30.4.

Table 5.39 shows only those rules that have a “match” of more than four, which means that the generated rule is drawn from at least four websites out of 24. The maximum matches of all generated rules were only eight, which means that the system could not find rules that applied to more than eight websites. This implies that rules satisfying everybody or even a majority of users cannot be identified. When the results of this strategy are compared with the previous strategy (“common rules” strategy), it can be seen that the rules of the previous approach are more efficient.

TABLE 5.39 RULES BASED ON TABLE 5.37 USER EVALUATIONS FOR ALL WEBSITES—SECOND STRATEGY—ATTRACTIVENESS

Kansei	Match	Strength	Attractive	Not Attractive	Certainty1	Certainty2	Coverage1	Coverage2	No. of Attributes	Font Size	Font Color	Font Type	Font Contrast	Image Size	Image Quality Resolution	Image Content	Page layout Density	Page Layout Color Combination	Picture	Line Length	Display Color
Att	8	0.33		8	0	1	0	0.8	3		2				1			2			
Att	8	0.33		8	0	1	0	0.8	2					1	1						
Att	7	0.29	7		1	0	0.5	0	2						2				3		
Att	6	0.25	6		1	0	0.43	0	3			1				2				1	
Att	6	0.25		6	0	1	0	0.6	1											2	
Att	6	0.25		6	0	1	0	0.6	3		2				1	2					
Att	6	0.25	6		1	0	0.43	0	3									1	3	1	
Att	5	0.21		5	0	1	0	0.5	2				2		1						
Att	5	0.21	5		1	0	0.36	0	1		1										
Att	5	0.21		5	0	1	0	0.5	3	2	2			1							
Att	5	0.21	5		1	0	0.36	0	3	2								1		1	
Att	5	0.21	5		1	0	0.36	0	3	2		1								1	
Att	5	0.21	5		1	0	0.36	0	3									1	3		3
Att	5	0.21	5		1	0	0.36	0	2						2	2					
Att	5	0.21	5		1	0	0.36	0	2						2		1				
Att	4	0.17		4	0	1	0	0.4	3	1	2				1						

By comparing Table 5.39 and Table 5.40, it can be realized that, in general, the strength of the rules generated by the "common rules" strategy is stronger than the "max strength" strategy. The same is true for the coverage of the rules, as shown in Table 5.40. The coverage of the rules was typically 100% (Table 5.40), while the coverage of the rules described in Table 5.39 was between 40% and 60%. For only two rules, it was 80%. As mentioned in Chapter 4, the coverage of a rule is the number of objects (here, websites) that have the same conditions and decision attribute values divided by the number of objects that have the same decision values. Less coverage of rules in Table 5.39 implies that rules cannot be extracted, even from a majority of websites that are attractive or not attractive. In the "common rules" strategy, by selecting websites using consistent discernible classes of users, the rules were able to cover almost 100% of websites with the same decision.

Another measure to compare the efficiency of the rules for two different strategies (the efficiency measure mentioned in Chapter 4), is the number of attributes that a rule contains. As shown in Table 5.39, the number of attributes for most of the rules was three, while this number in Table 5.40 was one. This is another reason that the "common rules" strategy is more efficient than the "max strength" strategy. Moreover, the rules of the "common rules" strategy were more descriptive and concise than the "max strength" strategy. This is because the "common rules" strategy provided rules for a specific discernible class of users and selected websites in which the relevant rules are extracted (Table 5.41) according to the characteristics of the discernible class of consistent users. That is why the coverage of the generated rules of the "common rules" strategy was greater than the "max strength" strategy. In summary, the rules of the "common rules" strategy provided more information, targeted specific groups, and were stronger and more succinct.

TABLE 5.40 DESIGN RULES GENERATED BY FIRST STRATEGY—ATTRACTIVENESS

Cluster (C)	Match	Strength	Attractive	Not Attractive	Certainty1	Certainty2	Coverage1	Coverage2	# of characteristic	Font Size	Font Color	Font Type	Font Contrast	Image Size	Image Quality Resolution	Image Content	Line Length	Display Color	Display Background Graphic
C1																			
AttG1	7	0.64	7		1	0	1	0	1						2				
AttG2	5	0.5	5		1	0	1	0	1						2				
AttG3	4	0.5	4		1	0	1	0	1						2				
AttG4	3	0.38	3		1	0	1	0	1						2				
AttG5	4	0.57	4		1	0	1	0	1						2				
C2																			
AttG1	7	0.64	7		1	0	1	0	2			1					1		
AttG2	5	0.5	5		1	0	1	0	2			1					1		
AttG3	4	0.5	4		1	0	1	0	2			1					1		
AttG4	3	0.38	3		1	0	1	0	2			1					1		
AttG5	4	0.57	4		1	0	1	0	2			1					1		
C3																			
AttG1	4	0.36		4	0	1	0	1	1						1				
AttG2	5	0.5		5	0	1	0	1	1						1				
AttG3	4	0.5		4	0	1	0	1	1						1				
AttG4	5	0.63		5	0	1	0	1	1						1				
AttG5	3	0.43		3	0	1	0	1	1						1				
C 4																			
AttG1	3	0.27		3	0	1	0	0.75	1								2		
AttG2	4	0.4		4	0	1	0	0.8	1								2		
AttG3	3	0.38		3	0	1	0	0.75	1								2		
AttG4	3	0.38		3	0	1	0	0.6	1								2		
AttG5	2	0.29		2	0	1	0	0.67	1								2		
C 5																			
AttG1	4	0.36	4		1	0	0.57	0	2			1				1			
AttG2	3	0.3	3		1	0	0.6	0	2			1				1			
AttG3	3	0.38	3		1	0	0.75	0	2			1				1			
AttG5	2	0.29	2		1	0	0.5	0	2			1				1			
C 6																			
AttG1	4	0.36	4		1	0	0.57	0	2							1			2
AttG2	3	0.3	3		1	0	0.6	0	2							1			2
AttG3	3	0.38	3		1	0	0.75	0	2							1			2
AttG5	2	0.29	2		1	0	0.5	0	2							1			2

TABLE 5.41USER CHARACTERISTICS LEADING TO SPECIFIC WEBSITES

Cluster	Corresponding Websites	Group Descriptors
AttG1	11, 17, 23, 26, 30, 33, 35, 38, 41, 46, 8	PWORK = Low; Hedonic = Low
AttG2	11, 13, 17, 24, 30, 33, 38, 41, 46, 8	PCOMM = High; Hedonic = Low
AttG3	11, 15, 17, 33, 38, 46, 50, 9	PWORK = Low; Involvement = Low
AttG4	15, 17, 23, 24, 29, 35, 41, 8	Student, PWORK = High; CVPA = High
AttG5	24, 26, 38, 41, 46, 8, 9	AGE = Above 40; PWORK = Low

5.8 Lead Users

This section discusses a process for identifying more introspective or predictive users, users who model the decisions (perceptions) of other users. Some users consistently represent the judgment of a group and are classed in different sets of users. Five groups, AttG1 through AttG5, were examined. Results are shown in Table 5.42 and Table 5.43. For AttG1, Table 5.42 shows that users 28, 39, 41, 46, 50, and 52 (users who have the same hedonic and participant's work) are consistent in their evaluation of websites 26, 38, 41, 30, 46, 35, 8, 23, 33, 11, and 17, but do not agree on websites 22, 5, etc.

TABLE 5.42 USERS IN ATTG1

User	STUCOM	Gender	Age	FREQ_USE	PENTER	PWORK	PCOMM	Skill	Utility	Hedonic	Involvement	CVPA	Attractive 26	Attractive 22	Attractive 38	Attractive 41	Attractive 5	Attractive 30	Attractive 46	Attractive 35	Attractive 8	Attractive 23	Attractive 33	Attractive 11	Attractive 17
28	1	2	1	2	1	1	2	1	1	2	1	2	1	2	1	1	1	1	1	1	2	2	1	2	2
39	2	2	2	1	1	1	1	1	1	2	1	2	1	2	1	1	2	1	1	1	2	2	1	2	2
41	2	2	2	1	2	1	2	2	1	2	2	2	1	2	1	1	1	1	1	1	2	2	1	2	2
46	2	1	2	1	2	1	2	2	1	2	2	1	1	1	1	1	2	1	1	1	2	2	1	2	2
50	2	2	2	1	1	1	1	2	1	2	1	2	1	2	1	1	2	1	1	1	2	2	1	2	2
52	2	2	1	1	1	1	1	1	1	2	2	1	1	1	1	1	2	1	1	1	2	2	1	2	2

TABLE 5.43 LEAD USERS IN ATTG1 THROUGH ATTG5

AttG1 User	AttG2 User	AttG3 User	AttG4 User	AttG5 User
28	12	31	1	14
39	16	38	5	39
41	28	41	10	41
46	36	46	19	46
50	41	52	24	50
52	46	61	25	53
		62		55
		63		
		69		
Total Users (AttG1 through AttG5)				24

The users who are in the other groups are identified and shown in Table 5.43. As can be seen, some users, such as users 41 and 46, are shown repeatedly in different consistent groups. They have characteristics to evaluate different websites (different situations). Therefore, they are called “lead” or “super users” who can model other cluster members’ behavior.

5.9 Kano Model

A Japanese professor, Noriaki Kano (1996), suggested a model to recognize different classifications of customer expectations: dissatisfiers, satisfiers, and exciters/delighters. This classification of customer needs is based on customer perceptions based on how the product performed with respect to a customer’s expectations. This model does not provide information regarding the interaction between customer satisfaction elements when there are multiple expectations.

For the information system shown in Table 5.44 , many decision rules can be generated. This study suggests choosing rules that have a single customer satisfaction element in their condition part and using them to identify “dissatisfiers,” “satisfiers,” and “exciters” elements, as Table 5.45 illustrates. In this table, only six rules are shown.

TABLE 5.44 INFORMATION SYSTEM TABLE FOR COMPANY'S WEBSITE (SATISFACTORY = 2, NOT SATISFACTORY = 1)

Subject No.	Customer Satisfaction Elements (Condition Attributes)										Decision Attributes
	Simple (c_1)	Interesting (c_2)	Well Designed (c_3)	Easy to Use (c_4)	… …	Good Use of Color (c_8)	Good Layout (c_9)	Imaginative (c_{10})	Attractive (c_{11})	Easy to Navigate (c_{12})	Overall Satisfaction
1	1	1	1	1		2	1	2	2	2	2
2	1	2	1	2		1	1	1	1	1	1
3	1	1	1	1		1	1	1	1	2	2
4	1	2	1	2		2	1	1	1	1	2
5	1	2	1	1		2	1	2	2	2	2
6	1	2	1	1		2	1	1	1	1	2
7	1	1	1	1		1	1	1	1	1	1
8	1	1	1	1		1	1	1	1	1	2
9	1	1	1	1		1	1	1	2	1	2
10	1	1	1	1		1	1	1	1	2	2
11	1	2	1	1		1	1	1	2	1	2
13	1	2	1	1		1	1	1	1	1	2
…	…	…	…	…	…	…	…	…	…	…	…
…	…	…	…	…	…	…	…	…	…	…	…
61	2	1	1	1		1	1	1	1	1	1
62	2	2	1	1		1	1	1	2	1	2
63	2	2	1	1		1	1	1	1	1	2

TABLE 5.45 EXAMPLES OF DECISION RULES EXTRACTED FROM INFORMATION SYSTEM (SATISFACTORY = 2, NOT SATISFACTORY = 1)

Rule No.	Satisfaction Elements							Type of Satisfaction Elements
	Simple (c_1)	Interesting (c_2)	Easy to Use (c_4)	……	Good Use of Color (c_8)	Imaginative (c_{10})	Overall Satisfaction	
1					2		2	Satisfiers
2					1		1	
3		2					2	Exciters
4	2	1				2	2	
5			1				1	Dissatisfiers
6	2		1			2	1	

Dissatisfiers are the elements that were expected in the websites. If they were realized in the website, the overall customer satisfaction should not be affected. If they were not met, it should cause dissatisfaction. Thus, rules 5 and 6 in Table 5.45 imply that they were dissatisfier (basic) elements of customer satisfaction, whereby if the customers were not satisfied with the "easy to use" website, then they became dissatisfied, even though the website met other satisfaction elements (“simple” and “imaginative”). Therefore, “easy to use” could be considered a basic customer satisfaction element.

Also, rules 1 and 2 in Table 5.45 can be interpreted as an indication of “satisfier” elements (performance), those which, if they exist in the website, create satisfaction, and if not, cause dissatisfaction. Therefore, “good use of color” is a “satisfier” element. Finally, those elements that, if not included in the website, do not cause dissatisfaction, but if they are included, positively influence customer satisfaction, are shown in rules 3 and 4 as “exciter” elements. Thus, “interesting” was an “exciter” satisfaction element.

5.10 Conclusions

In this chapter, the output of the RSBKE method was presented. Results showed that the proposed approach is capable of identifying influential user characteristics and product features. In addition, the approach generated efficient rules to identify discernible classes of users and design rules.

To ensure that the rules can provide useful information, rules were generated using randomized data and the corresponding statistics were compared. Results showed that the number of reduct sets and the number of rules were fewer than the same statistics for random data, while the size of the core sets of rules was much larger than the random data case. Therefore, the rules generated by the proposed approach were informative and represented the structure in data.

To validate the proposed approach, results were compared with the output of a statistical multiple regression approach. Regression did not capture the relationships between user characteristics and user perceptions for 18 out of 24 websites. For websites in which regression identified relationships, the power of connectivity was low, and the significant attributes were different than those identified by rough set theory. In addition, regression could not present “if-then” rules, as was done by RSBKE.

Finally, this chapter included an approach for identifying lead users who appeared in multiple discernible classes of users. These users had the same set of characteristics when they judged different websites. In fact, they represented the opinions of other cluster members. Identifying these users may be very valuable for surveys and designers.

CHAPTER 6

CONCLUSIONS AND FUTURE WORK

6.1 Conclusions

Fitting product features to customer needs has always been a challenge for designers. One of the most important challenges is created by the range of requirements that has originated from natural individuality of human beings. This individuality of users creates heterogeneity in the market. The distinctiveness of people produces unique perceptions of product characteristics. The single design of a product may address the needs of enough customers to make it successful, or multiple designs of a product may be generated to address individual requirements.

Many tools have been developed to understand customers and their requirements in a market research context. These tools are used to discover what customers want, need, believe, and even feel about the products. On the other hand, customer-oriented product development approaches, such as Kansei Engineering or quality function deployment, try to identify which product features respond to specific customer requirements.

Maximizing the satisfaction level of customers, while meeting economic objectives, is the ultimate goal of every company. Since the satisfaction of customer requirements (functional or non-functional) depends upon a specific set of product features, each specific set provides different levels of satisfaction for each customer. By identifying the characteristics of customers that have the same pattern of assessments, market decisions can be supported. Efficiently identifying a group of customers who are satisfied by a set of product features helps companies to map their markets and concentrate on their natural customer classes. This is useful for designers. On the other hand, identifying a set of product features that satisfies a specific targeted market allows the company to address specific customers. This is useful for marketers.

Either by finding the set of values that satisfies a certain customer group(s) or given certain customer characteristics, then determining the corresponding product values is necessary for product designers and marketers. These are identified as forward and backward strategies. In the forward strategy, given a specific set of user characteristics, what product features values respond to them? Conversely, the backward strategy implies a given set of product features and then identifies the user groups addressed by these features. By changing the values of product features, it can be determined how groups of customers are sensitive to changes in design. For example, in case it is necessary to change some product features due to technological or economic constraints, is it still possible to retain specific customer groups? The backward strategy also can help companies reveal hidden natural customer groups. This aids marketers in designing and planning market strategies for these natural clusters of customers and expanding the market. The method proposed in this study addresses these strategies.

The benefits of the results of the approach can be seen in the design/development process context.

Another contribution of this work is providing a link between the market research and customer-oriented product development approaches. Usually, market research techniques gather and prioritize customer requirements for customer-oriented product-development processes. Sometimes these techniques in customer-oriented product development context are called voice of the customer (VOC). The main aim of VOC is to provide appropriate customer information for the product development process. This is done through collection, translation, and interpretation of customer requirements, and then structuring, quantifying, and integrating them for use in the deployment process. Since large amounts of customer data are accumulated in the VOC process, the process must narrow the number of customer requirements according to their contributions to customer satisfaction. To do so, VOC uses "an average" to assign an importance to each requirement in order to identify the set of requirements that matters most to customers, and since the calculation of the mean is known, this is not a permissible operation for ordinal data. Furthermore, if a mean could be used for VOC, the fundamental approach of design for the average person is not valid. No one customer has an "average" set of characteristics or needs. The RSBKE does not use the mean but provides information to design for the most similar users in terms of the way they use or perceive the product. Moreover RSBKE is one of the few approaches in which both multiple users and multiple Kanseis cases are considered. Note that RSBKE can fit into the typical human-centered design/development process. Usually the typical customer-oriented product development process has two main globes: customer world and engineering world. Traditionally, in the customer world, or VOC, product users are grouped and targeted, and their needs are gathered and prioritized. On the engineering side, designers try to address user needs in the product. The two-stage subjective impression-based approach (RSBKE) is fitted into the typical customer-oriented product development. In the first stage, the approach helps marketers to define the natural classification or redefine the traditional user groups based on their assessments. In the second stage, the approach generates design rules for each group of customers.

RSBKE is an approach for reasoning under uncertainty that deals with imperfect information. The origin of imperfection in human decision-making exists in the range of user evaluation data. The vagueness of Kanseis notions makes users unable to uniquely classify each product into a specific class. That is why, in the specific data here for example, it was observed that there was no website upon which all users were able to agree with respect to "attractiveness." This made the case initially a P-indefinable set, since there was no lower approximation of classification, and the upper approximation was the universal set. Hence, this approach suggested determining from where these inconsistencies come. One information system table was built for each website with respect to one specific Kansei. Then, discriminatory characteristics were added to make users discernible in terms of their condition attributes. If every C-indiscernibility class has a unique value for the decision attributes, then these sets become crisp, and it is possible to go to the next step to extract perfect knowledge from imperfect knowledge. Otherwise, adding more discriminatory characteristics, deleting inconsistent objects, or surveying for those inconsistent objects should be repeated until inconsistency is resolved to prevent the system from producing non-deterministic rules, which are useless. Note that the approach started to encounter a P-indefinable set in the product information system table, going to the origination of vagueness in the user level, and was encountered again with the case of no lower approximation set for the system. Therefore, in both levels of product and user information system tables, there was vagueness and imperfect knowledge in which rough set theory could be applied. When the case was indefinable, RSBKE converted the situation to the case of a standard rough environment in which lower and upper approximations exist. Once the inconstancies among users were resolved, the second process of extracting the perfect knowledge was started from the user level. In this step, which it was called the first step in RSBKE, the imperfect knowledge was reduced to perfect knowledge and was able to produce rules to identify indiscernible classes of users in which there was no inconsistency among them on the vague Kansei concepts. Partitions were defined by decision attributes identified by the condition attributes. These condition-decision indiscernible classes became the input for the second step in which the condition-decision indiscernible classes at the product level were produced as product design rules. Note that RSBKE provides performance measures of the rules generated.

Although the main results of the RSBKE are the grouping policies for efficient users and product design rules for selected groups for a single/multiple Kansei(s), the flexibility of this approach provides other useful information. Identifying groups of users who are positively or negatively affected by a set of product features controlled by designers, helps designers to identify and evaluate the effects of different design plans (what-if analysis) on specific users (starting with specific rules, identifying specific groups). Also, this helps to identify cost-effective design plans for the most beneficial customer groups. Moreover, the system can provide rules corresponding to incremental group patterns in which rule change can be monitored as the number of groups and/or Kanseis change. This also assists designers in doing a cost benefit analysis if they want to attract other groups. By changing the group rule selection policy to different cases, RSBKE can provide design rules for each case. RSBKE is also able to identify lead users who are very important for an efficient customer survey. Moreover, another useful result is identifying “dissatisfier,” “satisfier,” and “exciter” elements in the Kano model using generated rules from an information system table corresponding to customer satisfaction elements.

Results from using random data in RSBKE showed that this method can effectively and accurately identify patterns in the data. Finally, the RSBKE approach was validated by comparing the results of this method with the results of a statistics-based approach. The results of RSBKE were accurate, efficient, and effective.

Below are listed some of the applications of this study:

- A way to use rough set theory in Kansei Engineering considering multiple users and/or multiple Kanseis.
- A method for performing perception-based customer classification as compared to function-based customer segmentation. While function-based segmentation may be suitable for functional requirement-based product development approaches, such as "quality function deployment," perception-based segmentation is helpful for perception-based customer-oriented product development approaches. Not many studies have investigated aspects of market segmentation based on customer perception.
- A way to group consistent sets of customers for QFD. Since most functional customer requirements are communicated through subjective perceptions, a rough set-based approach can identify the natural classifications of product users based on the similarity of their assessments of functional requirements.

The following applications make QFD and Kansei Engineering more robust and efficient, since they provide an enhanced means of defining customer groupings:

- Selecting elements critical to overall customer satisfaction can be another application of the proposed approach.
- Enhancing the Kano model.
- Identifying lead users.
- Using an integrated two-stage customer perception-based product development approach. This approach helps designers to plan for the desired customer groups (what the company wants), the natural classification of users, the majority of users for the single Kansei case or multiple Kanseis, and the incremental user group and/or Kanseis.

The limitations of the approach can be categorized into two types:

- Computational limitation, which mainly relates to software limitation. Particularly, RSES (the software used in this study) is able to analyze compound data such as reduction or rule calculation, the practical limit of which is around 5,000 objects, or 130–150 attributes. The calculation process time depends upon the computer RAM and processor.
- Specialization of the decision classes may decrease the efficiency of the approach. That is, if the number of decision classes increases, the number of generated rules dramatically increases. This results in low support for the rules that identify the group and in turn reduces the applicability of the approach. The point between specialization and generalization can be defined by designers.
- The approach did not propose any solution when there is no commonality across reducts and rules for different products.

Finally, the approach can be helpful for industries that intend to efficiently adopt the user-centered design philosophy into their product design and development process in the following ways:

- To identify influential users' characteristics based on their product selection by assessing one or multiple perception(s).
- To generate customer segmentation rules to identify or re-identify customer groups. Also, to identify the customer classification for possible non-existent products (backward strategy mentioned above).
- To identify influential product features that affect one or multiple impression(s) of one group or multiple groups of customers (forward strategy mentioned above).
- To generate design rules for one or multiple impression(s) of one group or multiple groups of customers.

- To generate both types of design rules, positive and negative, in order for designers to maximize the positive impact of their design and minimize the negative influence.
- To provide guidelines for designers to begin designing for one specific group and then expanding the design for other desired groups, with minimum changes in the original platform.
- To provide instructions in designing a universal platform using general design rules and customizing it for different users group, with minimum necessary changes.
- To provide flexible approaches for accommodating design restrictions, such as fixed product features, mandated standards, restrictions on specific features values, etc.
- To identify one specific platform that impressed two different groups—one positively and one negatively—at the same time, and also to identify relevant groups.

6.2 Future Work

This study can be extended by investigation in the following areas:

- The usefulness of RSBKE depends upon researchers' ability to collect the sufficient set of user characteristics and product features. If important attributes are missed, then the value of the proposed approach is reduced. An investigation to find a way to ensure that all relevant attributes are considered is suggested.
- Although the RSBKE is compared with a statistical approach, a comparison with artificial neural networks is suggested. Generating if-then rules using the regular back propagation method with ANNs is difficult and inefficient. The use of recent advanced studies in ANN (Liu, Zhang, & Wu, 2003) may provide a tool for extracting rules, which should be investigated in future research.

- The results of the case presented provide suggestions for a webpage design. User satisfaction of web pages based on this design could be compared with traditional webpage designs.
- As mentioned in section 5.8, the lead users are examined for five groups of users (AttG1 through AttG5) and for one Kansei (attractiveness). It is suggested that other generated groups (AttG6, 7, etc.) and other Kanseis (imaginativeness and interesting) be examined to see if they still hold the power of imitation for other evaluation situations. Also, their leadership ability should be validated externally by other means.
- As described in section 5.7, there was no significant difference between “mean of length of reduct” and “positive region” for the two cases—random and patterned data. Since these measures, especially “positive region,” are very important to the rough set theory concept, the validation of this measure should be investigated.
- In this study, the case where websites were fuzzy in which the majority of users classifies the site as neither positive nor negative was not examined. In such a case, it is possible to have a good impression of a website for one group of people and a bad impression of the same website for another group. Although there were few of these cases, it is anticipated that there could be different user groupings and different website features.
- A larger sample size could be developed in order to examine the accuracy of the classification by splitting the data set into two—one set for training and the other for testing. Using the training data, the rules are generated, and while using the testing data, the classification accuracy of the generated rules could be tested.
- To provide useful information for designers, it is suggested that the economic impact of adding more features be assessed. Including suggested features into a product design can

be expensive, so it would be useful to consider alternative features. By performing an economic analysis, the trade-off value of each feature could be assessed, thus helping designers to make a comprehensive decision.

- In this study, two strategies, “max strength” and “common rules,” for choosing user groups were introduced. Investigating other strategies and comparing them with current strategies is suggested.
- This study focused on the subjective impression (psychosocial requirements) of users as their decision attributes. RSBKE could easily be extended to include the case of functional customer requirements decision attributes (Ahmady, Malzahn, & Cheraghi, 2010).

LIST OF REFERENCES

REFERENCES

Ahmady, A. (2008). Review and Classification of Kansei Engineering and Its Applications. *Proceedings of the 4th Annual GRASP Symposium* (pp. 143-144). Wichita: Wichita State University.

Ahmady, A., & Zegordi, S. (2002). Application of Quality Function Deployment to the Design and Production of Persian Hand-Made Carpet. *proceeding of Eighth International Symposium on QFD* (pp. 329-338). Munich, Germany: QFD Institut Deutschland e.V., QFD-ID.

Ahmady, A., Malzahn, D., & Cheraghi, S. H. (2010). An Efficient Customer Group Selection for Quality Function Deployment . *Proceedings of The 16th International Symposium on QFD.* Portland: QFD Institute.

Akao, Y. (2004). *QFD: Quality Function Deployment - Integrating Customer Requirements into Product Design* . Productivity Press.

An, A., Shan, N., Chan, C., Cercone, N., & Ziarko, W. (1996). Discovering Rules for Water Demand Prediction:An Enhanced Rough-Set Appraoch. *Engineering Application Artificial Intellegent , 9* (6), 645-653.

Anitawati, M., Nor Laila, M., & Nagamachi, M. (2007). *Kansei Engineering:A Study on Perception of Online Clothing Websites.* Lund: Campus Helsingborg Lund University.

Anton, J., & Perkins, D. (1997). *Listening to the Voice of the Customer 16 Steps to a Successful Customer Satisfaction Measurement Program.* Alexander Communication Group, Inc.

Armstrong, G., & Kotler, P. (2006). *Marketing: An Introduction* (8th ed.). Prentice Hall.

Ayyub, B. M., & Klir, G. J. (2006). *Uncertainty Modeling and Analysis in Engineering and the science.* Boca Raton: Taylor & Francis Group, LLC.

Bianchi-Berthouze, N. (2001). Kansei-Mining:identifying visual impression as patterns in images. *IFSA World Congress and 20th NAFIPS International Conference*, (pp. 2183-2188). Vancouver.

Bouchafra, D., & Tan, J. (2003). Structural hidden Markov model and its application in automotive industry. *Artificial Intelligence and Decision Support Systems* , 155-162.

Brown, E. (1993). *Market Focus, Achieving and sustaining marketing effectiveness,.* London: Buttherworth-Heinmann Ltd.

Chaparro, B. (2009). Website technical features. (A. Ahmady, Interviewer)

Grimsaeth, K. (2006). *Kansei engineering, linking emotions and product features.* Department of product design, Norwegian University of science and technology.

Gutman, J. (1982). Means-End Chain Model Based on Consumer Categorization Processes. *Journal of Marketing , 46* (2), 60-72.

Hideyoshi, Y., & Fukuda, S. (2005). Interactive Reduct Evolutional Computation for Aesthetic Design. *Journal of Computing and Information Science in Engineering , 5*, 1-7.

Hideyoshi, Y., & Shuichi, F. (2007). Design optimization for customers' KANSEI requirement: application of interactive reduct evolutional computation to industrial design with curves. *Int. J. manufacturing technology and management , 10* (4), 360-379.

Hill, N., & Alexander, J. (2000). *Handbook of customer satisfaction and loyalty measurement* (Second ed.). Great Britain: Gower.

Hon Yee Siu, H., & KL Ho, J. (2005). Visual design for a webpage. *Proceedings of CybErg2005. The fourth International Cyberspace Conference on Ergonomics.* International ergonomics Association Press.

Horiguchi, A., & Suetomi, T. (1995). A Kansei engineering approach to a driver/vehicle system. *International Journal of Industrial Ergonomics , 15*, 25-37.

Imamura, K., Nomura, J., Tamura, H., & Nagamachi, M. (1996). Kitchen planning support system employing Kansei engineering and virtual reality. *Advances in Production Management System , 6*, 107-110.

International Organization for Standardization, (2005). *ISO 9000:2005, Quality management systems -- Fundamentals and vocabulary.* Geneva: International Organization for Standardization.

Inuiguchi, M., Hirano, S., & Tsumoto, S. (2003). *Rough Set Theory and Granular Competing.* Berlin , Germany: Springler-Verlag Berlin Hidelberg.

Ishihara, S., Ishihara, K., Nagamachi, M., & Matsubara, Y. (1997). An analysis of Kansei structure on shoes using self-organizing neural networks. *International Journal of Industrial Ergonomics , 19*, 93-104.

Ishihara, S., Ishihara, K., Nagamachi, M., & Matsubara, Y. (1995). An automatic builder for a Kansei engineering expert system using self-organizing neural networks. *International Journal of Industrail Ergonomics , 15*, 13-24.

ISO. (2000). ISO 9000:2000, Quality management systems – Fundamentals and vocabulary. Geneva: International Organization for Standardization.

Jindo, T., & Hirasago, K. (1995). Application studies for car interior of Kansei engineering. *International Journal of Industrial Ergonomics , 19*, 105-114.

Juran, J. M. (1992). *Juran on Quality by Design The New Steps for Planning Quality into Goods and Services.* The Free Press.

Kano, N. (1996). *Guide to TQM in Service Industries.* Productivity, Inc.

Kaplan, D. (2004). *The sage handbook of quantitative methodlogy for the social science* . London: Sage publication, Ltd.

Komorowski, J., Polkowski, L., & Skowron, A. (1998). *Rough sets: A tutorial.* Poland, Norway: the European Union 4th Framework Telematics project CardiAssist, by the ESPRIT-CRIT 2 project #20288, and by the 8T11C 02417 grant from the National Committee for Scientific research.

Kosaka, H., Nishitani, H., & Watanabe, K. (2005). Estimation of reaction force of keyboard switch based on Kansei information using neural networks. *Networking, Sensing and Control, 2005. Proceedings*, (pp. 425-430).

Kotler, P., & Armstrong, G. (2008). *Principle of Marketing.* New Jeresy: Pearson Prentice Hall.

Kudo, Y., & Murai, T. (2007). Semi-Structured Decision Rules in Object-Oriented Rough Set Models for Kansei Engineering. In *Rough Sets and Knowledge Technology* (pp. 219-227). Berlin : Springer Berlin / Heidelberg.

Liu, Y., Zhang, B., & Wu, G. (2003). Impulse Force Based Art Network With Ca Optimization. *IEEE Int. Conf. Neural Networks & Signal Processings,* (pp. 499-502). Nanjing, China.

Mano, M., & Kim, R. C. (2004). *Logic and Computer Design fundamentals.* Pearson Prentice Hall.

Margaret, B., & Biemans, W. G. (1995). *Product developement,meeting the chalenge of the designmaketing interface.* John Wiley & Sons Ltd.

Matsubara, Y., & Nagamachi, M. (1995). Hybrid Kansei engineering system and design support. *Advanced in Human Factors/Ergonomics , 20* (1), 161-166.

Matsubara, Y., & Nagamamachi, M. (1997). Hybrid Kansei Engineering System and design support. *International Journal of Industrial Ergonomics , 19*, 81-92.

Nagamachi, M. (2000). Application of Kansei engineering and concurrent engineering to a new cosmetic product. *Proceedings of the second International conference ERGON-AXIA 2000* (pp. 3-6). Warsaw: Central Institute for Labour Protection.

Nagamachi, M. (1994). Implication of Kansei engineering and its application to automotive design consultation. *Proceedings of the 3rd Pan-Pacific Conference on Occupational Ergonomics,* (pp. 171-175). Seoul-Korea .

Nagamachi, M. (2002). Kansei engineering as a powerful consumer-oriented technology for product development. *Applied Ergonomics , 33* (3), 289-294.

Nagamachi, M. (1997). Kansei engineering as consumer-oriented ergonomic technology of product developemnt. *Proceedings of the 13th triennial congress of the international ergonomics association*, (pp. 228-230). Tampere, Finland.

Nagamachi, M. (1999). Kansei engineering: A new consumer-oriented technology for product developemnet. In *The Occupational Ergonomic Handbook* (pp. 1835-1848). CRC press LLC.

Nagamachi, M. (1995). Kansei engineering: A new ergonomic consumer-oriented technology for product development. *International Journal of Industrial Ergonomics , 15*, 3-11.

Nagamachi, M., Okazaki, Y., & Ishikawa, M. (2006). Kansei engineering and application of the rough sets theory. *Proceedings of the Institution of Mechanical Engineers, Part 1, Journal of System and Control Engineering , 220*, 763-768.

Nagasawa, S. (2004). Present state of Kansei engineering in Japan. *IEEE international conference on systems, man and cybernetics* , 333-338.

Nishino, T., Nagamachi, M., & Tanaka, H. (2006). Variable Percision Baysian Rough Set model and its application to Knasei engineering. *Transaction on Rough sets V, LNCS 4100- Springer-Verlag Berlin Heidelberg* , pp. 190-206.

Oyama, H. (1997). Virtual reality for the palliative care of cancer. *Studies in health technology and informatics , 44*, 87-94.

Pawlak, Z. (1998). Rough set theory and its applications to data analysis. *Cybernetics and Systems: An International Journal , 29* (7), 661-688.

Pawlak, Z. (1991). *Rough sets: theoretical aspects of reasoning about data* . KLUWER ACADEMIC PUBLISHER.

Pawlak, Z., & Slowinski, R. (1994). Decision Analysis Using Rough Sets. *International Transaction Operational Research , 1* (1), 107-114.

Peters, J. F., & Skowron, A. (2004). *Transactions on rough sets.* Berlin ; New York: Springer.

Petiot, F., & Yannou, B. (2004). Measuring consumer perceptions for a better comprehension, specification and assesment of product semantics. *International Jouranl of Ergonomics , 33*, 507-525.

Phillips, C. M. (2007). *The Influence of Aesthetics on Website User Perceptions.* Wichita: Psychology Department, Wichita State University.

Polkowski, L. (2002). *Rough set, Mathematical Foundation.* New York: Physica-Verlag Heidelberg.

Rokeach, M. (1973). *The Nature of Human Values* . New York: Free Press.

Scheaffer, R. M. (1996). *Elementary survey sampling* (Fifth edition ed.). USA: Duxbury Press.

Schütte, S. (2005). *Engineering Emotional Values in Product Design-Kansei Engineering in Development.* Linköping, Sweden: Linköping Studies in Science and Technology.

Schutte, S. T., Eklund, J., Axelsson, J. R., & Nagamachi, M. (2004). Concepts, methods and tool in Kansei engineering. *Theoretical issues in Ergonomic Sience , 15* (3), 214-231.

Shigekazu, I., Keiko, I., & Nagamachi, M. (1996). Neural networks approach for Kansei analysis on milk carton design. *Human factors in organizational design and managemet*, (pp. 7-12).

Shillito, M. L. (2001). *Acquiring, Processing, and deploying Voice of the Customer.* CRC Press LLC.

Shillito, M. L. (1994). *Advanced QFD Linking Technology to Market and Company Needs.* John Wiley& Sons, INC.

Shimohara, S., & Shimazaki, T. (2005). Evaluation of walking space by rough set model. *Proceeding of the Eastern Asia society for transportation studies , 5*, 2114-2122.

Shimuzu, Y., Sadoyama, T., Kamijo, M., Hasoya, S., Hashimoto, M., Otan, T., et al. (2004). On-demand production system of apparel on the basis of Kansei engineering. *International Journal of Clothing Science and Technology , 16* (1/2), 32-42.

Shutte, S., & Eklund, J. (2005). Design of rocker switches for work-vehicles-an application of Kansei engineering. *Applied ergonomics , 36*, 557-567.

Soler, F. P. (2005). *Statistics: Understanding Uncertainty* (2nd edition ed.). Associated Research Consultants.

Suraj, Z. (2004, Dec). An Introduction to Rough Set theory and its Application- A tutorial. *Proceeding of ICEN Conference* , 1-39.

Takagi, M., Watada, J., & Yubazaki, N. (2004). Realization of a Comfortable Space Based on Kansei engineering. *IEEE International Conference on System, Man and Cybernetics*, (pp. 6463-6464).

Tanoue, C., Ishizaka, K., & Nagamachi, M. (1997). Kansei engineering: A study of vehicle interior image. *International Journal of Industrial ergonomic , 19*, 115-128.

Ulwick, A. (2005). *What customers want:Using outcome-driven innovation to create breakthrough product and services.* McGraw-Hill.

Weinstein, A. (2004). *Handbook of Market Segmentation, Strategic Targeting for Business and Technology Firms.* New York: The Haworth Press.

Wortman, B., Richardson, W., Gee, G., Williams, W., Pearson, T., Bensley, F., et al. (2007). The Certified Six Sigma Black Belt Primer. Indiana, USA: Quality Council of Indiana.

Yanagisawa, H., & Fukuda, S. (2007). Design optimization for customers' KANSEI requirement: application of interactive reduct evolutional computation to industrial design with curves. *Int. J. Manufacturing Technology and Management , 10* (4), 360-379.

Yanagisawa, H., & Fukuda, S. (2005). Interactive Reduct Evolutional Computation for Aesthetic Design. *Journal of Computing and Information Science in Engineering , 5*, 1-7.

Zadeh, L., & Kacprzyk, J. (1992). *Fuzzy logic for the management of uncertainty.* New York: John Wiley & Sons, Inc.

Zimmermann, H.-J. (2000). An application-oriented view of modeling uncertainty. *European Journal of Operational Research , 122* (2), 190-198.

APPENDICES

APPENDIX A

GLOSSARY

Abstract knowledge: A family of concepts or categories.

Accuracy of approximation (or degree of consistency of the decision table or the degree of dependency between attributes C (condition attributes) and D (decision attributes) :

With every set $Y \subseteq U$, an accuracy of approximation of set Y by P could be associated as follows

$$\alpha_p(Y) = \frac{card(\underline{P}Y)}{card(\overline{P}Y)}$$

if $\alpha_p(Y) = 1$, Y is crisp or precise with respect to P; otherwise, Y is rough or vague.

An important property: If P\` is a reduct of P, then neither $\{a\} \Rightarrow \{b\}$, nor $\{b\} \Rightarrow \{a\}$ holds, for every $a, b \in$ P\`. This means all attributes in a reduct are pairwise independent.

Bayes' theorem: This mechanism provides a solution to the problem of how to learn from data. Let H = hypothesis, D = data, P (H) = probabilistic statement of belief about H before obtaining data D (what is known about H without knowledge of data-priori distribution of H), and P(H\D) = probabilistic statement of belief about H after obtaining data D (what is known about H given knowledge of data-posterior distribution of H given D). Then Bayes' theorem is as follows:

$$P(H|D) = P(D|H) \times P(H)/P(D)$$

Bayes' theorem and rough set notation: Let S = (U, A) be a decision table. With every $B \subseteq A = C \cup D$, there is associated a set of formulas F or (B), which are built up from the attribute-value pair (a, v), where $a \in B$ and $v \in V_a$ by means of logical connectives $\wedge(and), \vee(or), \backsim (not)$ in the standard way. For any $\phi \in F\ or\ (B)$ by $\|\phi\|_S$, the set of all objects $x \in U$ satisfying ϕ in S defined inductively is denoted as follows: $\|(a, v)\|_S = \{x \in U: a(v) = x\}, \forall \in B\ and\ v \in V_a$.

Category or concept: $X \subseteq U$ is called a concept or a category in U.

Classification or partition: Concepts can form a partition (classification) of a certain universe, i.e., in families C = $\{X_1, X_2, \ldots, X_n\}$, $X_i \subseteq U$, $X_i \neq \emptyset$, $X_i \cap X_i = \emptyset$, for i#j, I,j=1,…,n and $\cup X_i$=U.

Categories of vagueness in rough set: The four basic classes of rough sets or vagueness are as follows:

$$\underline{P}Y \neq \emptyset\ and\ \overline{P}Y \neq U, iff\ Y\ is\ roughly\ P\_definable$$

$$\underline{P}Y = \emptyset\ and\ \overline{P}Y \neq U, iff\ Y\ is\ internally\ P\_indifinable$$

$$\underline{P}Y \neq \emptyset \text{ and } \overline{P}Y = U, iff\ Y\ is\ externally\ P\_definable$$
$$\underline{P}Y = \emptyset \text{ and } \overline{P}Y = U, iff\ Y\ is\ totally\ P\_indefinable$$

The last class, which is important in this dissertation, implies to the case where it is not possible to decide for any element of U whether it belongs to Y or –Y, using P.

Core of attributes: A core set of attributes includes the most important attributes in which none of them can be removed without affecting the classification power of attributes. The core is the intersection of all reducts and is denoted by the following: Core (P) = $\bigcap Red(P)$

Certainty factor of decision rule (confidence coefficient in data mining): With every decision rule $\rightarrow \psi$, a conditional probability is associated as follows:

$$cer_S(\phi,\psi) = \pi_S\,(\psi|\phi) = PU(\|\psi\|_S|\|\phi\|_S) = \frac{|(\|\phi \wedge \psi\|_S)|}{|(\|\phi\|_S)|}$$

where $\|\phi\|_S \neq \emptyset$
If $cer_S(\phi,\psi) = \pi_S\,(\psi|\phi) = 1$, then $\phi \rightarrow \psi$ is called a certain decision; otherwise, the decision rule will be referred to as an uncertain decision rule in S.

Also, without using decision language,

$$cer_x(C,D) = \frac{|C(x) \cap D(x)|}{|C(x)|} = \frac{supp_S(C,D)}{|C(x)|} = \frac{\sigma_x(C,D)}{\pi(C(x))}$$

where $\pi(C(x)) = \frac{|C(x)|}{|U|}$

Coverage factor of the decision rule:

$$cov_S(\phi,\psi) = \pi_S\,(\phi|\psi) = PU(\|\phi\|_S|\|\psi\|_S) = \frac{|(\|\phi \wedge \psi\|_S)|}{|(\|\psi\|_S)|}$$

where $\|\psi\|_S \neq \emptyset$. Also, without using decision language,

$$cov_x(C,D) = \frac{|C(x) \cap D(x)|}{|D(x)|} = \frac{supp_S(C,D)}{|D(x)|} = \frac{\sigma_x(C,D)}{\pi(D(x))}$$

where $\pi(D(x)) = \frac{|D(x)|}{|U|}$

Decision rule: A decision rule in L(S) (decision language) is an expression $\phi \rightarrow \psi$, where $\phi \in F\ or\ (C), \psi \in F\ or\ (D)$ (C and D are condition and decision attributes, respectively).

Dependency of attributes: A set of attributes D depends totally on a set of attributes C, denoted by $C \Rightarrow D$, if all values of attributes from D are uniquely determined by values of attributes from C (or there exist a functional dependency between values of D and C). If only some values of D are determined by values of C, then the partial dependency exists. Therefore, if D depends upon C in a degree k(0 ≤ k ≤ 1), denoted by $C \underset{k}{\Rightarrow} D$, if k = $\gamma(C, D)$. The notation k = 1 implies total dependency of D on C, whereas k ≤ 1 denotes partial dependency of C on D. The coefficient k expresses the ratio of all elements of the universe, which can be properly classified to blocks of the partition U/D, employing attributes C. Therefore, dependency of attributes shows the consistency of the decision table. Also, if $C \underset{k}{\Rightarrow} D$, where k = 1, then $I(C) \subseteq I(D)$, which means that the partition generated by C is finer than the partition generated by D.

Equivalence classes of IND (P): Equivalence classes are called basic categories (or concepts) of knowledge P. For example, if old and ill are elementary categories in some knowledge base, then old and ill is a basic category in knowledge base.

Generalization vs. specialization: Let K = (U, P) and K' = (U, Q) be two knowledge bases. Knowledge P is finer than knowledge Q, or Q is coarser than P, if $IND(P) \subset IND(Q)$. Also, P is a specialization of Q, and Q is a generalization of P. For example, if P, Q ∈ R and both U/P and U/Q are classifications of the universe with respect to color, but the classification of U/Q contains one category of green objects, whereas the classification of U/P contains more categories of green objects, each referring to a specific shade of green, then P is a specialization of Q, and Q is a generalization of P, provided that every category of the shades of green in U/P is included in the category of green in U/Q.

Indiscernibility relation over P: If $P \subseteq R$ and $P \neq \emptyset$, then $\cap P$ (intersection of all equivalence relations belonging to P) is the indiscernibility relation over P. Moreover, $[x]_{IND(P)} = \bigcap_{R \in P} [x]_R$. In other words, for any selected subset of attributes P, there will be sets of objects that are indiscernible based on those attributes. These indistinguishable sets of objects define an equivalence or indiscernibility relation, referred to as the P-*indiscernibility relation* [http://en.wikipedia.org/wiki/Rough_set].

Information system: An information system is a table in which the rows represent objects (actions, alternatives, candidate, patients, etc.) and the columns show attributes. The entries of the table are attribute values or descriptors (Inuiguchi, Hirano, & Tsumoto, 2003). The information system is called a decision table if the set of attributes is divided into two subsets: condition attributes (criteria, tests, symptom, etc.) and decision attributes (decision, classification, taxonomies, etc.).

Indispensable and dispensable attribute: Let P $\subseteq$ Q and let α belong to P. Then α is dispensable in P, if I (P) = I (P - {α}); otherwise, α is indispensible in P.

Independent set: Set P is independent if all its attributes are indispensible.

Indispensable and dispensable values of attribute: The value of attribute $\alpha \in B$ is dispensable for x (an object in universe), if $P(x) = P^a(x)$, where $P^a = P - \{\alpha\}$; otherwise, the value of attribute α is indispensable for x.

Indiscernibility matrix: The indiscernibility matrix is denoted by M (P), and each entry of this matrix c_{ij} consists of all attributes that discern object x_i and x_j, and are defined as follows: $c_{ij} = \{a \in C | a(x_i) \neq a(x_j)\}$ i, j = 1, 2,…,m (where m represents the number of attributes). Reduct is the minimal subset of attributes that discerns all objects as discernible by the whole set of attributes.

Indiscernibility function: An indiscernibility function f_A for information system A is a Boolean of m Boolean variables $a^*_1,\ldots,a^*_m$(corresponding to attributes $a_1,\ldots,a_m$) defined as follows: $(a^*_1,\ldots,a^*_m) = \wedge \{\vee c^*_{ij} | 1 \leq j \leq i \leq n, c_{ij} \neq \emptyset\}$, where $c^*_{ij} = \{a^* | a \in c_{ij}\}$, and n is the number of objects.

Knowledge: Knowledge is a family of various classification patterns of a domain of interest, which provides explicit facts about reality.

Knowledge base over U: Knowledge base over U is a family of classifications over U (different basic classification, e.g., according to color, size, shape, etc.).

Quality of approximation (quality of sorting): The quality of approximation is k = $\gamma(C, D)$ and is called calculated by the following equation: $\gamma_p(У) = \frac{\sum_{i=1}^{n} card(\underline{P}Y_i)}{card(U)}$

The numerator of the above equation is the cardinality of the P-positive region of partition U/У.

Orthogonal set: If $\forall\ \alpha \in P, a(x)$ is dispensable (the value of α is dispensable) for x, then P will be called the orthogonal for x.

Positive region: $POS_P(У) = \cup_{Y_i \in У} \underline{P}\, Y_i$ is called the P-positive region of partition U/У (or U/I (У) or indiscernibility У) with respect to P and applies to the set of all objects of the universe U, which can be certainly classified as the member of each class of У = {Y_1, Y_2 …Y_n} using knowledge P.

R: R is the equivalence relation over U.

Reduct set of attributes: Subset *P`* of *P* is a reduct of *P* and shown by *Red (P)* if *P`* is independent and *I (P`) = I (P).* The reduced set of attributes provides the same quality of sorting as the original set of attributes $\gamma_P(У) = \gamma_{P`}(У)$. In other words, the reduct is the minimal subset of attributes that has the same power of classification of elements as the whole set of attributes.

Reduct set of values: Subset P`⊆ P is a value reduct of P for x, iff P` is orthogonal for x and P`(x) = P(x).

Reduction of Categories: Like reducing unnecessary knowledge by elimination of equivalence relations, which are redundant to define all basic categories in knowledge P (reducing the set of attributes), the same idea is used to eliminate unnecessary elementary categories to define the basic categories. So let $F = \{X_1, X_2, \ldots, X_n\}$, where $X_i \subset U$. X_i is dispensable in F, if $\cap(F - \{X_i\}) = \cap F$; otherwise, the set X_i is indispensible in F. If all components of F are indispensable, then F is independent; otherwise, it is dependent. The family $H \subseteq F$ is a reduct of F, if H is independent and $\cap H = \cap F$. The family of all indispensible set sets in F is called the core of F, denoted by *CORE (F).*

Set Approximation: The lower approximation of a set is the union of all granules that are entirely included in the set. In other words, let P $\subseteq$ Q and Y $\subseteq$ U. Then, the P (subset of condition attributes Q, or equivalence relation P, or knowledge P) lower approximation of Y, denoted by $\underline{P}Y$, is defined as follows:

$$\underline{P}Y = \bigcup \{X \epsilon U | P : X \subseteq Y\}$$

The upper approximation is the union of all granules that have non-empty intersections with the set. The P-upper approximation of Y, denoted by $\overline{P}Y$, is defined as follows:

$$\overline{P}Y = \bigcup \{X \epsilon U | P : X \cap Y \neq \emptyset\}$$

Boundary region: The boundary region of a set is the difference between the upper and lower approximation of the set. The P-boundary (doubtful region) of set Y is defined as follows:
B_{np} (Y) = $\overline{P}Y - \underline{P}Y$. Set Y is crisp or exact with respect to P, if the boundary region of Y is empty; otherwise, set Y is rough or inexact with respect to P. In other words, $\underline{P}Y \# \overline{P}Y$.

Vagueness and uncertainty: Vagueness is the property of a set and can be explained by approximations, whereas uncertainty is the property of elements of a set and can be articulated by the rough membership function.

Some properties of approximations:

$$\underline{P}Y \subseteq Y \subseteq \overline{P}Y$$
$$\underline{P}\emptyset = \overline{P}\emptyset = \emptyset, \overline{P}U = \underline{P}U = U$$
$$\overline{P}(X \cup Y) = \overline{P}X \cup \overline{P}Y$$
$$\underline{P}(X \cap Y) = \underline{P}X \cap \underline{P}Y$$
$$X \subseteq Y\ implies \underline{P}X \subseteq \underline{P}Y\ and\ \overline{P}X \subseteq \overline{P}Y$$
$$\underline{P}(X \cup Y) \supseteq \underline{P}X \cup \underline{P}Y$$
$$\overline{P}(X \cap Y) \subseteq \overline{P}X \cap \overline{P}Y$$

Significance of attribute: The significance of an attribute can be measured by removing the attribute from an information system table and calculating it as follows:

$$\sigma_{(C,D)}(\alpha) = \frac{\gamma(C,D) - \gamma(C - \{a\}, D)}{\gamma(C,D)} = 1 - \frac{\gamma(C - \{a\}, D)}{\gamma(C,D)}$$

Obviously, $0 \leq \sigma_{(C,D)}(\alpha) \leq 1$, and the greater the number $\sigma_{(C,D)}(\alpha)$ the more important the attribute α. Also, the coefficient $\sigma_{(C,D)}(\alpha)$ can be considered an error that occurs when attribute α is dropped.

Significant of the set of attributes: If $B \subset C$ (C and D stand for set of condition attributes and decision attributes), then $\varepsilon(C, D)(B)$, a significant coefficient of the set of attributes or error of reduct approximation, and is calculated as follows: $\varepsilon(C, D)(B) = \frac{\gamma(C,D) - \gamma(C - B, D)}{\gamma(C,D)} = 1 - \frac{\gamma(C - B, D)}{\gamma(C,D)}$
If B is a reduct of C, then $\varepsilon(C, D)(B)$ = 1, which means that removing any reduct from a set of attributes will enable us to make certain decisions.

Any subset B of C is called an approximate reduct of C, and the error of reduct approximation, which is denoted as $\varepsilon(C,D)(B)$, is calculated as follows: $\varepsilon(C,D)(B)=\frac{\gamma(C,D)-\gamma(B,D)}{\gamma(C,D)}=1-\frac{\gamma(B,D)}{\gamma(C,D)}$
This measure tells us exactly how the set of attributes B approximates the set of condition attributes C. For any reduct set of C, $\varepsilon(C,D)(B) = 0$.

<u>Support of the decision rule</u>:

$$supp_s(\phi,\psi) = |(\|\phi\wedge\psi\|_s)|$$

The number is called support of the decision rule $\phi \rightarrow \psi$ in S.
Also, without using decision language, by referring only to the decision table, the number is called support of the decision rule $C \rightarrow_x D$

$$supp_x(C,D) = |C(x) \cap D(x)|$$

Where S= (U,C,D) is a decision table, and $x \in U$ determines a sequence $c_1(x), c_2(x), c_3(x), \dots, c_n(x), d_1(x), d_2(x), d_3(x), \dots, d_m(x)$, where $\{c_1, c_2, c_3, \dots, c_n\} = C$ and $\{d_1, d_2, d_3, \dots, d_n\} = D$.

<u>Strength of the decision rule</u>:

$$\sigma_s(\phi,\psi) = \frac{supp_s(\phi,\psi)}{|U|} = \pi_s(\psi|\phi) \times \pi_s(\phi)$$

Without using decision language:

$$\sigma_x(C,D) = \frac{supp_s(C,\ D)}{|U|}$$

<u>Universe</u>: $U \neq \emptyset$ is a finite set of objects.

<u>U/R</u>: U/R refers to the family of equivalence classes of R (or classification of U) or categories or concept of R. Also, $[x]_R$ denotes a category in R containing an element $x \in U$.

<u>U/IND (P) or U/P</u>: This is the family of all equivalence classes of the equivalence relation IND (P) and is called P-basic knowledge or basic knowledge about U in K = (U,R) (relational system).

APPENDIX B

SUPPLEMENTAL RESULTS TABLES

TABLE B.1 COMMON REDUCT SETS—INTERESTING

Reduct	User Characteristics									
	STUCOM	Gender	FREQ_USE	PENTER	PWORK	PCOMM	Skill	Hedonic	Involvement	CVTER
Group 1										
Int-20	1	1		1	1	1	1	1	1	1
Int-22	1	1		1	1	1	1	1	1	1
Int-30	1	1		1	1	1	1	1	1	1
Int-32	1	1		1	1	1	1	1	1	1
Int-41	1	1		1	1	1	1	1	1	1
Group 2										
Int-1	1	1	1	1	1	1	1	1	1	1
Int-15	1	1	1	1	1	1	1	1	1	1
Int-23	1	1	1	1	1	1	1	1	1	1
Int-8	1	1	1	1	1	1	1	1	1	1
Group 3										
Int-28	1	1		1		1	1		1	1
Int-46	1	1		1		1	1		1	1
Int-9	1	1		1		1	1		1	1
Group 4										
Int-29	1	1	1	1		1	1	1	1	1
Int-30	1	1	1	1		1	1	1	1	1
Int-5	1	1	1	1		1	1	1	1	1

TABLE B.2 COMMON REDUCT SETS—IMAGINATIVE

Reduct	User Characteristics									
	STUCOM	Gender	FREQ_USE	PENTER	PWORK	PCOMM	Skill	Involvement	CVTER	Hedonic
Group 1										
Im-13	1	1		1	1	1	1	1	1	1
Im-28	1	1		1	1	1	1	1	1	1
Im-30	1	1		1	1	1	1	1	1	1
Im-32	1	1		1	1	1	1	1	1	1
Im-5	1	1		1	1	1	1	1	1	1
Group 2										
Im-1	1	1	1	1	1	1	1	1	1	
Im-11	1	1	1	1	1	1	1	1	1	
Im-23	1	1	1	1	1	1	1	1	1	
Im-8	1	1	1	1	1	1	1	1	1	
Group 3										
Im-30	1	1	1	1		1	1	1	1	1
Im-33	1	1	1	1		1	1	1	1	1
Im-5	1	1	1	1		1	1	1	1	1
Im-8	1	1	1	1		1	1	1	1	1

TABLE B.3 USER GROUPING RULES FOR KANSEI “INTERESTING”

Adj	#	Match	Strength	1	2	Cer1	Cer2	Cov1	Cov2	# of Char	STUCOM	PWORK	PCOMM	Hedonic	Involvement
Group 1 (IntG1)															
Int	11	6	0.1		6	0	1	0	0.12	3	2	2			1
Int	17	6	0.1		6	0	1	0	0.11	3	2	2			1
Int	22	6	0.1	6		1	0	0.18	0	3	2	2			1
Int	24	6	0.1		6	0	1	0	0.11	3	2	2			1
Int	41	6	0.1	6		1	0	0.11	0	3	2	2			1
Int	46	6	0.1	6		1	0	0.12	0	3	2	2			1
Int	50	6	0.1	6		1	0	0.13	0	3	2	2			1
Group 2 (IntG2)															
Int	11	6	0.1		6	0	1	0	0.12	2			2	2	
Int	13	6	0.1		6	0	1	0	0.13	2			2	2	
Int	24	6	0.1		6	0	1	0	0.11	2			2	2	
Int	29	6	0.1	6		1	0	0.13	0	2			2	2	
Int	30	6	0.1	6		1	0	0.14	0	2			2	2	
Int	5	6	0.1	6		1	0	0.14	0	2			2	2	

TABLE B.4 USERS GROUPING RULES FOR KANSEI “IMAGINATIVE”

Adj	#	Match	Strength	1	2	Cer1	Cer2	Cov1	Cov2	# of Characteristics	Age	FREQ_USE	PENTER	PWORK	PCOMM	Skill	Utility	Hedonic
Group1(ImG1)																		
Im	11	6	0.1		6	0	1	0	0.12	2				1				2
Im	13	6	0.1		6	0	1	0	0.13	2				1				2
Im	17	6	0.1		6	0	1	0	0.11	2				1				2
Im	24	6	0.1		6	0	1	0	0.11	2				1				2
Im	34	6	0.1	6		1	0	0.11	0	2				1				2
Im	35	6	0.1	6		1	0	0.11	0	2				1				2
Im	38	6	0.1	6		1	0	0.11	0	2				1				2
Im	46	6	0.1	6		1	0	0.13	0	2				1				2
Im	9	6	0.1		6	0	1	0	0.13	2				1				2
Group 2 (ImG1)																		
Im	15	6	0.1		6	0	1	0	0.15	2	1							2
Im	17	6	0.1		6	0	1	0	0.11	2	1							2
Im	24	6	0.1		6	0	1	0	0.11	2	1							2
Im	28	6	0.1	6		1	0	0.12	0	2	1							2
Im	33	6	0.1	6		1	0	0.14	0	2	1							2
Im	38	6	0.1	6		1	0	0.11	0	2	1							2
Im	5	6	0.1	6		1	0	0.13	0	2	1							2
Im	50	6	0.1	6		1	0	0.14	0	2	1							2
Group 3 (ImG1)																		
Im	1	7	0.11		7	0	1	0	0.15	2					1	2		
Im	15	7	0.11		7	0	1	0	0.17	2					1	2		
Im	24	7	0.11		7	0	1	0	0.13	2					1	2		
Im	29	7	0.11	7		1	0	0.15	0	2					1	2		
Im	35	7	0.11	7		1	0	0.13	0	2					1	2		
Im	41	7	0.11	7		1	0	0.13	0	2					1	2		
Im	9	7	0.11		7	0	1	0	0.15	2					1	2		

TABLE B.5 COMMON REDUCT SET OF WEB FEATURES—DIFFERENT GROUP(S) OF CUSTOMERS—INTERESTING

Reduct	Font Size	Font Contrast	Font Color	Page Layout Color Combination	Image Size	Image Quality Resolution	Image Content	Picture	Page Layout Density	Line Length	Display Color
C 1											
IntG1		1						1			
IntG2		1						1			
C 2											
IntG1										1	
IntG2										1	
C 3											
IntG1								1			1
IntG2								1			1
C 4											
IntG1	1	1									

TABLE B.6 COMMON REDUCT SET OF WEB FEATURES—DIFFERENT GROUP(S) OF CUSTOMERS—IMAGINATIVE

Reduct	Image Size	Font Size	Font Color	Font Contrast	Image Content	Page Layout Density	Page Layout Color Combination	Picture	Display Color	Font Type	Display Background Graphic	Image Quality Resolution	Line Length
C1													
ImG1												1	
ImG3												1	
ImG4												1	
ImG5												1	
C 2													
ImG1	1												
ImG2	1												
ImG5	1												
C 3													
ImG2								1		1			
ImG4								1		1			
ImG5								1		1			
C 4													
ImG2										1			1
ImG4										1			1
ImG5										1			1
C 5													
ImG1		1		1				1	1				
ImG3		1		1				1	1				

TABLE B.7 DESIGN RULES FOR DIFFERENT USERS GROUPS AND FOR "INTERESTING"

										Web features											
Clusters (C)	Match	Strength	Interesting	Not Interesting	Certainty1	Certainty2	Coverage1	Coverage2	# of Characteristic	Font Size	Font Color	Font Type	Font Contrast	Image Size	Image Quality Resolution	Image Content	Page Layout Density	Page layout Color Combination	Picture	Line Length	Display Color
Cluster 1																					
IntG1	4	0.57	4		1	0	1	0	1						2						
IntG2	2	0.33	2		1	0	0.67	0	1						2						
Cluster 2																					
IntG1	4	0.57	4		1	0	1	0	1								2				
IntG2	1	0.17		1	0	1	0	0.33	1								2				
Group 3																					
IntG1	4	0.57	4		1	0	1	0	1											1	
IntG2	3	0.5	3		1	0	1	0	1											1	
Group 4																					
IntG1	4	0.57	4		1	0	1	0	2									1			3
IntG2	1	0.17	1		1	0	0.33	0	2									1			3
Group 5																					
IntG1	3	0.43	3		1	0	0.75	0	1							1					
IntG2	1	0.17	1		1	0	0.33	0	1							1					
Group 6																					
IntG1	3	0.43	3		1	0	0.75	0	2	1								1			
IntG2	1	0.17	1		1	0	0.33	0	2	1								1			
Group 7																					
IntG1	3	0.43		3	0	1	0	1	1											2	
IntG2	3	0.5		3	0	1	0	1	1											2	
Group 8																					
IntG1	2	0.29	2		1	0	0.5	0	1		1										
IntG2	1	0.17	1		1	0	0.33	0	1		1										
Group 9																					
IntG1	2	0.29		2	0	1	0	0.67	1										1		
IntG2	2	0.33		2	0	1	0	0.67	1										1		
Group 10																					
IntG1	1	0.14	1		1	0	0.25	0	2										3		3
IntG2	3	0.5	3		1	0	1	0	2										3		3
Group 11																					
IntG1	1	0.14		1	0	1	0	0.33	1												1
IntG2	1	0.17		1	0	1	0	0.33	1												1

TABLE B.8 DESIGN RULES FOR DIFFERENT USERS GROUPS AND FOR “IMAGINATIVE”

Cluster (C)	Match	Strength	Imaginative	Not Imaginative	Certainty1	Certainty2	Coverage1	Coverage2	# of Characteristic	Font Size	Font Color	Font Type	Font Contrast	Image Size	Image Quality Resolution	Image Content	Page Layout Density	Page layout Color Combination	Picture	Line length
Group 1																				
ImG1	5	0.56		5	0	1	0	1	1											2
ImG2	2	0.25		2	0	1	0	0.67	1											2
ImG3	2	0.29		2	0	1	0	0.5	1											2
ImG4	2	0.29		2	0	1	0	0.67	1											2
ImG5	2	0.33		2	0	1	0	0.67	1											2
Group 2																				
ImG1	4	0.44	4		1	0	1	0	1						2					
ImG2	4	0.5	4		1	0	0.8	0	1						2					
ImG3	3	0.43	3		1	0	1	0	1						2					
ImG4	4	0.57	4		1	0	1	0	1						2					
ImG5	3	0.5	3		1	0	1	0	1						2					
Group 3																				
ImG1	5	0.56		5	0	1	0	1	1						1					
ImG3	4	0.57		4	0	1	0	1	1						1					
ImG4	3	0.43		3	0	1	0	1	1						1					
ImG5	3	0.5		3	0	1	0	1	1						1					
Group 4																				
ImG1	4	0.44	4		1	0	1	0	1					3						
ImG2	3	0.38	3		1	0	0.6	0	1					3						
ImG4	3	0.43	3		1	0	0.75	0	1					3						
ImG5	3	0.5	3		1	0	1	0	1					3						
Group 5																				
ImG1	2	0.22	2		1	0	0.5	0	1										2	
ImG2	2	0.25	2		1	0	0.4	0	1										2	
ImG4	2	0.29	2		1	0	0.5	0	1										2	
ImG5	2	0.33	2		1	0	0.67	0	1										2	
Group 6																				
ImG2	4	0.5	4		1	0	0.8	0	1									1		
ImG3	2	0.29	2		1	0	0.67	0	1									1		
ImG4	3	0.43	3		1	0	0.75	0	1									1		
ImG5	3	0.5	3		1	0	1	0	1									1		
Group 7																				
ImG2	2	0.25	2		1	0	0.4	0	1								2			
ImG3	2	0.29	2		1	0	0.67	0	1								2			
ImG4	3	0.43	3		1	0	0.75	0	1								2			
ImG5	3	0.5	3		1	0	1	0	1								2			

TABLE B.9 DESIGN RULES FOR DIFFERENT USERS GROUPS AND FOR MULTIPLE KANSEIS

Clusters (C)	Match	Strength	Attractive	Not Attractive	Certainty1	Certainty2	Coverage1	Coverage2	# of Characteristic	Font Size	Font Color	Font Type	Font Contrast	Image Size	Image Quality	Image Content	Page Layout Density	Page layout Color	Picture	Line Length	Display Color
Cluster 1																					
AttG1	7	0.64	7		1	0	1	0	1						2						
AttG2	5	0.5	5		1	0	1	0	1						2						
AttG3	4	0.5	4		1	0	1	0	1						2						
AttG4	3	0.38	3		1	0	1	0	1						2						
AttG5	4	0.57	4		1	0	1	0	1						2						
ImG1	4	0.44	4		1	0	1	0	1						2						
ImG2	4	0.5	4		1	0	0.8	0	1						2						
ImG3	3	0.43	3		1	0	1	0	1						2						
ImG4	4	0.57	4		1	0	1	0	1						2						
ImG5	3	0.5	3		1	0	1	0	1						2						
IntG1	4	0.57	4		1	0	1	0	1						2						
IntG2	2	0.33	2		1	0	0.67	0	1						2						
Cluster 2																					
AttG1	3	0.27		3	0	1	0	0.75	1											2	
AttG2	4	0.4		4	0	1	0	0.8	1											2	
AttG3	3	0.38		3	0	1	0	0.75	1											2	
AttG4	3	0.38		3	0	1	0	0.6	1											2	
AttG5	2	0.29		2	0	1	0	0.67	1											2	
ImG1	5	0.56		5	0	1	0	1	1											2	
ImG2	2	0.25		2	0	1	0	0.67	1											2	
ImG3	2	0.29		2	0	1	0	0.5	1											2	
ImG4	2	0.29		2	0	1	0	0.67	1											2	
ImG5	2	0.33		2	0	1	0	0.67	1											2	
IntG1	3	0.43		3	0	1	0	1	1											2	
IntG2	3	0.5		3	0	1	0	1	1											2	
Cluster 3																					
AttG1	4	0.36		4	0	1	0	1	1						1						
AttG2	5	0.5		5	0	1	0	1	1						1						
AttG3	4	0.5		4	0	1	0	1	1						1						
AttG4	5	0.63		5	0	1	0	1	1						1						
AttG5	3	0.43		3	0	1	0	1	1						1						
ImG1	5	0.56		5	0	1	0	1	1						1						
ImG3	4	0.57		4	0	1	0	1	1						1						
ImG4	3	0.43		3	0	1	0	1	1						1						
ImG5	3	0.5		3	0	1	0	1	1						1						
IntG1	3	0.43		3	0	1	0	1	1						1						
Cluster 4																					
AttG1	4	0.36	4		1	0	0.57	0	1					3							
AttG2	2	0.2	2		1	0	0.4	0	1					3							
AttG3	3	0.38	3		1	0	0.75	0	1					3							
AttG5	3	0.43	3		1	0	0.75	0	1					3							
ImG1	4	0.44	4		1	0	1	0	1					3							
ImG2	3	0.38	3		1	0	0.6	0	1					3							
ImG4	3	0.43	3		1	0	0.75	0	1					3							
ImG5	3	0.5	3		1	0	1	0	1					3							
IntG1	3	0.43	3		1	0	0.75	0	1					3							

TABLE B.9 (continued)

Clusters (C)	Match	Strength	Attractive	Not Attractive	Certainty1	Certainty2	Coverage1	Coverage2	# of Characteristic	Font Size	Font Color	Font Type	Font Contrast	Image Size	Image Quality	Image Content	Page Layout Density	Page layout Color	Picture	Line Length	Display Color
Cluster 5																					
AttG1	4	0.36	4		1	0	0.57	0	1								2				
AttG3	2	0.25	2		1	0	0.5	0	1								2				
AttG4	2	0.25	2		1	0	0.67	0	1								2				
AttG5	3	0.43	3		1	0	0.75	0	1								2				
ImG2	2	0.25	2		1	0	0.4	0	1								2				
ImG3	2	0.29	2		1	0	0.67	0	1								2				
ImG4	3	0.43	3		1	0	0.75	0	1								2				
ImG5	3	0.5	3		1	0	1	0	1								2				
IntG1	4	0.57	4		1	0	1	0	1								2				
Cluster 9																					
AttG1	3	0.27		3	0	1	0	0.75	1									2			
AttG2	4	0.4		4	0	1	0	0.8	1									2			
AttG3	3	0.38		3	0	1	0	0.75	1									2			
AttG5	3	0.43		3	0	1	0	1	1									2			
ImG1	4	0.44		4	0	1	0	0.8	1									2			
ImG5	3	0.5		3	0	1	0	1	1									2			
IntG1	2	0.29		2	0	1	0	0.67	1									2			
Cluster 10																					
AttG2	3	0.3		3	0	1	0	0.6	1										1		
AttG4	2	0.25		2	0	1	0	0.4	1										1		
ImG1	3	0.33		3	0	1	0	0.6	1										1		
ImG2	2	0.25		2	0	1	0	0.67	1										1		
IntG1	2	0.29		2	0	1	0	0.67	1										1		
IntG2	2	0.33		2	0	1	0	0.67	1										1		
Cluster 11																					
AttG3	2	0.25	2		1	0	0.5	0	1										2		
ImG1	2	0.22	2		1	0	0.5	0	1										2		
ImG2	2	0.25	2		1	0	0.4	0	1										2		
ImG4	2	0.29	2		1	0	0.5	0	1										2		
ImG5	2	0.33	2		1	0	0.67	0	1										2		
IntG1	3	0.43	3		1	0	0.75	0	1										2		
Cluster 14																					
AttG1	5	0.45	5		1	0	0.71	0	2									1			3
AttG2	5	0.5	5		1	0	1	0	2									1			3
AttG3	4	0.5	4		1	0	1	0	2									1			3
ImG1	2	0.22	2		1	0	0.5	0	2									1			3
IntG1	4	0.57	4		1	0	1	0	2									1			3
Cluster 15																					
AttG2	2	0.2	2		1	0	0.4	0	1		1										
AttG3	2	0.25	2		1	0	0.5	0	1		1										
AttG5	2	0.29	2		1	0	0.5	0	1		1										
ImG2	3	0.38	3		1	0	0.6	0	1		1										
IntG1	2	0.29	2		1	0	0.5	0	1		1										

TABLE B.9 (continued)

Clusters (C)	Match	Strength	Attractive	Not Attractive	Certainty1	Certainty2	Coverage1	Coverage2	# of Characteristic	Font Size	Font Color	Font Type	Font Contrast	Image Size	Image Quality	Image Content	Page Layout	Page layout Color	Picture	Line Length	Display Color
Cluster 17																					
AttG1	4	0.36	4		1	0	0.57	0	2	1								1			
AttG2	2	0.2	2		1	0	0.4	0	2	1								1			
ImG1	2	0.22	2		1	0	0.5	0	2	1								1			
IntG1	3	0.43	3		1	0	0.75	0	2	1								1			
Cluster 22																					
AttG2	4	0.4		4	0	1	0	0.8	2		2					2					
AttG3	3	0.38		3	0	1	0	0.75	2		2					2					
ImG2	3	0.38		3	0	1	0	1	2		2					2					
IntG1	3	0.43		3	0	1	0	1	2		2					2					
Cluster 26																					
AttG5	3	0.43	3		1	0	0.75	0	1				2								
ImG4	3	0.43	3		1	0	0.75	0	1				2								
ImG5	3	0.5	3		1	0	1	0	1				2								
IntG2	2	0.33		2	0	1	0	0.67	1				2								
Cluster 33																					
AttG2	2	0.2	2		1	0	0.4	0	1					2							
ImG2	2	0.25	2		1	0	0.4	0	1					2							
IntG2	2	0.33	2		1	0	0.67	0	1					2							
Cluster 36																					
AttG4	2	0.25		2	0	1	0	0.4	2	2	2										
ImG3	2	0.29		2	0	1	0	0.5	2	2	2										
IntG1	2	0.29		2	0	1	0	0.67	2	2	2										

APPENDIX C

REGRESSION ANALYSIS

Regression Analysis of Some Websites That Showed Relationship in Phillips Study

Multiple Regressions—attrac11

Dependent variable: attrac11

Independent variables: S_N_S, Gender, Age, Freq, Enter, Work, Com, Skill, Utility, Hedonic, Involvement, CVPA

Parameter	*Estimate*	*Standard Error*	*T Statistic*	*P-Value*
CONSTANT	4.1801	1.42472	2.93399	0.0049
S_N_S	2.94065	0.653797	4.49781	0.0000
Age	-0.88386	0.273362	-3.23329	0.0021
Enter	1.73294	0.662703	2.61497	0.0115
Involv	-0.66792	0.234432	-2.8491	0.0062

Analysis of Variance

Source	Sum of Squares	Df	Mean Square	F-Ratio	P-Value
Model	112.034	4	28.0086	6.70	0.0002
Residual	225.694	54	4.17952		
Total (Corr.)	337.729	58			

R-squared = 33.1729 percent

R-squared (adjusted for d.f.) = 28.2228 percent

Standard Error of Est. = 2.04439

Mean absolute error = 1.53845

Durbin-Watson statistic = 2.02408 (P=0.4627)

Lag 1 residual autocorrelation = -0.0294523

Stepwise Regression

Method: backward selection

F-to-enter: 4.0

F-to-remove: 4.0

Step 0:

12 variables in the model. 46 d.f. for error.

R-squared = 36.71% Adjusted R-squared = 20.20% MSE = 4.6469

Step 1:

Removing variable Skill with F-to-remove =0.0157421

11 variables in the model. 47 d.f. for error.

R-squared = 36.69% Adjusted R-squared = 21.87% MSE = 4.54958

Step 2:

Removing variable Hedonic with F-to-remove =0.0333743

10 variables in the model. 48 d.f. for error.

R-squared = 36.64% Adjusted R-squared = 23.44% MSE = 4.45796

Step 3:

Removing variable Gender with F-to-remove =0.104333

9 variables in the model. 49 d.f. for error.

R-squared = 36.50% Adjusted R-squared = 24.84% MSE = 4.37648

Step 4:

Removing variable Com with F-to-remove =0.115823

8 variables in the model. 50 d.f. for error.

R-squared = 36.35% Adjusted R-squared = 26.17% MSE = 4.29908

Step 5:

Removing variable Freq with F-to-remove =0.071782

7 variables in the model. 51 d.f. for error.

R-squared = 36.26% Adjusted R-squared = 27.51% MSE = 4.22084

Step 6:

Removing variable CVPA with F-to-remove =0.144462

6 variables in the model. 52 d.f. for error.

R-squared = 36.08% Adjusted R-squared = 28.71% MSE = 4.15139

Step 7:

Removing variable Work with F-to-remove =0.588552

5 variables in the model. 53 d.f. for error.

R-squared = 35.36% Adjusted R-squared = 29.26% MSE = 4.11917

Step 8:

Removing variable Utility with F-to-remove =1.79126

4 variables in the model. 54 d.f. for error.

R-squared = 33.17% Adjusted R-squared = 28.22% MSE = 4.17952

Final model selected.

The output shows the results of fitting a multiple linear regression model to describe the relationship between attrac11 and 12 independent variables. The equation of the fitted model is

attrac11 = 4.1801 + 2.94065*S_N_S - 0.88386*Age + 1.73294*Enter - 0.66792*Involv

Since the P-value in the ANOVA table is less than 0.05, there is a statistically significant relationship between the variables at the 95% confidence level.

Multiple Regressions—attrac13

Dependent variable: attrac13

Independent variables: S_N_S, Gender, Age, Freq, Enter, Work, Com, Skill, Utility, Hedonic, Involv, CVPA

Parameter	*Estimate*	*Standard Error*	*T Statistic*	*P-Value*
CONSTANT	9.46439	1.18217	8.00596	0.0000
Enter	-1.70586	0.554033	-3.07899	0.0032
Work	-2.01258	0.595168	-3.38153	0.0013
Utility	1.04724	0.420693	2.48932	0.0157

Analysis of Variance

Source	Sum of Squares	Df	Mean Square	F-Ratio	P-Value
Model	99.8868	3	33.2956	7.81	0.0002
Residual	247.355	58	4.26474		
Total (Corr.)	347.242	61			

R-squared = 28.7658 percent

R-squared (adjusted for d.f.) = 25.0812 percent

Standard Error of Est. = 2.06513

Mean absolute error = 1.65566

Durbin-Watson statistic = 1.94587 (P=0.3901)

Lag 1 residual autocorrelation = 0.00852537

Stepwise Regression

Method: backward selection

F-to-enter: 4.0

F-to-remove: 4.0

Step 0:

12 variables in the model. 49 d.f. for error.

R-squared = 36.27% Adjusted R-squared = 20.66% MSE = 4.51629

Step 1:

Removing variable Hedonic with F-to-remove =0.00089654

11 variables in the model. 50 d.f. for error.

R-squared = 36.27% Adjusted R-squared = 22.25% MSE = 4.42604

Step 2:

Removing variable Involv with F-to-remove =0.00439783

10 variables in the model. 51 d.f. for error.

R-squared = 36.26% Adjusted R-squared = 23.77% MSE = 4.33964

Step 3:

Removing variable Freq with F-to-remove =0.249099

9 variables in the model. 52 d.f. for error.

R-squared = 35.95% Adjusted R-squared = 24.87% MSE = 4.27697

Step 4:

Removing variable Skill with F-to-remove =0.192912

8 variables in the model. 53 d.f. for error.

R-squared = 35.71% Adjusted R-squared = 26.01% MSE = 4.21184

Step 5:

Removing variable Com with F-to-remove =0.35277

7 variables in the model. 54 d.f. for error.

R-squared = 35.29% Adjusted R-squared = 26.90% MSE = 4.16136

Step 6:

Removing variable Gender with F-to-remove =0.46709

6 variables in the model. 55 d.f. for error.

R-squared = 34.73% Adjusted R-squared = 27.61% MSE = 4.12104

Step 7:

Removing variable S_N_S with F-to-remove =1.8156

5 variables in the model. 56 d.f. for error.

R-squared = 32.57% Adjusted R-squared = 26.55% MSE = 4.18106

Step 8:

Removing variable Age with F-to-remove =0.733027

4 variables in the model. 57 d.f. for error.

R-squared = 31.69% Adjusted R-squared = 26.90% MSE = 4.16148

Step 9:

Removing variable CVPA with F-to-remove =2.43926

3 variables in the model. 58 d.f. for error.

R-squared = 28.77% Adjusted R-squared = 25.08% MSE = 4.26474

Final model selected.

The output shows the results of fitting a multiple linear regression model to describe the relationship between attrac13 and 12 independent variables. The equation of the fitted model is

attrac13 = 9.46439 - 1.70586*Enter - 2.01258*Work + 1.04724*Utility

Since the P-value in the ANOVA table is less than 0.05, there is a statistically significant relationship between the variables at the 95% confidence level.

Multiple Regressions—attrac22

Dependent variable: attrac22

Independent variables: S_N_S, Gender, Age, Freq, Enter, Work, Com, Skill, Utility, Hedonic, Involv, CVPA

		Standard	*T*	
Parameter	*Estimate*	*Error*	*Statistic*	*P-Value*
CONSTANT	4.41644	1.15957	3.8087	0.0003
Age	-0.889701	0.216696	-4.10576	0.0001
Involv	0.525539	0.236597	2.22124	0.0303
CVPA	0.690964	0.343855	2.00947	0.0491

Analysis of Variance

Source	Sum of Squares	Df	Mean Square	F-Ratio	P-Value
Model	119.784	3	39.9279	8.29	0.0001
Residual	279.2	58	4.8138		
Total (Corr.)	398.984	61			

R-squared = 30.0222 percent

R-squared (adjusted for d.f.) = 26.4026 percent

Standard Error of Est. = 2.19404

Mean absolute error = 1.7591

Durbin-Watson statistic = 2.25968 (P=0.8348)

Lag 1 residual autocorrelation = -0.187384

<u>Stepwise Regression</u>

Method: backward selection

F-to-enter: 4.0

F-to-remove: 4.0

<u>Step 0:</u>

12 variables in the model. 49 d.f. for error.

R-squared = 38.65% Adjusted R-squared = 23.62% MSE = 4.99573

<u>Step 1:</u>

Removing variable Skill with F-to-remove =0.000261305

11 variables in the model. 50 d.f. for error.

R-squared = 38.65% Adjusted R-squared = 25.15% MSE = 4.89584

<u>Step 2:</u>

Removing variable Com with F-to-remove =0.0666652

10 variables in the model. 51 d.f. for error.

R-squared = 38.56% Adjusted R-squared = 26.52% MSE = 4.80625

Step 3:

Removing variable Work with F-to-remove =0.053127

9 variables in the model. 52 d.f. for error.

R-squared = 38.50% Adjusted R-squared = 27.86% MSE = 4.71873

Step 4:

Removing variable Freq with F-to-remove =0.215117

8 variables in the model. 53 d.f. for error.

R-squared = 38.25% Adjusted R-squared = 28.92% MSE = 4.64885

Step 5:

Removing variable S_N_S with F-to-remove =0.676362

7 variables in the model. 54 d.f. for error.

R-squared = 37.46% Adjusted R-squared = 29.35% MSE = 4.62099

Step 6:

Removing variable Hedonic with F-to-remove =0.841792

6 variables in the model. 55 d.f. for error.

R-squared = 36.48% Adjusted R-squared = 29.55% MSE = 4.60769

Step 7:

Removing variable Enter with F-to-remove =1.71339

5 variables in the model. 56 d.f. for error.

R-squared = 34.50% Adjusted R-squared = 28.66% MSE = 4.66639

Step 8:

Removing variable Utility with F-to-remove =1.58146

4 variables in the model. 57 d.f. for error.

R-squared = 32.65% Adjusted R-squared = 27.93% MSE = 4.71399

Step 9:

Removing variable Gender with F-to-remove =2.22795

3 variables in the model. 58 d.f. for error.

R-squared = 30.02% Adjusted R-squared = 26.40% MSE = 4.8138

Final model selected.

The output shows the results of fitting a multiple linear regression model to describe the relationship between attrac22 and 12 independent variables. The equation of the fitted model is

attrac22 = 4.41644 - 0.889701*Age + 0.525539*Involv + 0.690964*CVPA

Since the P-value in the ANOVA table is less than 0.05, there is a statistically significant relationship between the variables at the 95% confidence level.

Multiple Regressions—attrac24

Dependent variable: attrac24

Independent variables: S_N_S, Gender, Age, Freq, Enter, Work, Com, Skill, Utility, Hedonic, Involv, CVPA

		Standard	*T*	
Parameter	*Estimate*	*Error*	*Statistic*	*P-Value*
CONSTANT	7.19257	1.34663	5.34117	0.0000
S_N_S	1.50162	0.57867	2.59496	0.0121
Age	-0.461733	0.214418	-2.15343	0.0356
Freq	1.08358	0.413447	2.62085	0.0113
Work	-1.09079	0.525628	-2.07521	0.0426

Analysis of Variance

Source	Sum of Squares	Df	Mean Square	F-Ratio	P-Value
Model	76.7888	4	19.1972	5.44	0.0009
Residual	197.473	56	3.52631		
Total (Corr.)	274.262	60			

R-squared = 27.9983 percent

R-squared (adjusted for d.f.) = 22.8553 percent

Standard Error of Est. = 1.87785

Mean absolute error = 1.26161

Durbin-Watson statistic = 2.21134 (P=0.7337)

Lag 1 residual autocorrelation = -0.107757

Stepwise Regression

Method: backward selection

F-to-enter: 4.0

F-to-remove: 4.0

Step 0:

12 variables in the model. 48 d.f. for error.

R-squared = 35.86% Adjusted R-squared = 19.83% MSE = 3.66461

Step 1:

Removing variable Com with F-to-remove =0.125865

11 variables in the model. 49 d.f. for error.

R-squared = 35.70% Adjusted R-squared = 21.26% MSE = 3.59924

Step 2:

Removing variable Skill with F-to-remove =0.128602

10 variables in the model. 50 d.f. for error.

R-squared = 35.53% Adjusted R-squared = 22.63% MSE = 3.53651

Step 3:

Removing variable Gender with F-to-remove =0.163742

9 variables in the model. 51 d.f. for error.

R-squared = 35.32% Adjusted R-squared = 23.90% MSE = 3.47852

Step 4:

Removing variable CVPA with F-to-remove =0.516045

8 variables in the model. 52 d.f. for error.

R-squared = 34.66% Adjusted R-squared = 24.61% MSE = 3.44615

Step 5:

Removing variable Hedonic with F-to-remove =0.602751

7 variables in the model. 53 d.f. for error.

R-squared = 33.90% Adjusted R-squared = 25.17% MSE = 3.42032

Step 6:

Removing variable Utility with F-to-remove =0.799728

6 variables in the model. 54 d.f. for error.

R-squared = 32.91% Adjusted R-squared = 25.45% MSE = 3.40763

Step 7:

Removing variable Enter with F-to-remove =1.75531

5 variables in the model. 55 d.f. for error.

R-squared = 30.73% Adjusted R-squared = 24.43% MSE = 3.45443

Step 8:

Removing variable Involv with F-to-remove =2.16534

4 variables in the model. 56 d.f. for error.

R-squared = 28.00% Adjusted R-squared = 22.86% MSE = 3.52631

Final model selected.

The output shows the results of fitting a multiple linear regression model to describe the relationship between attrac24 and 12 independent variables. The equation of the fitted model is

attrac24 = 7.19257 + 1.50162*S_N_S - 0.461733*Age + 1.08358*Freq - 1.09079*Work

Since the P-value in the ANOVA table is less than 0.05, there is a statistically significant relationship between the variables at the 95% confidence level.

Multiple Regressions - attrac29

Dependent variable: attrac29

Independent variables: S_N_S, Gender, Age, Freq, Enter, Work, Com, Skill, Utility, Hedonic, Involv, CVPA

		Standard	*T*	
Parameter	*Estimate*	*Error*	*Statistic*	*P-Value*
CONSTANT	4.62002	1.23496	3.74102	0.0004
Utility	0.937318	0.386822	2.42312	0.0185
CVPA	-1.04504	0.334323	-3.12585	0.0028

Analysis of Variance

Source	Sum of Squares	Df	Mean Square	F-Ratio	P-Value
Model	85.1922	2	42.5961	9.67	0.0002
Residual	259.985	59	4.40653		
Total (Corr.)	345.177	61			

R-squared = 24.6807 percent

R-squared (adjusted for d.f.) = 22.1275 percent

Standard Error of Est. = 2.09917

Mean absolute error = 1.63374

Durbin-Watson statistic = 1.5933 (P=0.0502)

Lag 1 residual autocorrelation = 0.201103

Stepwise Regression

Method: backward selection

F-to-enter: 4.0

F-to-remove: 4.0

Step 0:

12 variables in the model. 49 d.f. for error.

R-squared = 48.01% Adjusted R-squared = 35.28% MSE = 3.66249

Step 1:

Removing variable S_N_S with F-to-remove =0.52705

11 variables in the model. 50 d.f. for error.

R-squared = 47.45% Adjusted R-squared = 35.89% MSE = 3.62784

Step 2:

Removing variable Enter with F-to-remove =1.32552

10 variables in the model. 51 d.f. for error.

R-squared = 46.06% Adjusted R-squared = 35.48% MSE = 3.651

Step 3:

Removing variable Skill with F-to-remove =1.03642

9 variables in the model. 52 d.f. for error.

R-squared = 44.96% Adjusted R-squared = 35.43% MSE = 3.65356

Step 4:

Removing variable Gender with F-to-remove =2.07456

8 variables in the model. 53 d.f. for error.

R-squared = 42.76% Adjusted R-squared = 34.13% MSE = 3.72763

Step 5:

Removing variable Involv with F-to-remove =2.0495

7 variables in the model. 54 d.f. for error.

R-squared = 40.55% Adjusted R-squared = 32.84% MSE = 3.80008

Step 6:

Removing variable Com with F-to-remove =1.7733

6 variables in the model. 55 d.f. for error.

R-squared = 38.60% Adjusted R-squared = 31.90% MSE = 3.85351

Step 7:

Removing variable Freq with F-to-remove =3.65102

5 variables in the model. 56 d.f. for error.

R-squared = 34.52% Adjusted R-squared = 28.68% MSE = 4.03593

Step 8:

Removing variable Hedonic with F-to-remove =3.20025

4 variables in the model. 57 d.f. for error.

R-squared = 30.78% Adjusted R-squared = 25.92% MSE = 4.19172

Step 9:

Removing variable Age with F-to-remove =2.18273

3 variables in the model. 58 d.f. for error.

R-squared = 28.13% Adjusted R-squared = 24.41% MSE = 4.2772

Step 10:

Removing variable Work with F-to-remove =2.78401

2 variables in the model. 59 d.f. for error.

R-squared = 24.68% Adjusted R-squared = 22.13% MSE = 4.40653

Final model selected.

The StatAdvisor

The output shows the results of fitting a multiple linear regression model to describe the relationship between attrac29 and 12 independent variables. The equation of the fitted model is

attrac29 = 4.62002 + 0.937318*Utility - 1.04504*CVPA

Since the P-value in the ANOVA table is less than 0.05, there is a statistically significant relationship between the variables at the 95% confidence level.

<u>Multiple Regression—attrac32</u>

Dependent variable: attrac32

Independent variables: S_N_S, Gender, Age, Freq, Enter, Work, Com, Skill, Utility, Hedonic, Involv, CVPA

		Standard	*T*	
Parameter	*Estimate*	*Error*	*Statistic*	*P-Value*
CONSTANT	3.05696	0.553023	5.52773	0.0000
Age	1.10149	0.246921	4.46089	0.0000

Analysis of Variance

Source	Sum of Squares	Df	Mean Square	F-Ratio	P-Value
Model	127.453	1	127.453	19.90	0.0000
Residual	384.289	60	6.40482		
Total (Corr.)	511.742	61			

R-squared = 24.9057 percent

R-squared (adjusted for d.f.) = 23.6541 percent

Standard Error of Est. = 2.53077

Mean absolute error = 1.9923

Durbin-Watson statistic = 1.4746 (P=0.0144)

Lag 1 residual autocorrelation = 0.226142

<u>Stepwise Regression</u>

Method: backward selection

F-to-enter: 4.0

F-to-remove: 4.0

<u>Step 0:</u>

12 variables in the model. 49 d.f. for error.

R-squared = 36.70% Adjusted R-squared = 21.20% MSE = 6.611

Step 1:

Removing variable Freq with F-to-remove =0.00787517

11 variables in the model. 50 d.f. for error.

R-squared = 36.69% Adjusted R-squared = 22.76% MSE = 6.47982

Step 2:

Removing variable Skill with F-to-remove =0.0191592

10 variables in the model. 51 d.f. for error.

R-squared = 36.66% Adjusted R-squared = 24.25% MSE = 6.3552

Step 3:

Removing variable S_N_S with F-to-remove =0.0244334

9 variables in the model. 52 d.f. for error.

R-squared = 36.63% Adjusted R-squared = 25.67% MSE = 6.23597

Step 4:

Removing variable Utility with F-to-remove =0.0471402

8 variables in the model. 53 d.f. for error.

R-squared = 36.58% Adjusted R-squared = 27.00% MSE = 6.12386

Step 5:

Removing variable Com with F-to-remove =0.158873

7 variables in the model. 54 d.f. for error.

R-squared = 36.39% Adjusted R-squared = 28.14% MSE = 6.02847

Step 6:

Removing variable Work with F-to-remove =0.470784

6 variables in the model. 55 d.f. for error.

R-squared = 35.83% Adjusted R-squared = 28.83% MSE = 5.97046

Step 7:

Removing variable CVPA with F-to-remove =1.03975

5 variables in the model. 56 d.f. for error.

R-squared = 34.62% Adjusted R-squared = 28.78% MSE = 5.9747

<u>Step 8:</u>

Removing variable Gender with F-to-remove =1.16621

4 variables in the model. 57 d.f. for error.

R-squared = 33.26% Adjusted R-squared = 28.57% MSE = 5.99212

<u>Step 9:</u>

Removing variable Enter with F-to-remove =1.36772

3 variables in the model. 58 d.f. for error.

R-squared = 31.66% Adjusted R-squared = 28.12% MSE = 6.03011

<u>Step 10</u>

Removing variable Involv with F-to-remove =2.90629

2 variables in the model. 59 d.f. for error.

R-squared = 28.23% Adjusted R-squared = 25.80% MSE = 6.22495

<u>Step 11:</u>

Removing variable Hedonic with F-to-remove =2.73371

1 variables in the model. 60 d.f. for error.

R-squared = 24.91% Adjusted R-squared = 23.65% MSE = 6.40482

Final model selected.

The output shows the results of fitting a multiple linear regression model to describe the relationship between attrac32 and 12 independent variables. The equation of the fitted model is

attrac32 = 3.05696 + 1.10149*Age

Since the P-value in the ANOVA table is less than 0.05, there is a statistically significant relationship between the variables at the 95% confidence level.

<u>Multiple Regressions—attrac17</u>

Dependent variable: attrac17

Independent variables: Age, Com, CVPA, Enter, Freq, Gender, Hedonic, Involvement, S_N_S, Skill, Utility, Work

		Standard	T	
Parameter	*Estimate*	*Error*	*Statistic*	*P-Value*
CONSTANT	8.67619	0.452917	19.1563	0.0000
Age	-0.477199	0.202225	-2.35975	0.0216

Analysis of Variance

Source	Sum of Squares	Df	Mean Square	F-Ratio	P-Value
Model	23.9215	1	23.9215	5.57	0.0216
Residual	257.756	60	4.29593		
Total (Corr.)	281.677	61			

R-squared = 8.49253 percent

R-squared (adjusted for d.f.) = 6.96741 percent

Standard Error of Est. = 2.07266

Mean absolute error = 1.59601

Durbin-Watson statistic = 2.19893 (P=0.7656)

Lag 1 residual autocorrelation = -0.107

Stepwise Regression

Method: forward selection

F-to-enter: 4.0

F-to-remove: 4.0

Step 0:

0 variables in the model. 61 d.f. for error.

R-squared = 0.00% Adjusted R-squared = 0.00% MSE = 4.61766

Step 1:

Adding variable Age with F-to-enter =5.56842

1 variables in the model. 60 d.f. for error.

R-squared = 8.49% Adjusted R-squared = 6.97% MSE = 4.29593

Final model selected.

The output shows the results of fitting a multiple linear regression model to describe the relationship between attrac17 and 12 independent variables. The equation of the fitted model is

attrac17 = 8.67619 - 0.477199*Age

Since the P-value in the ANOVA table is less than 0.05, there is a statistically significant relationship between the variables at the 95% confidence level.

Regression Analysis of Some Websites that Did Not Show Relationship in the Philips' (2007) study

Multiple Regressions—attrac23

Dependent variable: attrac23

Independent variables: Age, Com, CVPA, Enter, Freq, Gender, Hedonic, Involvement, S_N_S, Skill, Utility, Work

Parameter	*Estimate*	*Standard Error*	*T Statistic*	*P-Value*
CONSTANT	6.89475	2.68879	2.56425	0.0135
Age	-0.499891	0.346875	-1.44113	0.1559
Com	-0.317071	0.67846	-0.46734	0.6423
CVPA	0.297224	0.38889	0.764289	0.4484
Enter	-0.794498	0.767865	-1.03468	0.3059
Freq	1.38334	0.592327	2.33543	0.0237
Gender	0.585231	0.688897	0.849519	0.3997
Hedonic	-0.19898	0.451714	-0.440501	0.6615
Involv	-0.351696	0.282006	-1.24712	0.2183
S_N_S	0.509861	0.782229	0.651805	0.5176
Skill	0.358653	0.371281	0.96599	0.3388
Utility	-0.187782	0.541882	-0.346536	0.7304
Work	-0.0627882	0.792095	-0.0792685	0.9371

Analysis of Variance

Source	Sum of Squares	Df	Mean Square	F-Ratio	P-Value
Model	100.845	12	8.40377	1.57	0.1306
Residual	261.494	49	5.3366		
Total (Corr.)	362.339	61			

R-squared = 27.8318 percent

R-squared (adjusted for d.f.) = 10.1579 percent

Standard Error of Est. = 2.31011

Mean absolute error = 1.63079

Durbin-Watson statistic = 2.09006 (P=0.5251)

Lag 1 residual autocorrelation = -0.0621854

The output shows the results of fitting a multiple linear regression model to describe the relationship between attrac23 and 12 independent variables. The equation of the fitted model is

attrac23 = 6.89475 - 0.499891*Age - 0.317071*Com + 0.297224*CVPA - 0.794498*Enter + 1.38334*Freq + 0.585231*Gender - 0.19898*Hedonic - 0.351696*Involv + 0.509861*S_N_S + 0.358653*Skill - 0.187782*Utility - 0.0627882*Work

Since the P-value in the ANOVA table is greater or equal to 0.05, there is not a statistically significant relationship between the variables at the 95% or higher confidence level.

Multiple Regression—attrac30

Dependent variable: attrac30

Independent variables: Age, Com, CVPA, Enter, Freq, Gender, Hedonic, Involvement, S_N_S, Skill, Utility, Work

		Standard	*T*	
Parameter	*Estimate*	*Error*	*Statistic*	*P-Value*
CONSTANT	2.71557	2.40046	1.13127	0.2634
S_N_S	0.266934	0.698346	0.382237	0.7039
Gender	-0.323112	0.615022	-0.525366	0.6017
Age	0.452492	0.309677	1.46117	0.1504
Freq	0.755269	0.528808	1.42825	0.1596
Enter	-0.696408	0.685522	-1.01588	0.3147
Work	-0.284568	0.707154	-0.402413	0.6891
Com	-0.327032	0.605704	-0.539921	0.5917
Skill	0.595528	0.331466	1.79665	0.0786
Utility	-0.433307	0.483773	-0.895684	0.3748
Hedonic	0.0735776	0.403274	0.182451	0.8560
Involv	-0.061249	0.251765	-0.243278	0.8088
CVPA	0.433695	0.347187	1.24917	0.2175

Analysis of Variance

Source	Sum of Squares	Df	Mean Square	F-Ratio	P-Value
Model	51.3246	12	4.27705	1.01	0.4583
Residual	208.417	49	4.25341		
Total (Corr.)	259.742	61			

R-squared = 19.7599 percent

R-squared (adjusted for d.f.) = 0.109203 percent

Standard Error of Est. = 2.06238

Mean absolute error = 1.50014

Durbin-Watson statistic = 2.3632 (P=0.8783)

Lag 1 residual autocorrelation = -0.188788

The output shows the results of fitting a multiple linear regression model to describe the relationship between attrac30 and 12 independent variables. The equation of the fitted model is

attrac30 = 2.71557 + 0.266934*S_N_S - 0.323112*Gender + 0.452492*Age + 0.755269*Freq - 0.696408*Enter - 0.284568*Work - 0.327032*Com + 0.595528*Skill - 0.433307*Utility + 0.0735776*Hedonic - 0.061249*Involv + 0.433695*CVPA

Since the P-value in the ANOVA table is greater or equal to 0.05, there is not a statistically significant relationship between the variables at the 95% or higher confidence level.

Multiple Regressions - attrac33

Dependent variable: attrac33

Independent variables: Age, Com, CVPA, Enter, Freq, Gender, Hedonic, Involvement, S_N_S, Skill, Utility, Work

Parameter	*Estimate*	*Standard Error*	*T Statistic*	*P-Value*
CONSTANT	1.30251	2.37941	0.54741	0.5866
S_N_S	1.57902	0.692223	2.28109	0.0269
Gender	0.142464	0.60963	0.23369	0.8162
Age	-0.0479349	0.306962	-0.156159	0.8765
Freq	0.307916	0.524172	0.587433	0.5596
Enter	0.537863	0.679511	0.791543	0.4324
Work	0.568376	0.700953	0.810862	0.4214
Com	0.223775	0.600393	0.372713	0.7110
Skill	0.225123	0.32856	0.685182	0.4965
Utility	-0.66225	0.479531	-1.38104	0.1735
Hedonic	-0.0649809	0.399738	-0.162559	0.8715
Involv	-0.167689	0.249558	-0.671947	0.5048
CVPA	-0.232261	0.344142	-0.674896	0.5029

Analysis of Variance

Source	Sum of Squares	Df	Mean Square	F-Ratio	P-Value
Model	46.9633	12	3.91361	0.94	0.5196
Residual	204.779	49	4.17916		
Total (Corr.)	251.742	61			

R-squared = 18.6553 percent

R-squared (adjusted for d.f.) = 0.0 percent

Standard Error of Est. = 2.0443

Mean absolute error = 1.4915

Durbin-Watson statistic = 1.85426 (P=0.1891)

Lag 1 residual autocorrelation = 0.0548513

The output shows the results of fitting a multiple linear regression model to describe the relationship between attrac33 and 12 independent variables. The equation of the fitted model is

attrac33 = 1.30251 + 1.57902*S_N_S + 0.142464*Gender - 0.0479349*Age + 0.307916*Freq + 0.537863*Enter + 0.568376*Work + 0.223775*Com + 0.225123*Skill - 0.66225*Utility - 0.0649809*Hedonic - 0.167689*Involv - 0.232261*CVPA

Since the P-value in the ANOVA table is greater or equal to 0.05, there is not a statistically significant relationship between the variables at the 95% or higher confidence level.

Multiple Regressions—attrac5

Dependent variable: attrac5

Independent variables: Age, Com, CVPA, Enter, Freq, Gender, Hedonic, Involvement, S_N_S, Skill, Utility, Work

Parameter	*Estimate*	*Standard Error*	*T Statistic*	*P-Value*
CONSTANT	7.37624	3.10854	2.37289	0.0217
S_N_S	-0.121686	0.913041	-0.133276	0.8945
Gender	0.0615691	0.796162	0.0773324	0.9387
Age	-0.0153408	0.398095	-0.0385356	0.9694
Freq	-0.929058	0.678388	-1.36951	0.1772
Enter	0.088021	0.88282	0.0997044	0.9210
Work	-0.828987	0.917555	-0.903475	0.3708
Com	-1.20219	0.783804	-1.53379	0.1316
Skill	0.648419	0.425213	1.52493	0.1338
Utility	0.605495	0.620594	0.975671	0.3341
Hedonic	0.0851014	0.518157	0.164239	0.8702
Involv	-0.0408721	0.324391	-0.125996	0.9003
CVPA	-0.363231	0.447503	-0.811682	0.4210

Analysis of Variance

Source	Sum of Squares	Df	Mean Square	F-Ratio	P-Value
Model	85.0407	12	7.08672	1.01	0.4527
Residual	335.976	48	6.99949		
Total (Corr.)	421.016	60			

R-squared = 20.1989 percent

R-squared (adjusted for d.f.) = 0.248621 percent

Standard Error of Est. = 2.64566

Mean absolute error = 1.88732

Durbin-Watson statistic = 1.95445 (P=0.3167)

Lag 1 residual autocorrelation = 0.00221804

The output shows the results of fitting a multiple linear regression model to describe the relationship between attrac5 and 12 independent variables. The equation of the fitted model is

attrac5 = 7.37624 - 0.121686*S_N_S + 0.0615691*Gender - 0.0153408*Age - 0.929058*Freq + 0.088021*Enter - 0.828987*Work - 1.20219*Com + 0.648419*Skill + 0.605495*Utility + 0.0851014*Hedonic - 0.0408721*Involv - 0.363231*CVPA

Since the P-value in the ANOVA table is greater or equal to 0.05, there is not a statistically significant relationship between the variables at the 95% or higher confidence level.