LUT
Lappeenranta
University of Technology

Finance

Antti Sahikoski

Master's Thesis

Visual Customer Segmentation and Data Exploration with the Self-Organizing Map

Examiners: Professor Mikael Collan

M.Sc. Kaisa Kuokka

ABSTRACT

Author:	Antti Sahikoski
Title:	Visual Customer Segmentation and Data Exploration with the Self-Organizing Map
Faculty:	LUT, School of Business and Management
Major:	Finance
Year:	2015
Master's Thesis:	Lappeenranta University of Technology 72 pages, 15 figures
Examiners:	Prof. Mikael Collan M.Sc. Kaisa Kuokka
Keywords:	Self-organizing map, customer relationship management, RFM, segmentation

Advancements in information technology have made it possible for organizations to gather and store vast amounts of data of their customers. Information stored in databases can be highly valuable for organizations. However, analyzing large databases has proven to be difficult in practice.

For companies in the retail industry, customer intelligence can be used to identify profitable customers, their characteristics, and behavior. By clustering customers into homogeneous groups, companies can more effectively manage their customer base and target profitable customer segments.

This thesis will study the use of the self-organizing map (SOM) as a method for analyzing large customer datasets, clustering customers, and discovering information about customer behavior. Aim of the thesis is to find out whether the SOM could be a practical tool for retail companies to analyze their customer data.

TIIVISTELMÄ

Tekijä:	Antti Sahikoski
Tutkielman nimi:	Visuaalinen asiakassegmentointi ja datan tutkinta itseorganisoituvaa karttaa käyttäen
Tiedekunta:	LUT, Kauppatieteellinen tiedekunta
Pääaine:	Rahoitus
Vuosi:	2015
Pro gradu -tutkielma:	Lappeenrannan teknillinen yliopisto 72 sivua, 15 kuvaa
Tarkastajat	Prof. Mikael Collan KTM Kaisa Kuokka
Hakusanat:	Itseorganisoituva kartta, asiakkuudenhallinta, RFM, segmentaatio

Kehitys informaatioteknologiassa on mahdollistanut organisaatioiden kerätä ja varastoida suuret määrät dataa asiakkaistaan. Tietokantoihin tallennettu tieto voi olla hyvin arvokasta organisaatioille. Suurten tietokantojen analysointi on kuitenkin osoittautunut hankalaksi käytännössä.

Vähittäiskauppa-alan yritykset voivat käyttää asiakastietoa kannattavien asiakkaidensa ja näiden ominaisuuksien ja käytöksen tunnistamiseksi. Klusteroimalla asiakkaat homogeenisiin ryhmiin, yritykset voivat tehokkaammin hallinnoida asiakaskuntaansa ja kohdistaa voimavarojaan kannattaviin asiakassegmentteihin.

Tämä tutkielma keskittyy itseorganisoituvan kartan (SOM) käyttöön suurten asiakastietokantojen analysoimisessa, asiakkaiden klusteroinnissa ja asiakaskäyttäytymisen tutkimisessa. Tutkielman tavoitteena on selvittää, voisiko SOM olla käytännöllinen työkalu vähittäismyyntialan yrityksille analysoida asiakasdataansa.

ACKNOWLEDGMENTS

It has been a long journey with both ups and downs. Working on this thesis has been exciting and interesting, and I have learned a lot during the process. But it has also been a demanding journey, and I have often felt that there are not enough hours in the day. Therefore, I am extremely happy now that the work has been done.

I wish to thank Professor Mikael Collan for all the help, comments, and suggestions during this process. Your help has been invaluable, and I greatly appreciate it. I also wish to thank the case company for the opportunity to work on such an interesting project and with a dataset that provided so many opportunities.

Most of all, I would like to thank Elina and our daughters, Hilla and Peppi, for all the support, patience, and understanding during these past months.

In Helsinki, November 29th, 2015

Antti Sahikoski

TABLE OF CONTENTS

1 INTRODUCTION

Advancements in information technology have made it possible for organizations to gather and store vast amounts of data from different sources. As data is more abundant than ever, a new dilemma has arisen: how to gain insights from databases that contain ample amounts of information. A specific industry that is facing this issue is the retail industry. Retail companies are able to gather vast amounts of data of their customers through loyalty programs and bonus cards. Data gathered of customers and their purchasing behavior can be extremely valuable for retail companies, and can be used to drive effective customer relationship management. However, even large amounts of data are of little value unless there is a practical way to analyze the data and benefit from the information that it can provide. This study will focus on a possible solution to this issue by examining whether a data exploration method called the self-organizing map could be a practical way for retail companies to analyze their customer data.

1.1 Background of the Study

Customer relationship management (CRM) has been a growing focus for companies for a number of years now. The goal of CRM is to strengthen the relationship between a company and its customers and consequently maximize customer lifetime value (Ha 2007). Companies are facing ever growing competition from both online and offline competitors, which has led them to focus more on improving the loyalty of their existing customers (Phan, Vogel 2010). Many companies are now thriving to build long-lasting relationships with their customers and are, in essence, transforming from a product-oriented view to a customer-oriented view (Rygielski, Wang et al. 2002).

Successful practice of CRM requires the company to know its customers and their needs. The ongoing advancements in information technology have made it possible for companies to utilize tools and technologies such as data mining, data warehousing, and other CRM techniques, which enable companies to follow the concepts of customer relationship management (Rygielski, Wang et al. 2002).

The CRM framework can be divided into two groups: operational CRM and analytical CRM. Analytical CRM is focused on analyzing the behavior and characteristics of customers, and how such knowledge can be exploited to support the company's customer management strategies (Ngai, Xiu et al. 2009).

Segmentation is a central concept of marketing and one of the key applications of analytical CRM. Creating marketing strategies for individual customers is not an economically sound practice in the majority of cases. Therefore, it is practical to segment customers into homogeneous segments with similar behavior, values, and characteristics. Optimized marketing strategies can then be created to target the segments and individual customers within them.

Variables that are used as the basis of creating customer segments can be organized into four groups: geographic, demographic, psychographic, and behavioral variables (Kotler, Keller 2011, p.214). Many studies highlight the importance of behavioral variables in segmenting customers as they provide the best means of predicting customers' future behavior (Hiziroglu 2013, Bose, Chen 2009). A set of behavioral variables that are very often used in segmentation studies are known as the RFM variables, which refer to recency, frequency, and monetary. The RFM variables are based on customers' past purchasing data. They are especially useful in creating customer segments as they can be used to quantify customer behavior and to identify the company's most valuable customers (Cheng, Sun 2012).

1.2 The Self-Organizing Map

The self-organizing map (SOM) is an unsupervised neural network based method that can be used in data exploration, visualization, and clustering. The SOM works by mapping high-dimensional input data onto a two-dimensional output map (Kohonen 2001). The results of a SOM analysis are presented visually, using e.g. the U-matrix and component planes.

It has been suggested that when used as a data mining method maybe the most important beneficial feature of the SOM is its ability to visualize complex multidimensional data and make it simple and easily understandable (Vesanto 1999). SOM visualizations can be used to, for example, discover correlations, patterns, and clusters within the underlying data. Another beneficial feature of the SOM is its usefulness as a tool for exploratory data analysis (Collan, Eklund et al. 2007). The SOM can be used to quickly explore and analyze large datasets, and find areas that should to be further studied.

The SOM is also a useful method in performing clustering and has been used in academic studies to create customer segments (Bloom 2004, Hong, Kim 2012, Yao, Sarlin et al. 2014). Compared to cluster analysis, the SOM has a number of advantages that make it a better alternative for customer segmentation studies, including robustness and the limited a priori information of the sample data that is needed to perform a SOM analysis (Bloom 2004).

Despite this study focusing on the use of the SOM as a tool to study customer data, it should be mentioned that the SOM has been used in many other fields as well. Kohonen estimated that over 10,000 scientific papers related to the SOM have been published (Kohonen, 2014, p.18). The areas where SOM research has been done are very varied, including engineering, biology, economics, medicine, and speech recognition (Oja, Kaski et al. 2003).

1.3 Data and Empirical Research

We have the privilege of being able to study the use of the SOM with real-life data. We will study a loyalty program member dataset provided to us by our case company, a Finnish retailer with over 150,000 customer loyalty program members. The dataset provided to us by the case company consists of demographic and behavioral data, as well as data related to how customers consume the company's email newsletters.

In the empirical research part of this study, we will create multiple self-organizing maps in order to study the use of the SOM in data clustering and exploration. We will experiment with different SOM visualizations, but try to focus on making each one of them practical and easy to interpret. All SOM visualizations in this study have been made with the MATLAB program using the SOM Toolbox program package.

1.4 Focus and Contribution of the Study

The focus of this study can be seen to consist of two different elements. (1) Firstly, we will study the use of the self-organizing map, and how it can be used to explore and cluster customer data. Our motivation is to see whether the SOM could be a practical tool for companies to use in the analysis of large datasets. (2) Secondly, our focus will be on analytical CRM, as we will analyze customer data with the intention of gaining insights that could be used to create more effective customer relationship management strategies. In the area of analytical CRM, we will focus on customer segmentation, and especially on the use of the RFM variables.

The information gained from this study will provide the case company with more information about their customers, but the results and information gained will also be useful from an academic perspective. (1) Ngai et al. (2009) called for more studies to focus on the use of visual data mining methods in CRM. With this study, we hope to provide more information on how visual data mining methods can be used. (2) The SOM has already been used in customer segmentation studies. However, even though there have been some studies that have applied the self-organizing map to cluster customers based on their RFM values (e.g. Hsieh 2004, Wei, Lin et al. 2012), the number of studies is not abundant. We will provide more evidence on the use of the SOM as a method to perform RFM analysis. (3) Lastly, we will study a non-public dataset from a real-life case company. Data is unique to this study, and we are able to provide more evidence from the real-world.

1.5 Research Questions

In this thesis, we will examine whether the SOM could be a useful tool for companies to explore and visualize multi-dimensional customer data. We will study this broader question by answering three very specific research questions that are listed below. By answering these three questions, we are able to the gain more information about the overall usability of the SOM as a practical tool for exploring customer data.

RQ1: Is the SOM a useful method for conducting customer segmentation using the RFM variables?

RQ2: Is the SOM a useful method for conducting customer segmentation using the RFM and demographic variables?

RQ3: Is the SOM a useful method for exploring email newsletter data?

Research questions one and two resemble each other, but by answering these questions separately, we can highlight similarities and differences that we might find when studying them. Research question three is very specific, but it can also be seen to provide more information in relation to the larger problem: whether the SOM is a practical general-purpose method for analyzing customer data.

1.6 Limitations and Considerations

Broadly speaking, this thesis will study the use of visual data clustering and exploration methods in CRM. We are, however, limiting the study to focus on a much smaller area. (1) This thesis will study the use of only the self-organizing map in performing clustering and data exploration analyses. Other methods are not tested, nor are they presented. (2) We will focus on the retail industry, and only perform tests using the data from a single case company. Results of the analyses are therefore specific to this case company.

In particular, to the data that is used to carry out the empirical tests, some limitations must be mentioned. Included in the dataset, are only the case company's loyalty program members who had made at least two transactions during the study period and had spent a minimum of 20€ on their purchases. All other customers are excluded from the analysis.

In this thesis, we will talk about customer segments, and how they can be created using the SOM. It should, however, be noted that the SOM is most of all an exploratory model that can be used to quickly identify the kinds of clusters that exist in the data. The results of our analyses, therefore, should be considered only as a possible starting point for creating actual targetable customer segments.

1.7 Structure of the Study

The remainder of this study is organized as follows. Chapter 2 will present the theoretical background of the study. The concepts of customer relationship management and customer segmentation will be discussed with a special focus on the RFM variables.

Chapter 3 will present the methodology used to carry out the empirical tests. An introduction to the self-organizing map will be given and a brief look at recent research articles will be presented. We will also present examples of SOM visualizations and how they can be interpreted.

Chapter 4 will present the research data and variables. Data preparation and preprocessing will be discussed. Lastly, in order for the reader to get a better understanding of the research data, characteristics of data variables will be presented.

Chapter 5 will present the results of the empirical research. Firstly, the SOM of RFM variables will be presented. Secondly, the SOM of RFM and demographic variables will be presented. And lastly, three self-organizing maps that were created using the newsletter data will be presented.

Chapter 6 will conclude the study and answer the research questions. Suggestions for further research will also be presented.

2 THEORETICAL BACKGROUND

In this chapter, we will introduce the theoretical background of the thesis. We will give an introduction to customer relationship management and discuss customer segmentation – a central concept of analytical CRM – more closely. We will especially focus on the RFM variables, which are often used to cluster customers based on their purchasing behavior.

2.1 Introduction to Customer Relationship Management

A universal definition for customer relationship management (CRM) does not exist. Kotler and Keller (2011, p.135) defined CRM as the process of managing customer information - and all occasions when a customer is in contact with the product or brand of a company - with the intention of maximizing customer loyalty. In their literature review, Ngai et al. (2009) presented a number of definitions for CRM, but summarized that the common aspect of these definitions was that they emphasized the importance of seeing CRM as a comprehensive process of acquiring and retaining customers, and maximizing customer value to the company through the use of business intelligence. Ling and Yen (2001) summarized that the most important aspects of a CRM definition are customer value, technology empowerment, and a holistic approach. For companies, the goal of CRM is to forge closer relationships with customers and to maximize customer lifetime value (Ha 2007).

The importance of CRM has been a growing focus for companies for multiple years. Some of the reasons that are driving companies to focus more on CRM include the high cost of acquiring new customers, which has led companies to focus more on their existing customers (Rygielski, Wang et al. 2002); and the growing competition that companies face from both online and traditional businesses, which has made it more important for companies to focus on making sure their current customers are satisfied and to maintain their customer loyalty (Phan, Vogel 2010). As a result, marketing is now more focused on improving and deepening the relationship with a customer rather than acquiring a larger customer base. Companies are transforming

from a product-oriented view to a customer-oriented view, moving from mass marketing to one-to-one marketing, and are thriving to build long-lasting relationships with their customers. (Rygielski, Wang et al. 2002)

The growing popularity of CRM can also be explained by recent advancements that have made it easier to manage information needed to practice CRM, and by the sheer amount of data companies are able to gather, quite easily, of their customers and their purchasing behavior these days. The ongoing advancements in information technology have made it possible for companies to make use of immense potential of customer relationships that was not possible before (Phan, Vogel 2010). Tools and technologies such as data mining, data warehousing, and other CRM techniques are making it easier for companies to follow the concepts of customer relationship marketing (Rygielski, Wang et al. 2002). For example, developments in data mining and data warehousing have made it easier for companies to practice cross-selling and up-selling to their existing customers (Wang, Chiang et al. 2009). Successful CRM requires the company to know its customers, their wants, and needs. Gathering vast amounts of customer data is a relatively simple process these days but was simply impossible, on such a large scale, only a number of years ago.

Successful implementation of CRM offers companies a number of benefits to help them stay competitive in the world of ever tightening competition. With effective CRM, companies have better knowledge of their valued customers and can customize services, market offerings, and advertising to these customers (Kotler, Keller 2011, p.214). All the aforementioned elements are used to provide better service for the company’s customers. Companies can also make use of the detailed customer intelligence they have gathered to predict future behavior of customers, which can support the company in making proactive decisions that are driven by knowledge and data. Furthermore, customer intelligence can be used to identify profitable customers, who the company should focus more on, as well as

unprofitable customers, who are not worth targeting anymore. (Rygielski, Wang et al. 2002)

The CRM framework can be divided into two groups that are focused on different aspects of CRM. These two groups can be labeled as operational CRM and analytical CRM. Operational CRM is focused on the automation of business processes. Analytical CRM, on the other hand, is focused on the analysis of the company's customers, their behavior, characteristics, and how knowledge gained from various analyses can be exploited to support the company's customer management strategies. (Ngai, Xiu et al. 2009)

This thesis is focused on analytical CRM, and especially focused on customer segmentation, which is an important application of analytical CRM. Customer segmentation is discussed more thoroughly in the next part of this thesis.

2.2 Customer Segmentation

Segmentation is a central concept of marketing with abundant amount of research done on the topic. In general, segmentation aims to maximize the within-segment homogeneity and the between segment heterogeneity (Jonker, Piersma et al. 2004). Customer segmentation is used to divide a market into smaller segments, based on the needs and wants of the customers (Kotler, Keller 2011, p.214). Customer segmentation is needed for a company to be able to economically manage customers with different characteristics, behavior, and customer value. After segmenting customers, the customer segments can be profiled, the future behavior of customers in a segment can be predicted, and an organization's marketing strategy can be optimized to target specific customer segments in order to better meet their needs and desires.

2.2.1 Segmentability and Effective Segmentation

Segmenting a market can be done in many ways, but one should keep in mind that the outcome of customer segmentation should always be actually useful in practice. The term segmentability refers to the question when, and under which conditions, it is possible to segment a market (Hiziroglu 2013). Kotler and Keller (2011, p.231) listed five criteria that are required to be fulfilled in order to perform effective segmentation:

1. **Measurable** Different variables related to the segments - such as the size, purchasing power, and characteristics of the segments - must always be measurable.
2. **Substantial** The segments that are created need to be large and profitable enough to serve. Aim of segmentation is not to find niche groups, but to find the largest possible homogeneous groups of customers that are worth targeting with custom-made marketing programs.
3. **Accessible** The segments need to be such that they can be reached and served effectively.
4. **Differentiable** The segments need to be distinguishable from each other. In practice this means that the segments respond differently to different elements and programs of the marketing-mix.
5. **Actionable** The segments need to be such that marketing programs can be effectively created to attract and serve them.

There are naturally many other attempts to classify the characteristics that comprise an effective segmentation study. For a more thorough discussion on the topic, readers can refer to the literature review by Hiziroglu (2013). In his review, Hiziroglu summarized that a comprehensive analysis of the key criteria does not currently exist in scientific literature. However, the five criteria, which are listed above, are the five common criteria that are found in various classification efforts. Hiziroglu went on to add that from a clustering perspective homogeneity should be added to the list as the sixth criteria.

Various researchers have pointed out flows in previous research where the criteria for effective segmentation were not fulfilled or the outcome of the research was not ideal. For example, Chan (2008) criticized past studies for not taking into consideration the link between customer segmentation and the implementation of efficient campaign plans to target profitable customers. Jonker et al. (2004) noted that from a direct marketing perspective, the goal of an effective segmentation study should be to identify the segments that will help the marketer to maximize profits, instead of a segmentation that is focused on maximizing the within segment homogeneity.

2.2.2 Segmentation Variables

Customers can, in theory, be segmented into groups based on any imaginable variable. When performing customer segmentation, one should, however, keep in mind the purpose of the segmentation analysis and not get lost in the limitless possibilities that the vast amount of data in modern databases can offer. Therefore, it is important that the variables that are used in segmentation should be decided based on the goal of segmentation.

Kotler and Keller (2011, p.214) organized segmentation variables into four groups:

1. **Geographic segmentation** The market can be divided according to geographical units e.g. country, city, or postal code.
2. **Demographic segmentation** Segmenting customers based on e.g. age, gender, income, or nationality.
3. **Psychographic segmentation** Customers can be segmented into groups based on their values, lifestyle, or personality traits.
4. **Behavioral segmentation** Customers can be segmented based on their behavior e.g. knowledge of a product, their usage of a product, or their attitudes and loyalty toward a product or brand.

Demographics and socio-economic variables are easily obtainable. However, citing multiple articles, Hiziroglu (2013) suggested that these variables should rarely be the basis of segmentation as they are not effective in predicting future consumer behavior, nor are they effective in providing guidance for communication strategies or product development.

More useful segmentation results can be gained, if customers are segmented based on their purchasing behavior. For example, Bloom (2004) suggested that in strategic marketing, customer segments should be created based on behavioral variables. He went on to add that demographic, value, or lifestyle characteristics should, instead, be used to determine sub-segments that are used to drive tactical marketing initiatives. Bose and Chen (2009) came to the same conclusion and concluded their literature review by stating that future behavior of customers could be better predicted if, instead of only using demographic data, marketers used both demographic and purchasing information to predict customers' future behavior.

Cheng and Sun (2012) gave a practical example that highlighted the reason why customers should be segmented based on their consumption behavior, rather than demographics or attitudes: In the telecommunications and internet services business, high value customers with high monthly expenditure are not necessarily the ones with high income; the most valuable customers could even be students with very little disposable income, but high expenditure on telecommunications services. The importance of analyzing customer behavior and purchasing patterns can also be seen in academic studies; only very few recent studies on direct marketing have relied solely on geographic, demographic, or lifestyle data. Instead, these three types of data have often been combined with data of customer behavior (Bose, Chen 2009).

2.2.3 The RFM Model and RFM Variables

The RFM model is one of the most often used segmentation models by marketers and has also been extensively used in academic studies (Olson, Chae 2012). The

RFM variables (recency, frequency, monetary) are based on customers' past transaction data, and can be used to quantify customer purchase behavior and to identify the most valuable and profitable customers of a company (Cheng, Sun 2012).

Depending on the study, the RFM variables may be defined differently, but the following definition used by e.g. Liu et al. (2009) and Shim et al. (2012) is quite common. R stands for recency, and is the amount of time since the customer's last transaction; a low value for R is preferred, as a low value indicates a higher probability of the customer making repeat transactions. F stands for frequency, and is the sum of transactions during a period of time; a higher frequency of transactions indicates stronger loyalty toward the company. M stands for monetary, and is the sum of transaction value during a period of time; a high monetary value is preferred, and the company should focus on retaining its high spending customers.

There are a number of ways the RFM analysis can be performed. An often used method is the basic RFM model, which does not require any advanced methods to be used. The basic RFM model works by sorting customers in an ascending order based on each of the RFM variables and then dividing the list of customers into quintiles. Each quintile will be given a score of 1-5, and as a result, each customer will have a score of 1-5 for each of the RFM variables. The most valuable customers of a company will belong to the group with an RFM score of 5-5-5. Once the process is complete, all customers will have been divided into 125 cells. (Chan 2008, Chan, Cheng et al. 2011)

Although it is possible to perform an RFM analysis by using the basic RFM model, better results can be obtained if the RFM variables are used as input data for more advanced models. Olson and Chae (2012) compared the performance of the basic RFM model to more sophisticated data mining techniques, namely logistic regression, decision tree, and neural networks, which used the RFM variables as input data. In accordance with previous research, Olson and Chae found the more

sophisticated data mining techniques to outperform the basic RFM model in both prediction accuracy and customer gains. The basic RFM model has also been criticized by e.g. Hiziroglu and Sengul (2012), who found the basic structural model to outperform the RFM model when the segmentation structures produced by each model were compared. Despite the criticism towards the basic RFM model, Olson and Chae (2012) highlighted that the RFM variables are very important in understanding customer purchasing behavior, and that these variables alone can be used to create effective customer response models. In addition to decision tree, neural networks, and logistic regression, the RFM variables have been used together with other advanced methods as well e.g. K-means clustering algorithm (Li, Dai et al. 2011) and the self-organizing map (Hsieh 2004, Wei, Lee et al. 2013).

The popularity of the RFM variables can be explained by the simplicity of the RFM model and the availability of data; RFM variables are often easily obtainable from customers' transaction records. Another good quality of the RFM model is that it can be easily modified or extended to better work with the data that is available, and to better serve the sought outcomes of a study. For example, in recent studies, recency has been defined as the latest purchase amount (Chan 2008) or been replaced with average usage time (Cheng, Sun 2012), and monetary has been defined as the average expenditure of a customer (Chen, Chiu et al. 2005). The RFM variables can also be weighted if it is seen necessary. For example, it has been suggested that the weight allocation of the variables should be R>F>M (Chan, Cheng et al. 2011, Chen, Chiu et al. 2005). A number of studies have also developed extended RFM models that add an additional variable to the model (Li, Dai et al. 2011, Khajvand, Zolfaghar et al. 2011).

2.2.4 Recent Segmentation Studies Utilizing the RFM Variables

The RFM variables have often been used to study the behavior of retail customers in order to create more effective CRM strategies. Chan et al. (2011) developed pricing and promotion strategies for an online store that were based on the concepts of CRM. Customers were first segmented based on their value to the company, and

were then given the opportunity to bargain over the list prices of products; the more valuable a customer was for the online store, the greater the price concession was that he/she could be given when purchasing goods. Liu et al. (2009) developed a new recommendation method that combined two existing methods: collaborative filtering and sequential rule-based filtering. In this study, the RFM variables were used to cluster customers, and sequential rules were then extracted from purchase sequences to make recommendations for each of the segments. The study assumed that customers with similar RFM values were likely to make similar purchases and have comparable purchasing behavior. Chen et al. (2005) highlighted the importance of detecting and predicting changes in customer behavior in order to develop long-term relationships with retail customers. They used demographic and RFM variables to build a customer profile, and then studied the relationship between customer profile and the products that customers purchased. Finally, they developed a method for identifying patterns in customer behavior over time, which could be used to develop more impactful marketing strategies.

Although especially popular in the retail field, the RFM variables have been used to study customer behavior in many other industries as well. Cheng and Sun (2012) developed a version of the RFM model that could be used to study the behavior of 3G subscribers and their value to the service provider. After segmenting customers, the intra-cluster and inter-cluster behavior of subscribers were studied in order to find out more information of the ways consumers in different value segments consume 3G services. Hsieh (2004) studied the behavior of a bank's credit card owners. Credit card owners were segmented based on their repayment behavior and RFM variables; association rules were then employed to create customer profiles in each of the segments. These segments could then be targeted with more effective marketing strategies. Li et al. (2011) used an extended RFM model to classify the customers of a textile manufacturing business. In their extended model, relationship length was added as an additional variable to the RFM model to create the customer segments. Motivation for the study was to be able to create effective discriminative marketing strategies. Khajvand et al. (2011) also applied an extended

RFM model to segment the customers of a manufacturing business. However, the extended model, which included a variable called count item, was not found to have an effect on the segmentation results when compared with the results of using a standard RFM model.

Although, it has been suggested that advanced methods should be used to perform RFM analysis (Olson, Chae 2012), there have only been a small number of studies that have used the SOM to cluster customers based on their RFM values. Wei et al. (2012) clustered customers of a dental clinic based on an extended RFM model. Hsieh (2004) used the SOM to segment bank customers based on their RFM values and repayment behavior. Wei el al. (2013) combined the use of SOM and K-means clustering to segment customers of a hair salon based on customers' RFM values. RFM values and SOM were also used in two studies that focused on the migration of customers from segment to segment (Ha 2007, Ha 2002).

3 THE SELF-ORGANIZING MAP

This part of the thesis gives an introduction to the self-organizing map, presents recent studies where the SOM has been used in CRM context, and gives an introduction to SOM visualizations and how they can be interpreted.

3.1 Introduction to the Self-Organizing Map

The self-organizing map (SOM) is an unsupervised neural network based method that can be used in data exploration, visualization, and clustering. The SOM works by mapping high-dimensional input data onto a two-dimensional output map. The topological relations of the original input data are preserved in the output map (Kiang, Hu et al. 2006). Because of its capability to simultaneously cluster and project data - in contrast to many other clustering algorithms - the cluster structures produced by SOM are understandable without the need for post-processing (Yao, Sarlin et al. 2014).

The SOM is a two-layer artificial neural network consisting of the input and output layers. The output layer is a regular often two-dimensional array of nodes, and the SOM models are associated with the nodes of the array. Each observation of the input data is broadcasted to all of the SOM models and matched with the model that best matches with the input item, illustrated in Figure 1. The best matching model, known as the winner model, and its spatial neighbors in the array are then modified for better matching. As a result of this process, models that resemble each other will be associated with nodes that are located close to each other on the SOM array, and less similar models will be located further away from each other on the array. (Kohonen 2013)

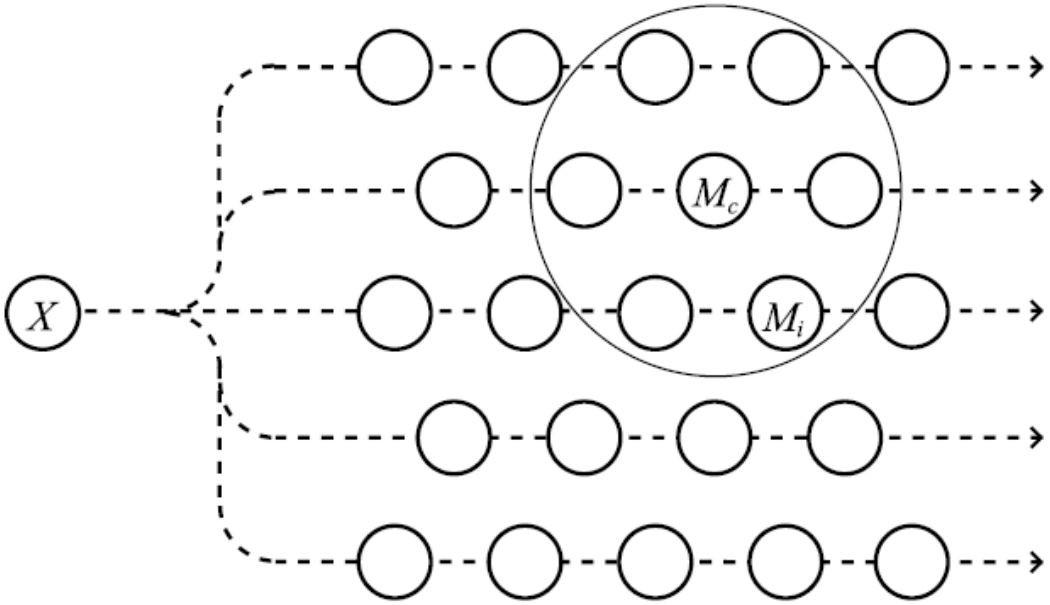

Figure 1 A self-organizing map. The input data X is broadcasted to a set of models (M_i). The best matching unit (M_c) and the units in its neighborhood (large circle) are modified to better match with input unit X. (Kohonen 2013)

The output layer of the SOM is often visualized with a two-dimensional array of nodes. The number of nodes can be predefined so that all the input data is mapped to an array, the size of which is defined by the person performing the SOM analysis. The array size should be chosen based on the size and complexity of the input data. Estimating the definite size of the array beforehand is not possible, and a trial-and-error method should be used to determine the final array size (Kohonen 2013). Each node of the array can be considered its own cluster (Yao, Sarlin et al. 2014). If the data set consists of only a small number of clusters, it is possible to visualize these clusters sufficiently with only a small number of nodes. However, information might be lost if a complex set of data is mapped onto an array that has a resolution too small to display the fine structures in the input data (Kohonen 2013). If, however, the SOM is constructed using a large number of nodes, the processing of the SOM begins to move further away from cluster analysis, and it becomes more difficult to identify the clusters in the data (Yao, Sarlin et al. 2014). A second level of clustering is often applied, in order to cluster the nodes of a larger SOM array (Sarlin, Yao 2013). Applying a second-level of clustering does not have a negative impact on

clustering accuracy when compared with results that are obtained by directly clustering the data (Vesanto, Alhoniemi 2000).

3.2 Key Beneficial Features of the SOM

The SOM has a number of beneficial features that make it a good method for performing segmentation and data mining studies. Bloom (2004) listed three key advantages that the SOM has over cluster analysis when performing market segmentation studies:

1. The self-organizing map is more robust when compared to traditional clustering techniques; missing values in the sample data do not have a significant effect on the performance of the SOM. In comparison, cluster analysis would be more greatly affected by missing elements in the data, and as a result the effective use of cluster analysis would be curtailed.
2. A priori information about the underlying distribution of the sample data is not needed when performing a SOM analysis. Cluster analysis requires the analyst to make various assumptions about the data before performing the analysis.
3. Cluster analysis requires the analyst to specify the number of clusters beforehand. The self-organizing map does not require any such specifications to be made.

In addition to the advantages listed by Bloom, Vesanto (1999) noted that, when used as a data mining method, maybe the most important beneficial feature of the SOM is its ability to visualize complex multidimensional data and make it simple and easily understandable. Also, an important feature of the SOM is its usability as a quick data exploration method. Because the SOM is trained using unsupervised learning and it organizes itself based on identified patterns, it is a great tool for exploring data; the SOM can be used to quickly identify and analyze key features of a dataset (Collan, Eklund et al. 2007).

3.3 Recent Segmentation Research Utilizing the SOM

The self-organizing map has been used in a number of recent studies to create and study customer segments. These studies have applied the SOM to study customer characteristics and behavior from many different industries e.g. online (Hong, Kim 2012) and offline retail (Yao, Sarlin et al. 2014), tourism (Bloom 2004), telecommunication (Kiang, Hu et al. 2006), and banking (Hsieh 2004). Below we present some of the recent studies.

The self-organizing map has been used to study changes in segment structure and examine how individual customers migrate from segment to segment over time. Ha (2007) tracked customer migration from segment to segment and developed a method to predict segment behavior patterns. The self-organizing map was used to segment customers based on their RFM values, and afterwards, decision-tree technique was applied to predict the most likely transition paths of customers. The probability of a customer migrating from one segment to another was also studied in a previous study by Ha (2002), which also used the SOM to segment customers based on RFM variables. In a more recent study, Yao et al. (2014) used the SOM to study how customers migrated from segment to segment during sales campaigns. A study by Sarlin and Yao (2013) did not make the assumption that customer segment structures remain stable over time. In their study, Sarlin and Yao (2013) focused on examining changes in segment structure - the emergence and disappearance of segments.

A number of researchers have focused on studying the performance of the SOM compared to other clustering methods. Kiang et al. (2006) compared the effectiveness of an extended SOM network and a two-step method, which combined factor analysis and k-means clustering, to uncover customer segments. The SOM was found to outperform the two-step procedure when measured by within cluster variance. Shin and Sohn (2004) compared the performance of three clustering methods: SOM, k-means, and fuzzy k-means. Compactness of clusters was used

to measure the performance of the clustering methods. Although all the methods were found to perform well, fuzzy k-means was found to be the most robust method.

The ability of the SOM to visualize multidimensional data and make it easy to interpret has been exploited in academic studies. For example, the study by Yao et al. (2014) relied heavily on the visual interpretation of results; component plane visualization was first used to profile customer segments, and migration planes were used to follow segment migration patterns. Visual inspection of the self-organizing map, by examining e.g. component planes and the U-matrix, has also been used in other recent studies e.g. Sarlin and Yao (2013), Hanafizadeh and Mirzazadeh (2011), and Holmbom et al. (2011).

3.4 Visualization and Interpretation of the SOM

As has already previously been mentioned, one of the best qualities of the SOM is that it is a visual analysis method, and the maps that the SOM produces are easy to understand and interpret. Multiple ways to visualize the SOM exist, and various software packages offer different methods of visualization. We will next introduce three often used methods for the visualization of the SOM: The U-matrix, component planes, and hit histograms. A more thorough introduction to possible visualization methods has been presented by Vesanto (1999).

The U-matrix (unified distance matrix) visualizes, with a single image, the distribution of all variables in a SOM network, and is the most often used distance matrix technique (Vesanto 1999). It is used to analyze the distance between a node and its neighboring nodes. This information can be used to identify clusters and cluster borders within the data.

The U-matrix visualization works in practice by adding additional nodes between the original nodes of the SOM. These nodes are represented by, for example, light colors in cases where the mean distance between neighboring nodes is small, and

by dark colors where the mean distance is large (Kohonen 2014). Clusters on the map are thus represented by uniform areas where the value of mean distance is low, and the underlying data is homogeneous.

Figure 2 presents a U-matrix that was created using the well-known Iris dataset, which is also one of the example datasets used by the SOM Toolbox. Examining the U-matrix in Figure 2, we can see that the data appears to include two clusters. These two clusters are separated by the yellow cells on the map, which represent high mean distance values. Dark blue areas, with low mean distance values, represent the actual clusters.

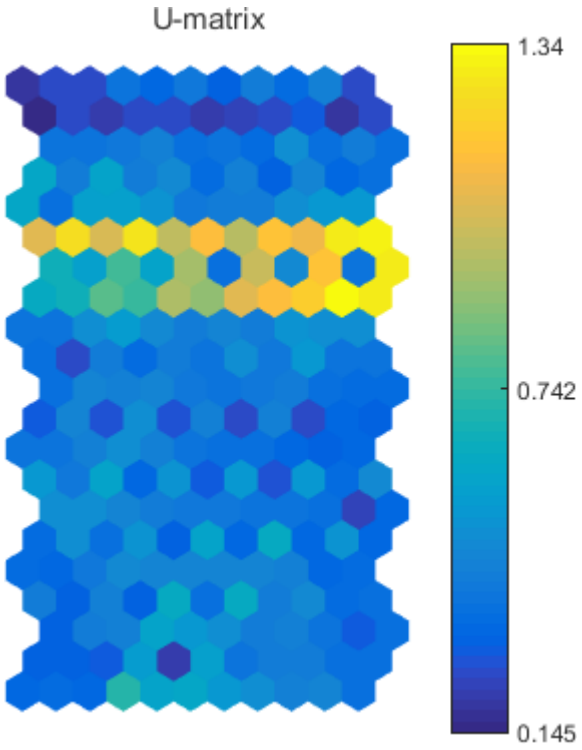

Figure 2 U-matrix created using the Iris dataset.

Component planes are used to provide more information on how values of individual variables in the data are distributed amongst the nodes. By examining component planes, it is possible to study the characteristics of clusters, and to see what kind of values different variables have in each of the clusters. Another use for the component plane visualization is its use in correlation hunting. Similar patterns in the same cells of separate component planes might indicate correlation between the variables. However, the results of correlation hunting using component planes might not be accurate, as the perceived correlation might be caused by, for example, the cluster structure of the map (Vesanto 1999). That being said, component planes make it easy to discover variables that could be correlated with each other, and these variables can then be chosen to be further studied.

Examining the component planes in Figure 3, we can see that flowers with large values for sepal length, petal length, and petal width form one of the clusters; flowers with large values for sepal width, but small values for the other variables form the other cluster.

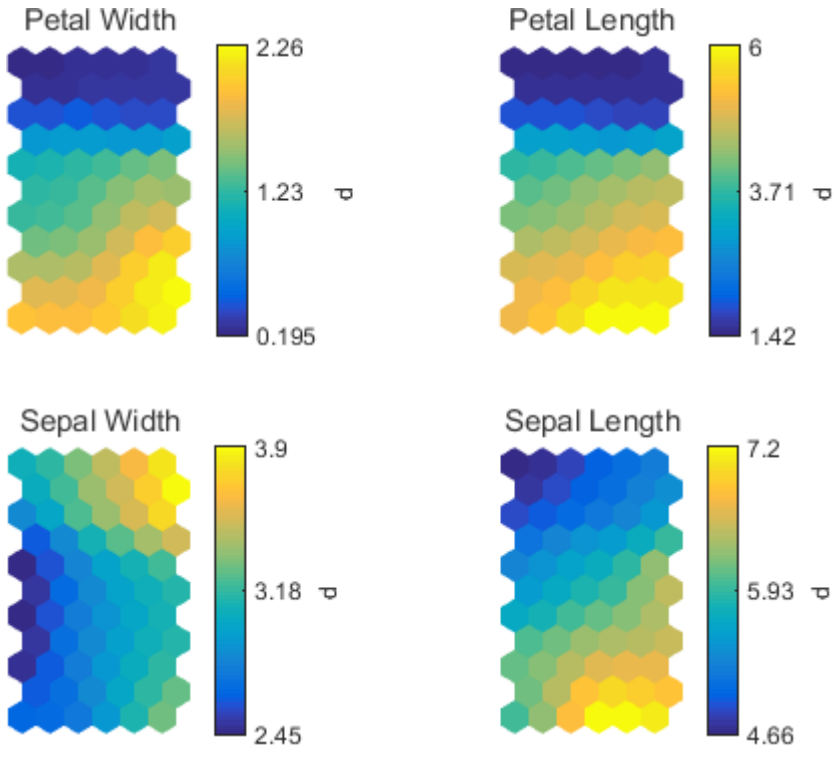

Figure 3 Component plane visualization of the Iris dataset.

Hit histograms can be used to visualize how the input vectors are distributed among the nodes of the map. This is important as the input vectors are not distributed equally. Some nodes will have a large proportion of the input vectors, while some nodes could have no input vectors at all. Multiple ways of visualizing hit maps exist. In Figure 4, the amount of or input vectors, or hits, for each node is visualized by the size of a hexagon; a larger hexagon denotes a larger amount of input vectors for the node.

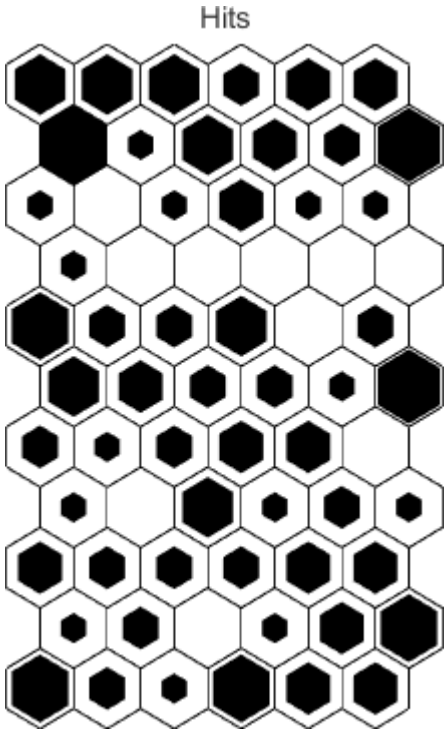

Figure 4 A hit histogram showing the number of hits for each node. A larger hexagon is used to represent a larger amount of hits for a node. The empty nodes on the map do not have any hits.

4 RESEARCH DATA

This part of the thesis will present the variables that were used to perform the analyses, discuss data preparation and preprocessing, and give an overview of the data characteristics in order to give the reader a better understanding of the type of data that was studied.

4.1 Data Variables

The research data used to perform the SOM analyses was provided to us by a Finnish retailer. The dataset consists of all the purchases made in the retailer's stores in Finland by the members of the retailer's own customer loyalty program. Each customer loyalty program member has been given their own customer ID, and all purchases made by a member can be linked to their customer ID. Purchases in both the online store and the brick & mortar stores are included in the dataset. The dataset spans the full calendar year 2014, and consists of the following type of data:

- Demographic data: gender, age, city of residence, registration date for the loyalty program.
- Purchasing behavior data: e.g. total purchase value, number of transactions, offline and online store purchase value, and the transaction dates.
- Email newsletter data: number of newsletters sent to the recipient, number of newsletters opened by the recipient, and the number of newsletters clicked by the recipient.

The data provided by the company was also used to calculate additional variables, such as 'recency' (of last transaction) and 'online value %' (share of online transaction value of total transaction value).

The full list of variables that were used to create various self-organizing maps are briefly explained below:

- **Age** in years
- **Gender** 0 for female, 1 for male
- **Store Presence** 0 if the case company's physical store does not exist in the customer's city of residence, 1 if a physical store exists in the customer's city of residence.
- **Membership Tenure** Number of days the customer has been a member of the loyalty program.
- **Recency** Number of days since the customer's last transaction, counted from December 31st 2014.
- **Frequency** Number of transactions during 2014. Maximum value set at 15.
- **Monetary** Total spending amount during calendar year 2014. Maximum value set at 1000€.
- **Online Value** Total spending amount in the company's online store. Maximum value set at 1000€.
- **Online Value %** Online spending amount / total spending amount
- **Newsletters Received** Number of newsletters received
- **Open Rate** Number of newsletters opened / number of newsletters received
- **Click-Through Rate** Number of newsletters clicked / number of newsletters received

4.2 Data Preparation and Preprocessing

In order to have reliable data mining results, the data needs to be preprocessed. This is an important and time-consuming task that needs to be completed in order to deal with noise and erroneous values that are bound to exist in real life business transaction data (Bose, Chen 2009).

Erroneous values for some of the variables were found in the dataset that we received from the case company. For example, the data for customer age was not

always reliable. In the dataset, the ages listed for the loyalty program members ranged from -7,984 years to 215 years. In order to deal with this issue, we removed all the values that were not in the range of 16 to 89 years. All the other data for these customers were kept and used in the analyses, only the erroneous value for age was removed. Similar logic was used to deal with erroneous values in other demographic variables as well.

The number of erroneous values in the dataset was not alarmingly big; the number of customers, whose demographic data was impacted, was an extremely small percentage of the total customer base. Because missing values in data do not have a significant impact on the performance of the self-organizing map (Bloom 2004), deleting the erroneous values was, in our view, the best way to handle the issue.

The dataset also included a number of variables with outliers. For example, the variable ‘monetary’ had outliers that could have affected the performance of the SOM. In this case, it was not appropriate to delete the value for the customers with extremely high values, as one of the key motivations for segmenting customers based on their RFM values, is to identify the most valuable customers of a company. After all, the highest spending customers - including the outliers - are the ones that companies should especially focus on when developing CRM strategies.

In order to deal with the outliers, we set a maximum value of 1000€ for the monetary variable. The data for all customers whose total spending exceeded 1000€ was manually modified and given the value of 1000€. By following this logic, we were able to include the high value customers in the dataset, but at the same time, make sure the SOM performed reliably by limiting the effects of the outliers. In order to decide the cut-off point, three RFM SOMs were created: one using a maximum value of 1000€; one using a maximum value of 2000€; and one with the original monetary values. After running these tests, it was decided that 1000€ should be used as the cut-off point. Similar logic was also applied to the variable ‘recency’. Maximum value for the recency variable was set at 15.

In order to ensure the reliability of the results, regarding the behavioral aspects of the study, some of the customers were filtered out of the final dataset based on their purchase frequency and monetary values. Customers who had made only one purchase during the one-year time period were not included in the dataset. The assumption for using behavioral variables to predict future behavior is that the customers' behavior will stay the same in the future (Liu, Lai et al. 2009). It was decided that the information gathered from just one transaction was not reliable enough to make assumptions on the future behavior of the customer. For this reason, at least two purchases during the year were required for the customer to be included in the dataset. Also, a minimum of 20€ spent during the year was required for the customer to be included in the dataset.

These limitations resulted in the dataset shrinking considerably. Only 48 percent of the total amount of customers in the original dataset were included in the dataset that was used to create the SOM analyses. However, one of the main motivations for carrying out an RFM analysis in practice, is to identify the most valuable customers of a company. The limitations that were applied did not have as large an impact in this regard, as the remaining customers accounted for 75 percent of loyalty program members' total purchase value. From a segmentation point of view, the customers who were excluded from the final dataset could be seen as a segment of their own. A rough description would be that they are one-time customers whose total purchase value is low. Because of the very large number of one-time customers, the case company should consider possible ways of converting some of them into repeat customers.

Euclidean metric is used to measure the distance between vectors in the SOM algorithm (Hanafizadeh, Mirzazadeh 2011). Therefore, variables in the dataset needed to be scaled in order to make them equally important. We scaled all variables linearly to make their variances equal to one. In the visualizations that are later presented, unscaled values are shown in order to make the interpretation of results simpler.

4.3 Data Characteristics

The final dataset that was used to create the self-organizing maps, included the data of 77,423 customer loyalty program members. The loyalty program members are mostly female, with men only making up 9 percent of the total amount of customers. Most of the customers (58%) live in a city where a retailer's brick and mortar store is also located. Online store usage has not widely caught up yet; only 8 percent of the customers had made a purchase in the company's online store during 2014. The value share of online purchases was 8% of total purchase value. Email newsletters are popular with the loyalty program members. 80 percent of the customers received at least one newsletter during 2014; 58 percent of all customers opened at least one newsletter; and 50 percent of all customers clicked a hyperlink in at least one of the newsletters.

Figure 5 presents bar charts of some of the key variables: (A) age, (B) purchase duration, (C) membership tenure, (D) recency, (E) frequency, and (F) monetary. Age of customers is normally distributed. Customers were on average 48 years old (median 49). Purchase duration shows the amount of days between the first and last transaction of a customer. 17 percent of the customers had a purchase duration of 30 days or less, otherwise customers are distributed relatively equally throughout duration lengths. The average purchase duration for customers was 161 days (median 160). Average loyalty program membership tenure was 1311 days (median 1459). The bar chart for the recency variable shows that a large percentage (35%) of the customers had made purchases in the last 30 days of the study period. This figure can be explained by seasonality; December is a busy sales period for the case company. Average value for recency was 85 days (median 66). Frequency of transactions is skewed with 67 percent of customers making only 2-3 transactions during the study period. Average value for frequency was 3 transactions (median 3). Monetary value of transactions is also heavily skewed. The total spending amount of 38 percent of customers was less than 100€, while the average spending

amount was 200€[1] (median 131€) for the study period. Binning of monetary values of over 1000€ explains the spike at the end of the bar chart.

[1] Average monetary value without setting a maximum value of 1000€ was 219€ (median 131€); average frequency value without setting a maximum value of 15 transactions was 4 (median 3).

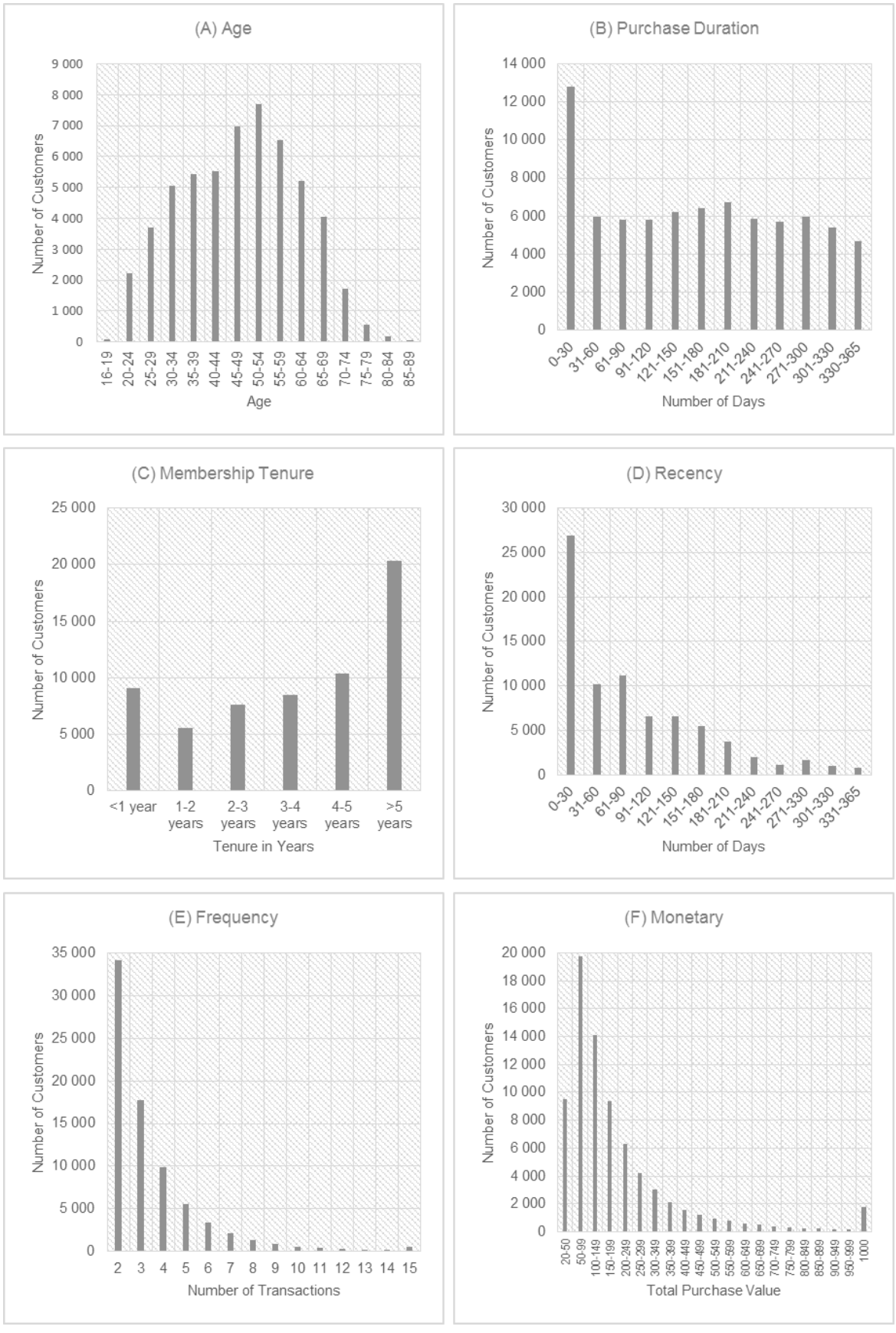

Figure 5 Bar charts of selected variables: (A) Age of Customer; (B) Purchase Duration; (C) Membership Tenure; (D) Recency; (E) Frequency; and (F) Monetary.

5 RESULTS OF THE EMPIRICAL RESEARCH

In this chapter, we will present the various self-organizing maps that were created using the data provided by the case company. Our ambition was to study the use of the SOM as a data clustering and exploration method in a CRM context, focusing especially on the RFM variables and the practical usability of the SOM.

All of the self-organizing maps in this thesis were created with the MATLAB program using the SOM Toolbox program package.

5.1 The Self-Organizing Map of RFM Variables

We first clustered the loyalty program members into segments by their purchasing behavior. In particular, we segmented the customers using the RFM variables, and focused on identifying the most valuable customers and customer segments among the loyalty program members.

After experimenting with different grid sizes for the SOM, it was decided that a map of approximately 100 nodes should be used. As we did not have prior knowledge of how many segments existed in the data or had any specific need to segment the customers into a specified number of segments, it was decided that the map should have a relatively large number of nodes, and that we should afterwards apply a second level of clustering to segment the SOM. Also, we did not want to create a map that had an abundance of nodes as it could have made it more difficult to use the map to perform segmentation.

Figure 6 presents the U-matrix and component planes of the SOM. Examining the U-matrix, and where it has its highest values, we can see that the nodes on the lower right corner of the map are clearly separated from the rest of the nodes on the map. Examining the individual component planes, we can see that the nodes with the highest values for the recency variable are concentrated on the upper right

corner of the map. Customers who had made a purchase very recently, and have thus a small recency value, are located at the bottom left corner of the map. Examining the component planes of the two other RFM variables, we can see that the frequency and monetary variables correlate with each other. Correlation between these two variables can be considered quite expected. The highest values for both, recency and frequency, are located at the bottom right corner of the map.

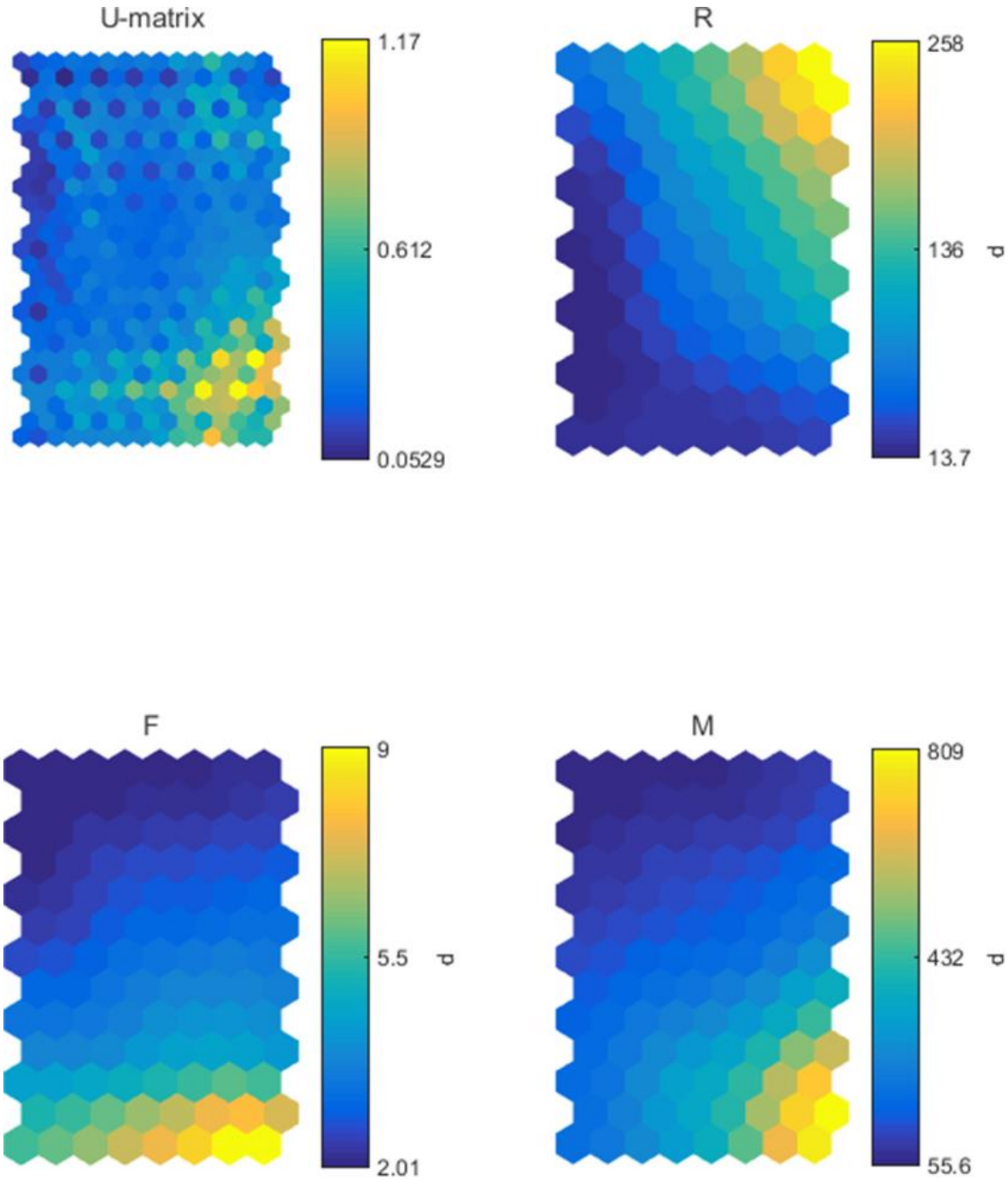

Figure 6 U-matrix and component planes of the SOM used to segment customers based solely on their purchasing behavior. Component planes are shown for the three RFM variables: R(ecency), F(requency), and M(onetary).

To create homogeneous segments that are large enough to serve, we applied a second level of clustering to segment the SOM. K-means clustering was used to perform the second level of clustering. The results of the second level of clustering are presented in Figure 7. Figure 7 also presents a hit histogram of the SOM. Seven

segments were created as a result of the second level of clustering, and will be discussed below.

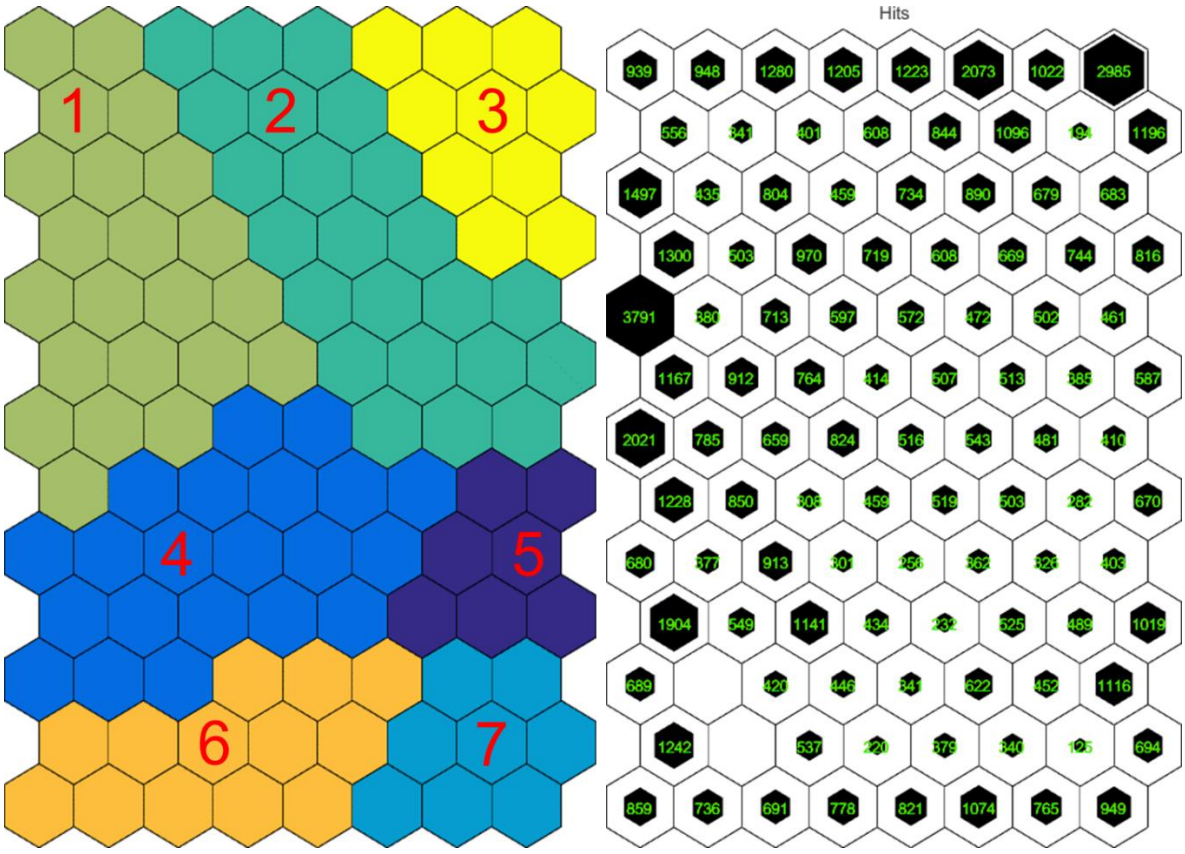

Figure 7 On the left, the seven segments that were created using the second level of clustering. On the right, a hit histogram showing the number of hits – the number of best matching units - for each node.

5.1.1 Customer Segments

- Segment 1 is comprised of loyalty program members who are the least valuable customers of the company. Customers in this segment spent the least money and made the fewest number of transactions during 2014. Customers in this segment have not made a transaction in a long time. Segment 1 is the largest of the seven segments, consisting of 28 percent of all customers.
- Segment 2 is comprised of customers whose transaction frequency and monetary value are low. Customers in this segment have not made a transaction in a long time. 19 percent of the loyalty program members belong to Segment 2.

- Segment 3 is quite similar to the two previous segments: the total purchase value and transaction frequency of customers in this segment is very low. Customers in this segment, have the highest values of any segment for the recency variable, which infers that they are very unlikely to make a repeat purchase when compared to the other segments. Customers in Segment 3 make up 15 percent of the total amount of customers.
- Segment 4 differs significantly from the three segments that were presented previously; customers in Segment 4 spend more and make transactions more frequently. Customers in this segment have also made their latest purchase very recently. 16 percent of all customers belong to Segment 4.
- Segment 5 is the smallest of the segments with 5 percent of customers. Despite its small size, this is a very interesting segment. Customers in this segment are the second highest spending customers of the company. They are, however, not frequent customers. In fact, frequency of transactions is very similar between customers in segments 4 and 5.
- Segment 6 is comprised of customers who make purchases often, and have made their latest purchase most recently of all the segments. Monetary value wise there is variance between customers in this segment, and some of the customers have very high values. 10 percent of the customers belong to Segment 6.
- Segment 7 is the company's most valuable customer segment. Customers in this segment spend significantly more than customers in any other segment. They make purchases the most often, and have made their latest transaction very recently. 7 percent, approximately 5,500 customers, belong to Segment 7.

5.1.2 Experiments with Different Grid Sizes

Before deciding to use the self-organizing map that was presented above, a number of maps with varying amounts of nodes were created in order to see how grid size affected the results and readability of the self-organizing map. Figure 8 shows the U-matrix and component planes for three maps with varying grid sizes. (A) Was created to segment the customers into four nodes, with the result being a 2x2 map;

(B) was created to segment the customers into ten nodes, with the result being a 2x5 map; and (C) was created by letting the SOM Toolbox decide the number of nodes into which the data were fitted, the result is a 48x29 map with 1392 nodes. We will not discuss these maps in detail as the final segmentation was done with a map of 100 nodes, and the results have already been presented. We will, however, quickly present the findings from running these tests.

Map A and B in Figure 8 present the results of assigning the customers into a predefined number of nodes. As each node of a SOM can be considered its own segment, this is an interesting way of segmenting customers, provided that it is known beforehand how many segments need to be created. In Figure 8, Map A shows the results of dividing the customer base into four segments. The node/segment on the lower right corner of this map is the most interesting one; the highest spending and most frequent customers are found in this segment. From the U-matrix we can see that the lower right corner and the upper right corner of the map are very clearly separated from each other. Customers in the upper right corner are the company's most infrequent customers with low values for the monetary and frequency variables. Map B shows the results of segmenting the customers into ten segments. From this map we can even more clearly see the correlation between the frequency and monetary variables. This map is also more useful in practice, if looking for the most valuable customers; it more clearly divides the customer base into smaller groups based on purchasing behavior, and high-spending customers are more clearly separated from the rest.

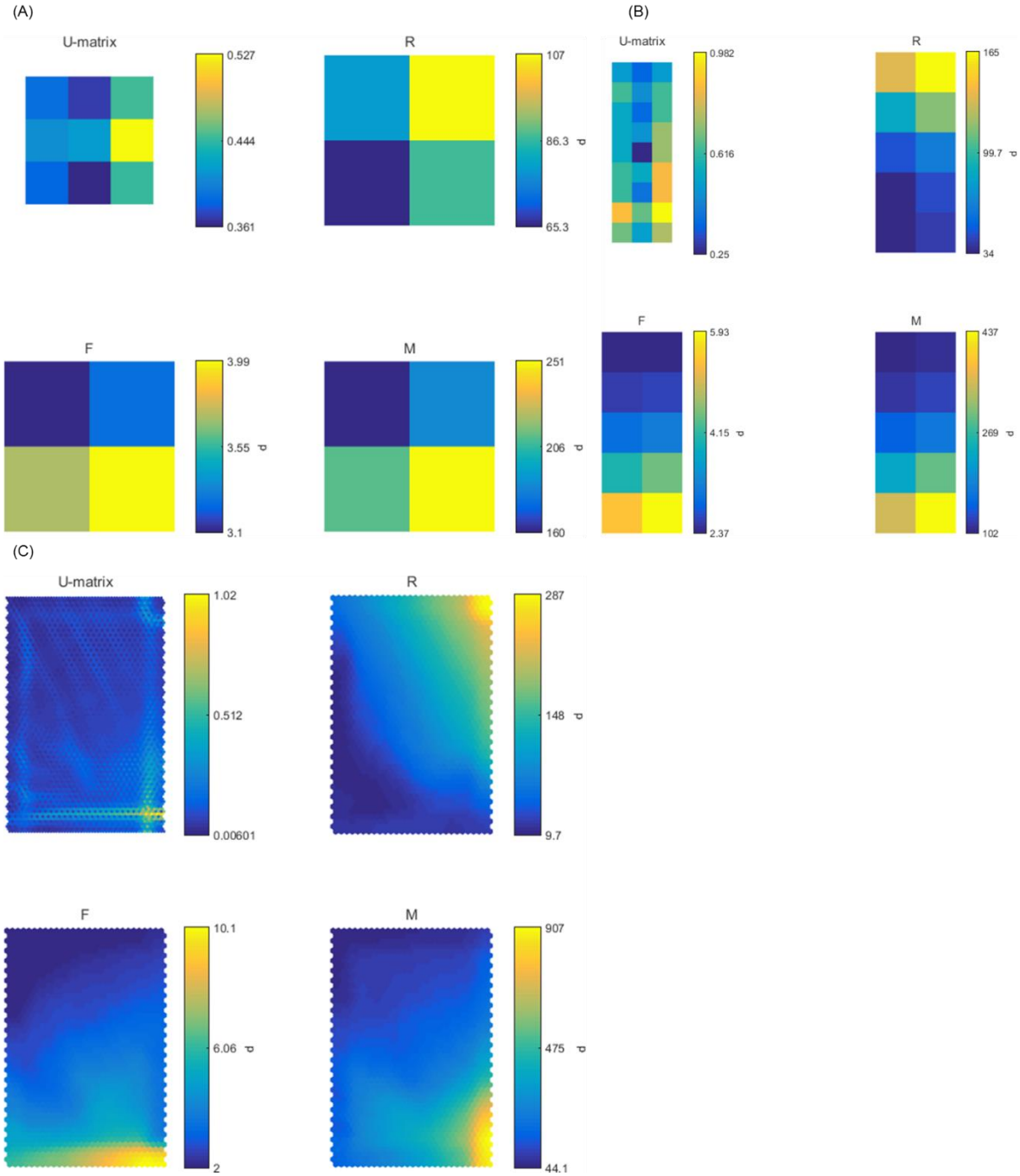

Figure 8 The U-matrix and component planes of three self-organizing maps. (A) a map with 4 nodes; (B) a map with 10 nodes; and (C) a map with 1392 nodes. Component plane R refers to Recency; F refers to Frequency; and M refers to Monetary.

Because maps A and B have a small number of nodes, some of the information in the data might be lost. Displaying fine details of the data is simply not possible with a limited number of nodes. To discover hidden data, we created one more SOM:

Map C, in Figure 8, has a much larger amount of nodes which makes it possible to display more information. However, it also makes it more difficult to use the map in segmentation tasks, as the individual nodes should no longer be considered segments of their own. When examining the U-matrix and component planes of map C, we can see information that was not visible in the maps A and B: the highest spending customers are divided into two groups based on the frequency of transactions.

Examining the different self-organizing maps that were created, we can see all of them displaying the same information with different levels of detail. All of them have their pros and cons, and a use case for each of them can be thought of. In this thesis, we decided to use a grid size of approximately 100 nodes to visualize most of the maps. For the dataset that we have, a map of this size provided a good level of detail and could still be used to visualize additional information, such as bar-planes and hit-histograms, without the map becoming too cluttered and difficult to interpret.

5.1.3 Discussion

The results of the SOM analysis show that it can be successfully applied to cluster customers into segments based on their RFM values. In this case, we identified seven customer segments. From the case company's perspective, the most interesting segments are segments 5, 6, and 7. The most valuable customers of the company belong to Segment 7; these customers should be the number one priority for the case company, and focus should be on ensuring that they remain satisfied and loyal to the company. The monetary value of customers in Segment 5 is also high. However, customers in this segment are not purchasing very frequently. The case company could consider possible ways, for example cross-selling, to increase the frequency of transactions for customers in Segment 5. Customers in Segment 6 make transactions frequently, but the total spending value of these customers is significantly lower than customers in segments 5 and 7. To increase the overall

value of customers in Segment 6, the case company could consider, for example up-selling, as a potential way of increasing the transaction value of these customers.

5.2 The Self-Organizing Map of RFM and Demographic Variables

In the previous analysis, customers were segmented into seven segments solely on the basis of their purchasing behavior. In the analysis that will be presented next, we added also demographic variables to the SOM. Our aim was to study and discover the demographic attributes of different customer segments and especially the customers whose transaction value is the highest, who purchase products the most often, and have made purchases most recently.

The SOM was created by including the RFM variables and four demographic variables (age, gender, store presence, and membership tenure) in the analysis. The U-matrix and component planes of the SOM are presented in Figure 9. Figure 10 presents a bar-plane visualization of the data. A bar-plane visualization is used to display the values of individual variables and to examine the characteristics of nodes and segments. Thus, the bar-plane visualization is an alternative way of visualizing the same information as component planes. In Figure 10, the values for each variable are scaled variable wise in order to fit the bars inside their corresponding node. In practice this means that each bar will be displayed with maximum range for the node where the variable gets its highest value, and with minimum range where the variable gets its lowest value. Results of the second level of clustering and a hit histogram are presented in Figure 11.

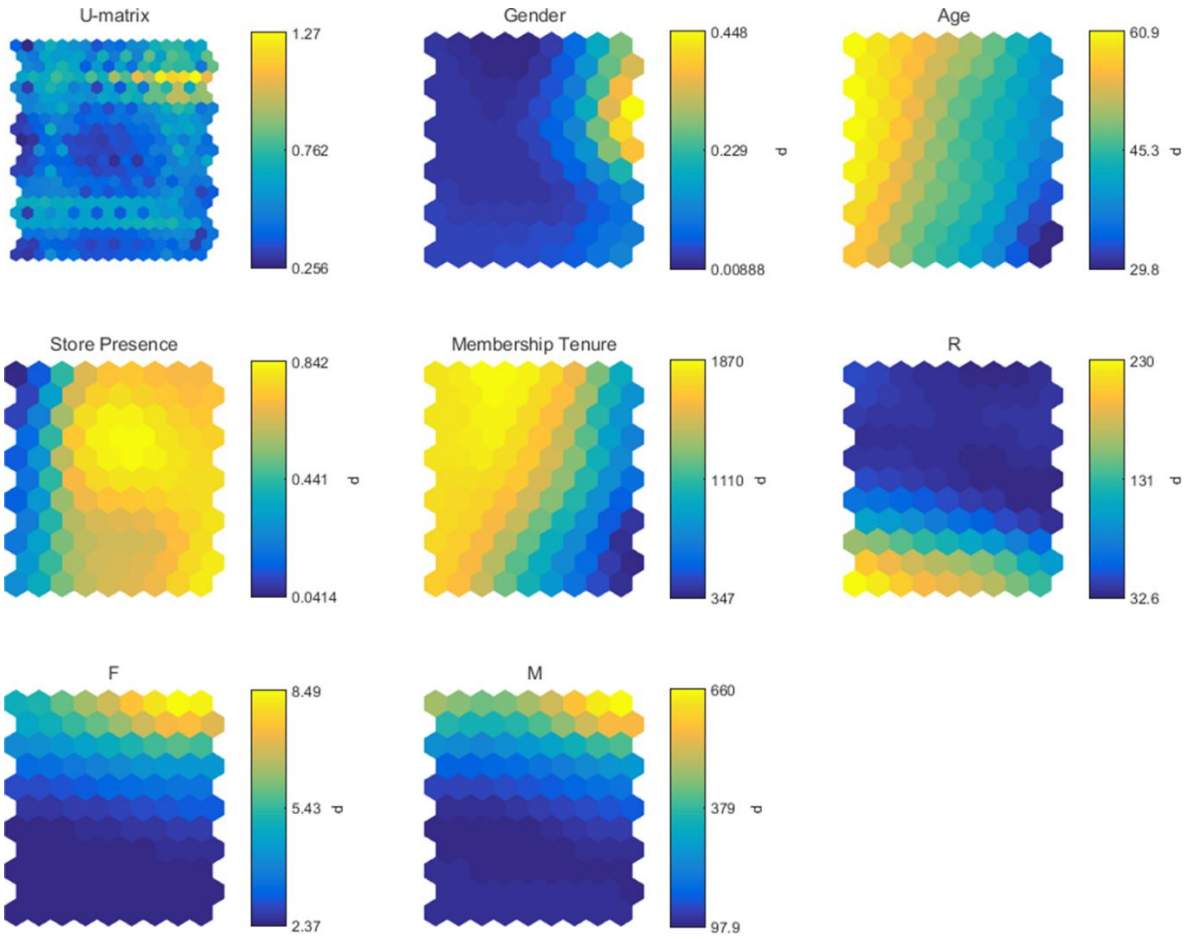

Figure 9 U-matrix and component planes of the SOM that was created using demographic and RFM variables: Gender, Age, Store Presence, Membership Tenure, (R)ecency, (F)requency, and (M)onetary.

Examining the U-matrix in Figure 9, we can see that the highest values are located at the top right corner of the map, implying that the underlying data in this part of the map differs greatly from other parts of the map. Nodes with relatively high values, separating clusters from each other, are also found in other parts of the map. A large area in the middle of the map, where nodes have very low values, implies that the underlying data is very homogeneous.

Examining the individual component planes in Figure 9 or the bar-plane visualization in Figure 10, we can see that most of the male customers are grouped into relatively few nodes on the right side of the map. Low values for age and membership tenure correlate with each other; the youngest and least senior customers are located at the lower right corner of the map. Customers who almost exclusively live in cities

with retail store presence are located in the center of the map. Frequency and monetary variables correlate heavily with each other, and the highest values for these two variables are located at the upper right corner of the map. The highest values for recency are located directly opposite of the high frequency and monetary values at the lower left corner of the map. Adding demographic variables to the SOM resulted in some loss in detail regarding the RFM variables. Examining the SOM that was created, we can no longer see the split between high-spending frequent and relatively in-frequent customers.

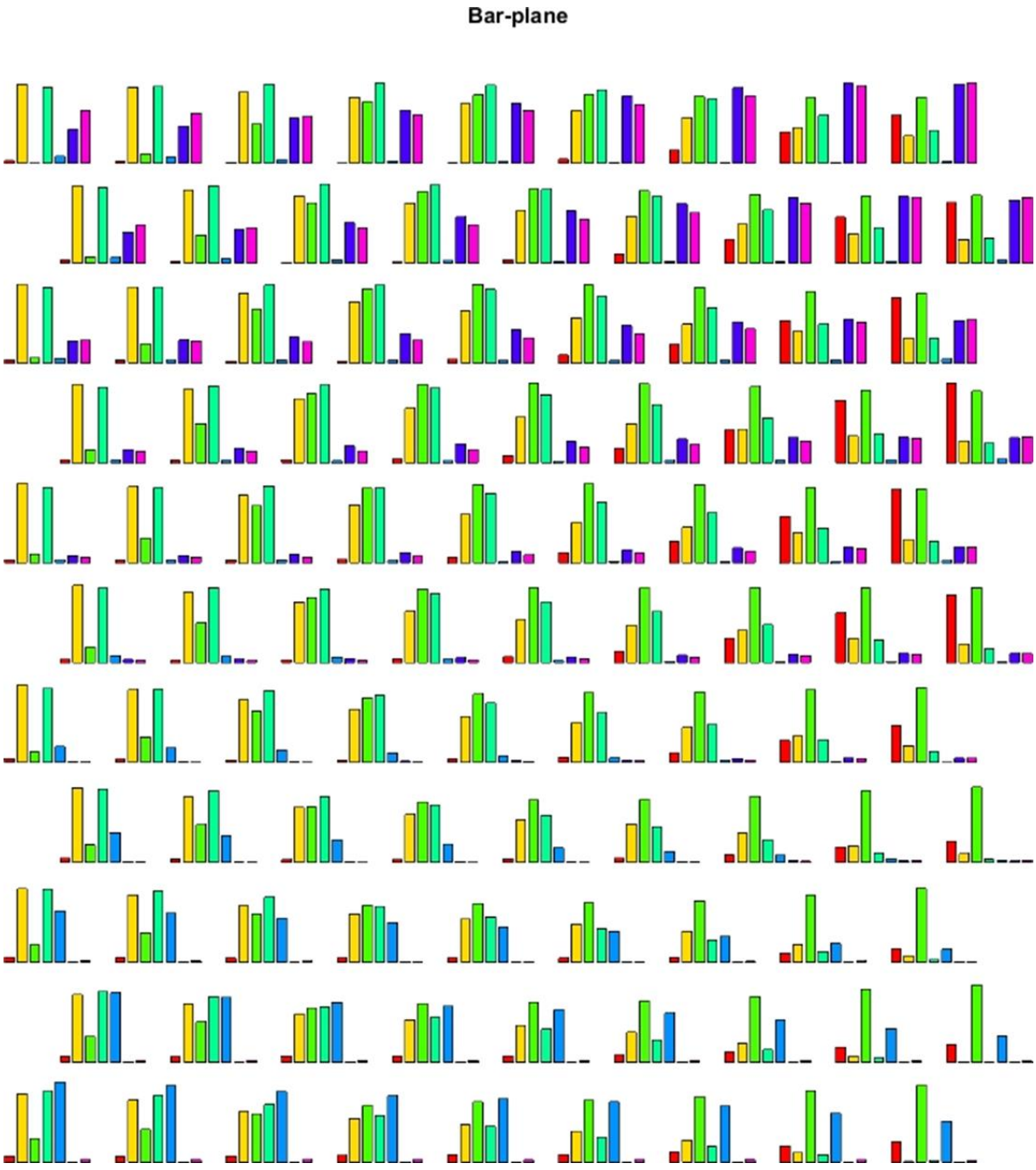

Figure 10 Bar-plane visualization of the demographic and RFM variables. Variables relating to each bar (from left to right): 1. Gender; 2. Age; 3. Store Presence 4. Membership Tenure; 5. Recency; 6. Frequency; 7. Monetary. The values are scaled variable wise to fit them inside their corresponding node. The bar gets its maximum range for the node where the variable gets its highest value, and minimum range in the node where the variable gets its lowest value.

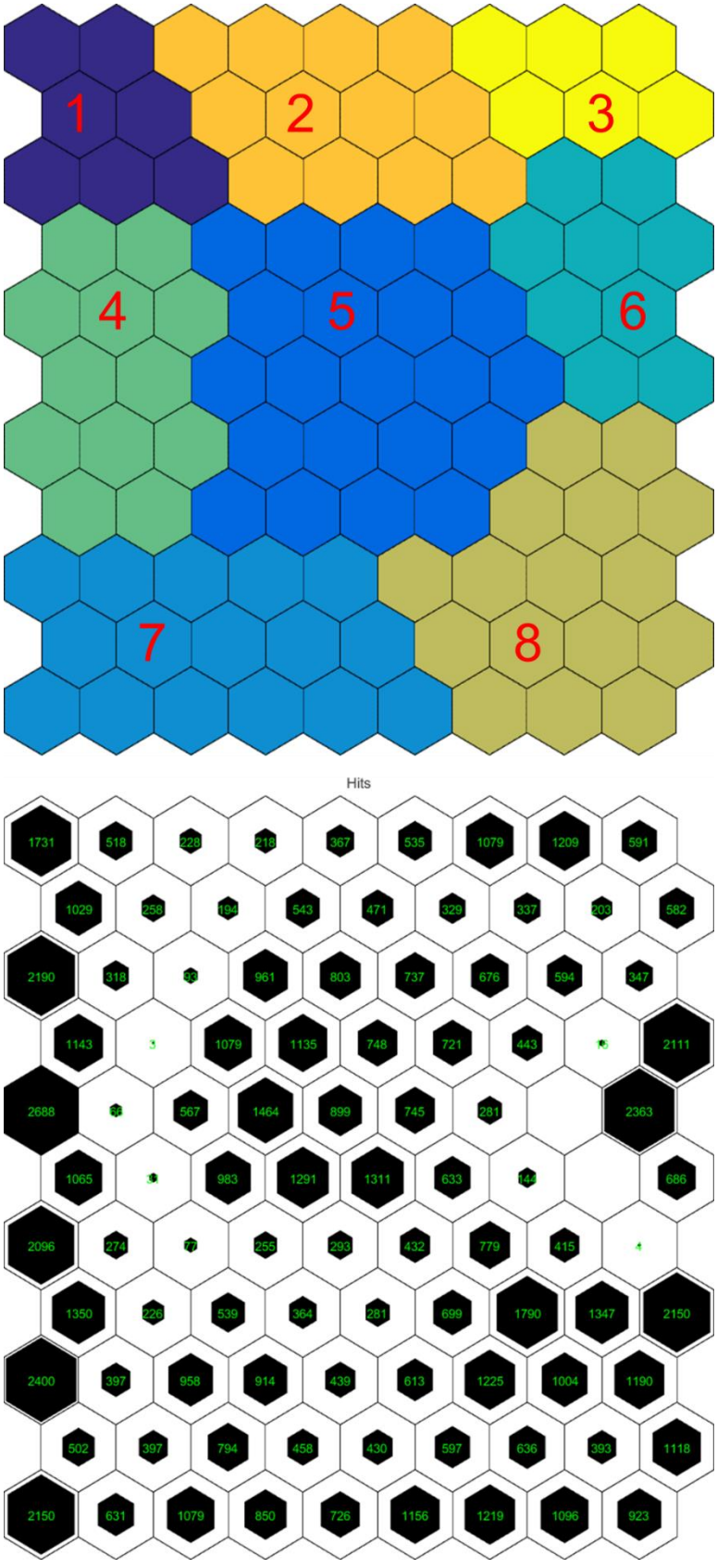

Figure 11 On the top: the eight segments that were created after applying the second-level of clustering. On the bottom: a hit histogram showing the number of hits – the number of best matching units - for each node.

A second-level of clustering was again applied to cluster the self-organizing map. As a result of applying the second level of clustering, the loyalty program members were segmented into eight segments based on their purchasing behavior and demographic characteristics. The eight segments are presented in Figure 11. Characteristics of each segment will be discussed below.

5.2.1 Customer Segments

- Segment 1 is comprised of female customers who are in, or close to, their sixties and have been members of the loyalty program for a long time. Customers in this segment live in a city or a municipality where the case company's retail store is not present. Despite not living in a city with a retail store, they are relatively frequent customers, and spend large amounts on purchases. Customers in this segment have made their latest purchase very recently. Segment 1 holds 8 percent of the total number of customers.
- Segment 2 is comprised of female customers who are approximately in their fifties. Customers in this segment have been members of the loyalty program for the longest time of any segment, and mainly live in a city with a physical store. Customers in this segment are also frequent shoppers whose purchase value is the second highest out of all the segments. The latest purchase of these customers was made very recently. Segment 2 holds 8 percent of the customers.
- Segment 3 is the segment with the company's most valuable customers. Customers in this segment spend significantly more on purchases than customers in other segments. They also make transactions much more frequently than customers in other segments. In addition, their latest purchase was also made very recently. Customers in this segment are mostly female and younger than customers in segments 1 and 2. Customers in this segment mostly live in cities where a retail store is located. Membership tenure wise, there are both relatively new and relatively old members in this segment. Segment 3 is the smallest of the eight segments; 5 percent of the loyalty program members belong to Segment 3.

- Segment 4 is comprised of customers whose demographic attributes resemble those of customers in Segment 1. However, customers in these two segments differ from each other significantly when analyzed by their purchasing behavior. Customers in Segment 4 spend significantly less and make far fewer transactions than customers in Segment 1. Segment 4 holds 12 percent of the total number of customers.
- Segment 5 is comprised of female customers in their forties. Customers in this segment live in a city that has a retail store, and they have been loyalty program members for a long time. Most of the customers in this segment have made their latest purchase very recently. Frequency and monetary value wise, customers in this segment are not among the most valuable ones of the company. 19 percent of the customers belong to Segment 5.
- Segment 6 holds the largest proportion of male customers out of all the segments. Customers in this segment are younger and their membership tenure is shorter than customers in most other segments. Based on the RFM variables, customers in this segment are quite valuable for the case company. Segment 6 holds 8 percent of the customers.
- Segment 7 holds the least valuable customers of the company based on the RFM variables. Customers in this segment spend the least money, make transactions most infrequently, and have not made any purchases recently. Examining the component planes of demographic variables, we can see that Segment 7 is comprised of female customers, some of whom live in cities that do not have a retail store. Regarding age and membership tenure, there is deviation between the customers within this segment, with values for these two variables running between the middle and high-ends of the value scales. Segment 7 holds 18 percent of the customers.
- Segment 8 is comprised of the youngest customers of the company. Customers in this segment have been members of the loyalty program for the shortest time. These customers are mostly female and live in a city that has a physical retail store. Judging by their purchasing behavior, customers in this segment are among the least valuable customers of the company. They make transactions infrequently and spend significantly less than customers on average. They have also not made purchases very recently.

Segment 8 is the largest of the segments with 20 percent of the total number of customers.

5.2.2 Discussion

By creating the SOM that was described above, we were able to segment the customers into eight segments based on their purchasing behavior and demographic attributes. The most valuable segments were identified with Segment 3 clearly standing out from all the other segments. Segments 1, 2, and 6 also gained relatively good values in regards to the RFM variables. From the case company's perspective, Segment 8 is also interesting as it is the largest segment and consists of the youngest demographic. Customers in this segment are new members of the loyalty program. Winning these customers over, transforming them into lifetime customers, and increasing the frequency and monetary value of their transactions presents both a challenge and an opportunity for the case company.

Adding demographic variables to the analysis not only made it possible to analyze the characteristics of customers who make up the most valuable segments, but it also made it possible to examine the purchasing behavior of certain demographic groups. For example, in this case the male customers of the company were mostly clustered into one of the eight segments and by examining the purchasing behavior of customers in that segment, we could also examine the purchasing behavior of the male customers.

5.3 The Self-Organizing Maps of Newsletter Consumption

The dataset provided by the case company included very interesting data related to the company's email newsletters and how they are consumed by the recipients. In the next analyses, we will use the SOM to explore information related to the email newsletters. We will create three SOMs: one that uses RFM variables and newsletter variables, one that uses online purchase variables and newsletter variables, and one that uses demographic variables and newsletter variables. With

these tests we will be able to study the practicality of the SOM as a method for performing data exploration.

Two of the three newsletter variables are percentages of total amount of newsletters received. In order to gain reliable information about customers' behavior related to how often they open and click hyperlinks on email newsletters, we filtered the original data to only include customers who had received at least ten newsletters during the year 2014. Otherwise, the behavioral data for customers who had only received very few newsletters would have been unreliable. The dataset used in this part of the thesis consists of data from 50,083 loyalty program members. On average these customers received 36 newsletters (median 38), opened 40 percent of received newsletters (median 31%), and clicked a hyperlink in 13 percent of all received newsletters (median 8%).

5.3.1 SOM of Newsletter and RFM Data

To create the first self-organizing map, we used RFM variables and newsletter variables. The SOM was constructed without limiting the size of the SOM grid. This was done in order to better visualize the fine structures that might exist within the data. The U-matrix and component planes of the SOM are presented in Figure 12. Because the resulting SOM is very large and the values of different nodes can be difficult to interpret easily, we visualized the same data on a U-matrix and component planes that only utilize four colors. The dramatically reduced number of available colors on the map makes it easier to get a quick view of the data and see a simplified representation of the values that different variables have in different parts of the map. This visualization is presented in Figure 13.

Examining the component planes of either Figure 12 or Figure 13, we can see information that was discovered already earlier: the frequency and monetary variables correlate heavily with each other, and have negative correlation with the recency variable. Customers who have received between 10 and 30 newsletters are split into two relatively small areas on the map. The component plane 'open rate'

clearly splits the map in the middle to customers who have opened approximately 45 percent or more of received newsletters, and customers who have opened a smaller percentage of received newsletters. Customers who have clicked a hyperlink on approximately 60 percent of the newsletters they received are grouped to a relatively small space on the bottom left corner of the map. Customers with a low click-through rate occupy the majority of the map area.

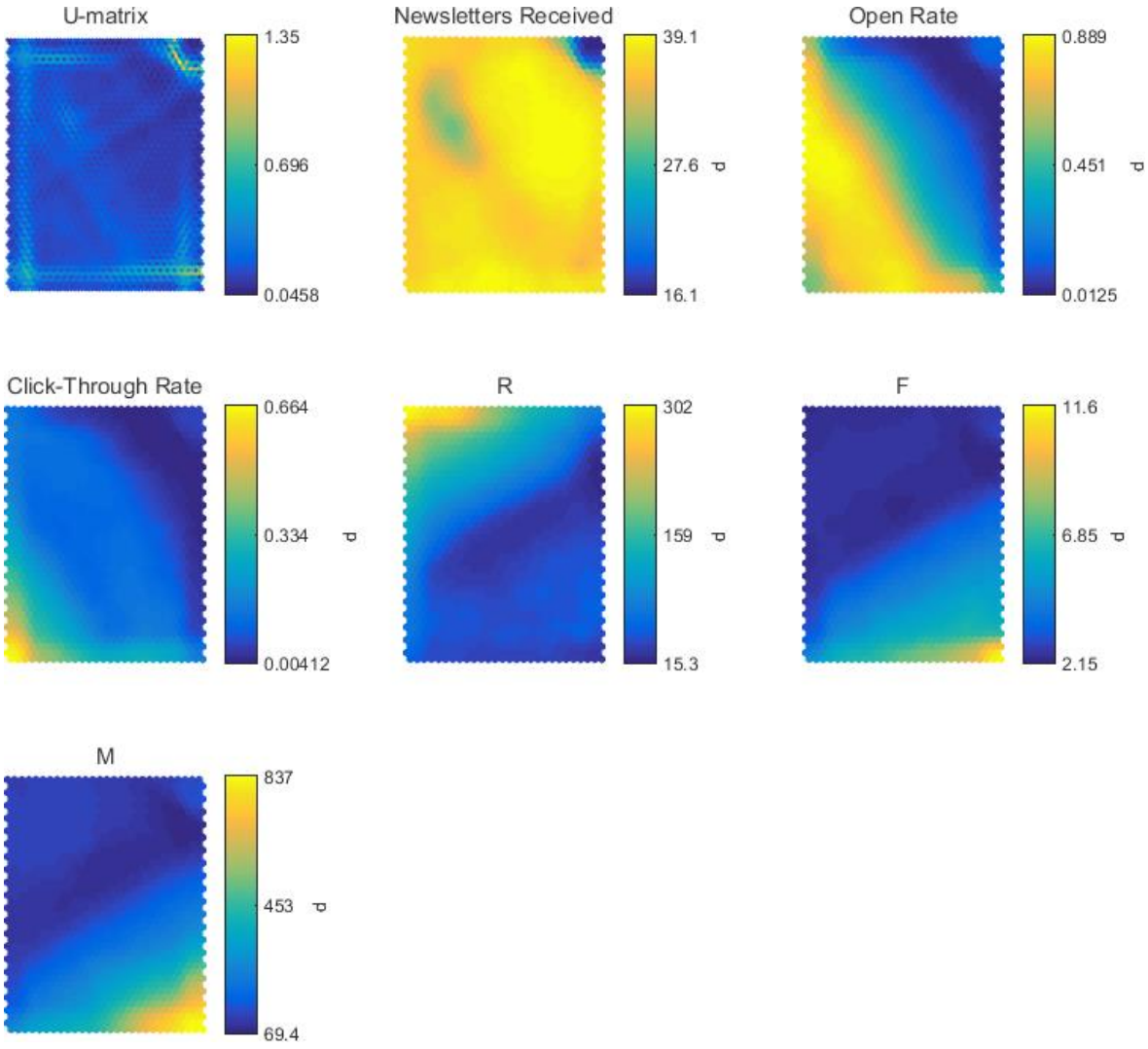

Figure 12 The U-matrix and component planes for the variables Newsletters Received, Open Rate, Click-Through Rate, R(ecency), F(requency), and M(onetary).

Examining the component planes we cannot see any direct correlation between the variables related to purchasing behavior and variables related to newsletter consumption. For example, the highest values for click-through rate are located at the bottom left corner of the SOM, and the highest values for the monetary variable are located at the bottom right corner of the SOM. There is, however, overlapping between the high values. For example, the highest spending customers are divided into those who received almost 40 emails during 2014 and opened approximately 80 percent of them, and those who received the same amount of emails, but opened almost none of them. This information shows that, although almost all of the loyalty program members are receiving the newsletters via email, some of the most valuable customers are not opening, let alone, reading them. Thus, different means of communication should be considered for those customers who are not reached, in practice, via newsletters. However, it is also important to note that that many of the high-value customers are receiving and reading the company's newsletters.

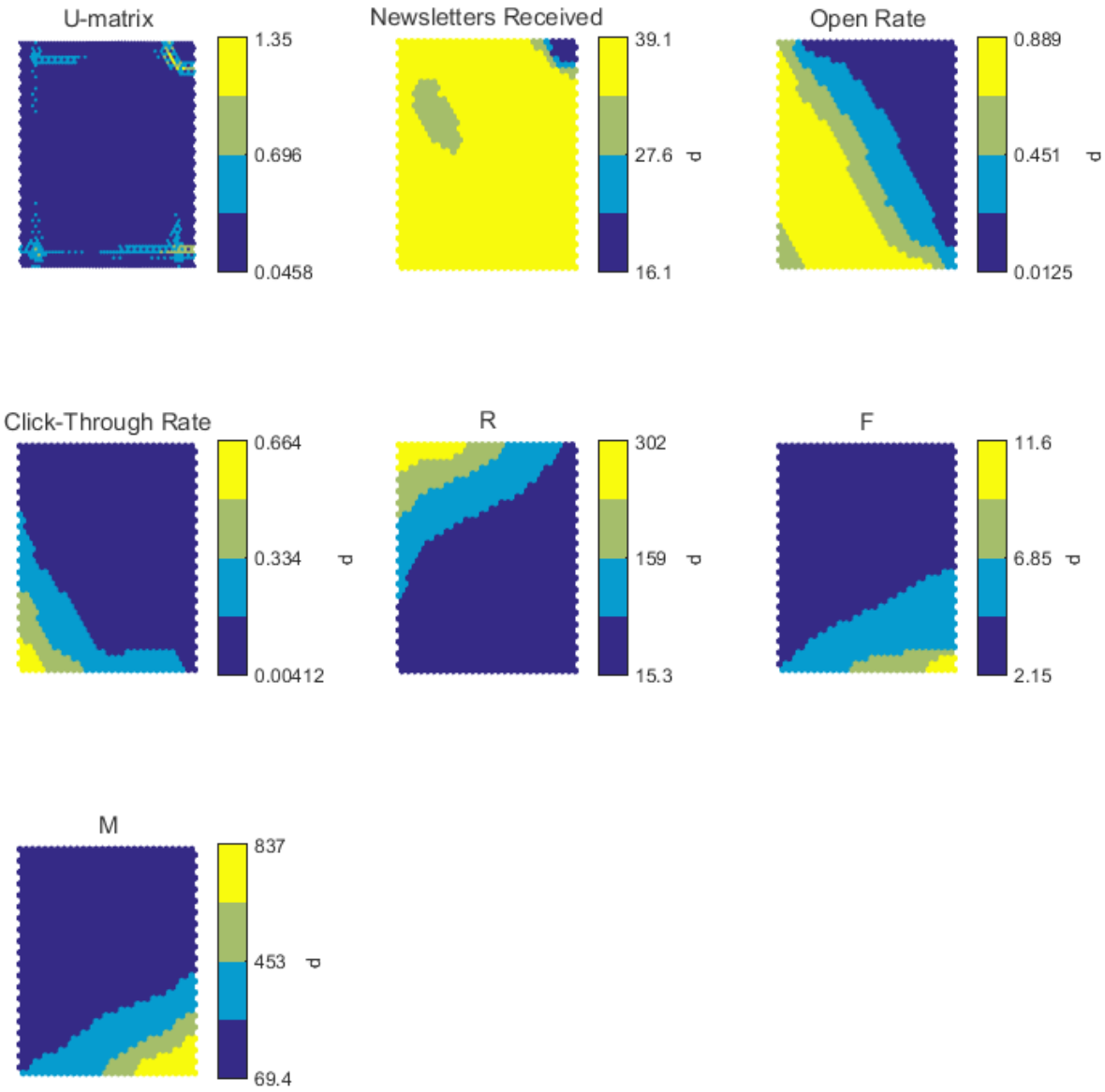

Figure 13 The U-matrix and component planes for the variables Newsletters Received, Open Rate, Click-Through Rate, R(ecency), F(requency), and M(onetary). The U-matrix and component planes are visualized by utilizing a color map of four colors.

5.3.2 SOM of Newsletter and Online Purchase Data

After examining the newsletter and RFM variables, we wanted to study how the results would be affected if we focused on online purchasing behavior. Clicking a hyperlink in an email newsletter will take the customer into the company's online store. Therefore, it was interesting to see whether we could see correlation between the click-through rate and online purchase value.

A similar analysis to the one presented before was run, this time using newsletter variables and online purchasing variables. The new variables used in this analysis are ‘online value’ which is the total value of online purchases and ‘online value %’ which is the online purchase share of a customer’s total purchase value. It should be noted that customers who did not make any online purchases during 2014 are also included in the analysis. The U-matrix and component planes of the analysis are presented in Figure 14, and the results are briefly discussed below.

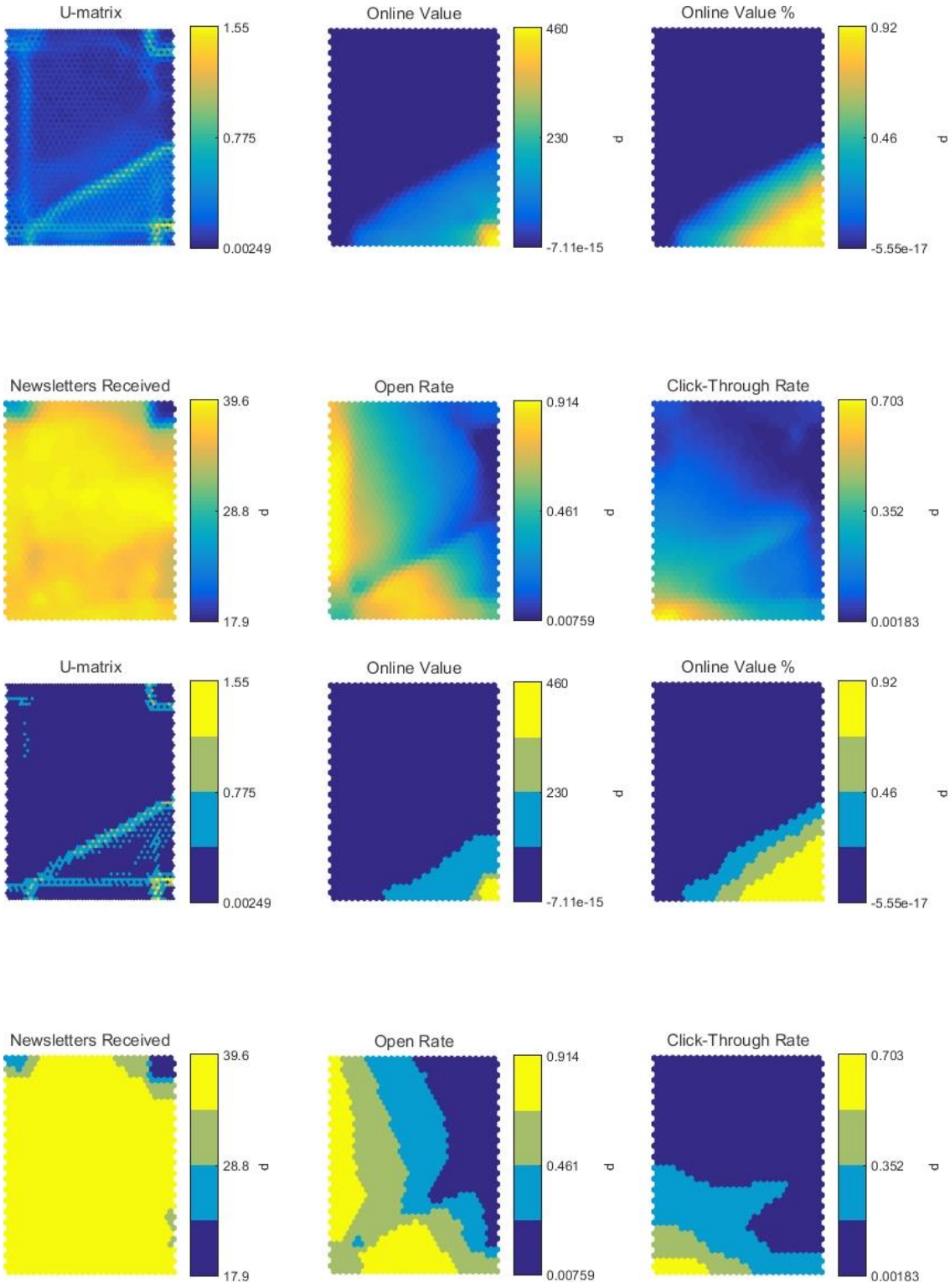

Figure 14 The U-matrix and component planes for the variables Online Value, Online Value %, Newsletters Received, Open Rate, and Click-Through Rate. U-matrix and component planes visualized with 64 colors (above) and 4 colors (below).

In Figure 14, by examining the U-matrix, we can see that the SOM is divided very clearly into two parts based on the variable ‘online value %’. The upper part of the map is comprised of customers who did not make any online purchases during 2014; the bottom right part of the map holds customers who made at least one purchase online. While not being mirror images of each other, the map halves do resemble each other regarding customers’ open rate and click-through rate. This means that there are no clear differences in newsletter consumption between customers who had made at least one online purchase and customers who had not made a single online purchase during 2014. Regarding the correlation between newsletter variables and online purchases, the results are similar to the previous analysis: no direct correlation exists. The customers who spent the most on online purchases opened approximately 45 percent of received emails and clicked a hyperlink in less than one third of received emails. Maybe surprisingly, customers with the highest click-through rates, made very few online purchases during 2014.

5.3.3 SOM of Newsletter and Demographic Data

Finally, as the last analysis, we created a SOM using newsletter variables and demographic variables to study who are the customers that are most likely to open newsletters and click the hyperlinks on them. In this analysis we built the map using a grid size of 11x9 in order to make the interpretation of the results simpler. The U-matrix and component planes are presented in Figure 15, and the results of this analysis are discussed below.

The U-matrix shows that the greatest distances between nodes are found at the upper left and bottom left of the map. Meaning that the customers who have been a loyalty program member for a short time and consequently received a smaller amount of newsletters, and the customers who are the most likely ones to click a hyperlink on a newsletter, are separated clearly from the rest of the customers. The first component plane shows that gender does not have a great effect on how likely a person is to open/click a newsletter. When compared to the SOM that was built using RFM and demographic variables, there is a clear difference on how the male

customers are situated on the map. In Figure 15, the male customers do not a form a uniform segment. In fact, they are spread - quite evenly - across the map.

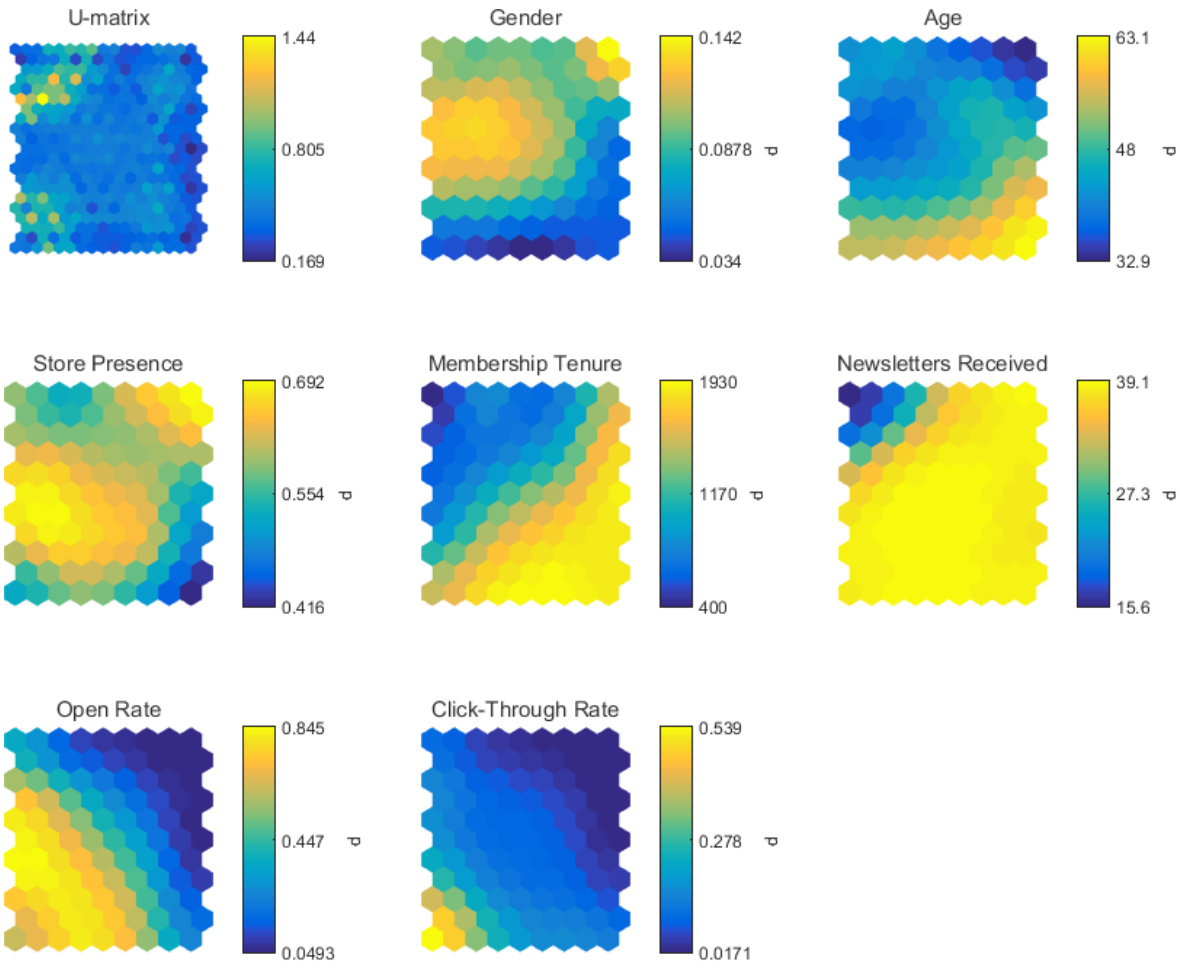

Figure 15 The U-matrix and component planes of demographic and newsletter variables: Gender, Age, Store Presence, Membership Tenure, Newsletters Received, Open-Rate, and Click-Through Rate.

The youngest group of customers is situated at the upper right corner of the map and the oldest group of customers is located at the bottom right of the map. The high values for store presence are situated at two separate parts of the map, on the left side and on the top right corner of the map. The lowest values are at the bottom right of the map. Membership tenure is shortest at the upper left corner of the map and peaks at the bottom of the map.

Judging by the component planes for the demographic variables, it is difficult to identify a unified group of customers who are more likely than others to open a newsletter they have received. The only variable whose highest values correlate with the highest values for opened newsletters is store presence; customers who live in a city with a retail store are more likely to open a newsletter. Otherwise the customers who are likely to open newsletters comprise of both male and female customers, aged between 30 and 60 years old, with a great variety in their membership tenure. Maybe a little surprisingly the least likely customers to open a newsletter are the youngest customers of the company. However, it must be noted that there is also a large group of young people who are among the customers that are the most likely ones to open a newsletter.

Customers who click on a hyperlink in more than 30 percent of received newsletters are located in an area of only a few nodes at the bottom left corner of the map. Based on the component planes for the demographic variables, we can identify these customers as being mostly female, around the age of 50, who have been members of the loyalty program for approximately four years or more. The least likely demographic to click on hyperlinks is situated at the top right corner of the map and is comprised of the company's youngest customers.

5.3.4 Discussion

In this part of the thesis, three self-organizing maps were created. The first SOM was created in order to study newsletter variables and RFM variables. The second SOM was used to study newsletter variables and online purchase variables. Finally, the third SOM was created using newsletter variables and demographic variables.

We did not find correlation between the newsletter variables and RFM/online purchase variables. The most valuable customers of the company are comprised of people who actively read the newsletters, as well as customers who do not read any of the newsletters. Therefore, in order to reach all valuable customers, the case company could consider the use of additional communication methods to

complement the newsletter communication. Whether the appropriate way of communication should be other ways of direct marketing, social media, or other such methods, is an interesting question. The last SOM analysis studied newsletter data and demographic variables. We did not find a homogeneous group of customers that was more likely than others to open newsletters.

6 DISCUSSION AND CONCLUSIONS

In this thesis, we studied how the self-organizing map could be used to explore and analyze a case company's loyalty program member data. The motivation for the study was to find out whether the SOM could be a useful and practical tool for companies to analyze their customer data, especially data related to customers' purchasing behavior, which was studied by focusing on the well-known RFM variables. In order to gain better knowledge related to these issues, we focused on three particular research questions that together can provide more information about the usability of the SOM in the context of analytical CRM.

The three research questions are answered below.

RQ1: Is the SOM a useful method for conducting customer segmentation using the RFM variables?

According to our experience from studying this question and creating the analysis, we can say that the SOM should be regarded as a useful method for conducting customer segmentation using the RFM variables. By creating the SOM analysis, we were able to quite easily divide the case company's customers into 7 segments based on their RFM values. We did not compare the results of our analysis to other clustering methods, but we can point out a couple of benefits that using the SOM has over the basic RFM model when performing this kind of an analysis in practice. (1) In the basic RFM model, customers are divided into equally sized quintiles based on each of the RFM variables. This would not have provided the best results when studying the case company's data as the monetary and frequency values were heavily skewed; dividing the customers into quintiles would have caused the most valuable customer to be binned into the same quintile with less valuable customers. Using the SOM as our analysis method, we were able to use the actual values for all studied variables and not bin them into quintiles. This made it possible to more accurately identify, for example, the most valuable customer segment, which was comprised of seven percent of the case company's customers. (2) The basic RFM model divides customers into 125 cells, which can then be grouped to create

segments. By using the SOM, we were able to easily and quickly divide the customers into just seven segments. (3) The results gained from the SOM analysis can be presented visually, which made the results easy to interpret.

RQ2: Is the SOM a useful method for conducting customer segmentation using the RFM and demographic variables?

Based on our analysis, the SOM should be regarded as a useful method for conducting customer segmentation using the RFM and demographic variables. This example even more clearly emphasized the usability of the self-organizing map in clustering customer data. Creating a similar analysis with an extended version of the basic RFM model - by including four additional variables - would have made the analysis very time consuming to perform, and the results difficult to interpret. By using the SOM as our analysis method, creating this analysis was very quick and easy.

Instead of the basic RFM model, other more advanced clustering methods could have been used to create this segmentation analysis with multiple variables. One of the advantages of the SOM over other advanced methods is that it can provide a visualization of the results. The clustering that was made divided the customer base into eight segments based on seven variables. Being able to visualize the segments and values for different variables made it very easy to compare the differences and similarities between the eight segments.

RQ3: Is the SOM a useful method for exploring email newsletter data?

In our view, the SOM should be considered a useful method for exploring email newsletter data. With the chosen method, we were able to very quickly create three different analyses related to the case company’s newsletter data. Because the SOM is an unsupervised data mining method, we did not need to have a priori information about the newsletter data. When analyzing the data, it was very useful that we could quickly create an analysis, grasp an understanding of the results, and move on to

the next analysis. Therefore, the SOM should be considered a very useful method for companies to explore large datasets, as it can be used to create multiple analyses very quickly.

As data volumes are ever increasing, practical ways to analyze large data masses are needed. We see the self-organizing map to have great potential in being used by companies to explore their customer data. Based on the analyses that were created, the greatest benefits of the SOM are (1) its ability to visualize multidimensional data and make the results easy to interpret, (2) and the fact that the SOM can be used to create new analyses with different variables very quickly without the need for a priori information regarding the data, which makes the SOM a very useful tool in data exploration.

Our findings on the positive aspects of the SOM are in line with previous studies that have focused on the use of the SOM as a clustering and data exploration tool. The capability of the SOM to visualize multidimensional data and make the results easy to understand has been described as one of the best aspects of the tool (Vesanto 1999), and has been highlighted also in the customer segmentation context (Holmbom, Eklund et al. 2011, Hanafizadeh, Mirzazadeh 2011). The fact that the SOM does not require a priori information has been highlighted previously by Bloom (2004) and Yao et al. (2010), who also commended the use of the SOM as a data exploration method.

Although we found the SOM to be a practical method for retail companies to cluster customers, it must be noted that these results are limited in regards to how actionable they are in practice. Yao et al. (2010) pointed out that, because the clusters created with the SOM are based on unsupervised learning, they should perhaps not be considered actionable in relation to marketing strategy. Instead, supervised learning methods should be used to create more actionable customer segments. The customer segments that were created in this thesis, therefore, only

provide an overall understanding of the case company's customers, and should be used as a first step in creating actual targetable customer segments.

6.1 Suggestions for Further Research

As we focused on the practical use of the SOM and how it could be used to study the case company's loyalty program data, our suggestions for further research are also heavily tied to the use of the SOM as a practical data analysis tool for companies. We will suggest a few ideas, how the case company could use the SOM method to gain more knowledge of their customers. These suggestions could also be studied with data from other case companies.

In this thesis, the SOM was found to be an efficient tool for clustering customers, especially when using both the RFM and demographic variables and thus creating the analysis with a relatively large number of variables. In the future, it would be interesting to add product-related data to the analysis, such as brand, product category, and price. From the case company's point of view, this information could be used to examine which products different customer segments tend to purchase. Also, results could be used to aid the decision making process of which product segments the company should focus more on, and to identify typical purchasing behavior and demographic characteristics of customers who prefer certain brands.

The newsletter data that we studied was very interesting. However, the insights that were gained were quite limited and did not provide a lot of information on how to improve communication with the customers in practice. In the future, it would be interesting to study the effectiveness of individual newsletters. For example, to study the open rate and click-through rate of individual newsletters. This could be used to identify the types of newsletters that are most likely to receive clicks. Also, it is expected that various customer segments will react differently to newsletters with different content. Therefore, by studying individual newsletters and click-through rates within different segments, we could identify newsletter content that is appropriate for each customer segment. Finally, the most interesting analysis would

be to study whether individual targeted newsletters can lead to increased sales for the company.

Finally, and maybe most importantly, future research should be done in close cooperation with the case company. This thesis was carried out with little knowledge of the case company's business. By conducting the research in cooperation, more reliable and actionable results could be achieved. Most importantly, it would be possible to follow through with the results of the SOM analyses; to implement marketing strategies that are based on the findings of the analyses, and to measure the effectiveness of the implemented strategies.

7 LIST OF REFERENCES

BLOOM, J.Z., 2004. Tourist market segmentation with linear and non-linear techniques. *Tourism Management,* **25**(6), pp. 723-733.

BOSE, I. and CHEN, X., 2009. Quantitative models for direct marketing: A review from systems perspective. *European Journal of Operational Research,* **195**(1), pp. 1-16.

CHAN, C.-.H., CHENG, C. and HSIEN, W., 2011. Pricing and promotion strategies of an online shop based on customer segmentation and multiple objective decision making. *Expert Systems with Applications,* **38**(12), pp. 14585-14591.

CHAN, C.C.H., 2008. Intelligent value-based customer segmentation method for campaign management: A case study of automobile retailer. *Expert Systems with Applications,* **34**(4), pp. 2754-2762.

CHEN, M., CHIU, A. and CHANG, H., 2005. Mining changes in customer behavior in retail marketing. *Expert Systems with Applications,* **28**(4), pp. 773-781.

CHENG, L. and SUN, L., 2012. Exploring consumer adoption of new services by analyzing the behavior of 3G subscribers: An empirical case study. *Electronic Commerce Research and Applications,* **11**(2), pp. 89-100.

COLLAN, M., EKLUND, T. and BACK, B., 2007. Using the Self-Organizing Map to Visualize and Explore Socio-Economic Development. *EBS Review,* (22), pp. 6-15.

HA, S.H., 2007. Applying knowledge engineering techniques to customer analysis in the service industry. *Advanced Engineering Informatics,* **21**(3), pp. 293-301.

HA, S.H., BAE, S.M. and PARK, S.C., 2002. Customer's time-variant purchase behavior and corresponding marketing strategies: an online retailer's case. *Computers & Industrial Engineering,* **43**(4), pp. 801-820.

HANAFIZADEH, P. and MIRZAZADEH, M., 2011. Visualizing market segmentation using self-organizing maps and Fuzzy Delphi method – ADSL market of a telecommunication company. *Expert Systems with Applications,* **38**(1), pp. 198-205.

HIZIROGLU, A., 2013. Soft computing applications in customer segmentation: State-of-art review and critique. *Expert Systems with Applications,* **40**(16), pp. 6491-6507.

HIZIROGLU, A. and SENGUL, S., 2012. Investigating Two Customer Lifetime Value Models from Segmentation Perspective. *Procedia - Social and Behavioral Sciences,* **62**(0), pp. 766-774.

HOLMBOM, A.H., EKLUND, T. and BACK, B., 2011. Customer portfolio analysis using the SOM. *International Journal of Business Information Systems,* **8**(4), pp. 396-412.

HONG, T. and KIM, E., 2012. Segmenting customers in online stores based on factors that affect the customer's intention to purchase. *Expert Systems with Applications,* **39**(2), pp. 2127-2131.

HSIEH, N., 2004. An integrated data mining and behavioral scoring model for analyzing bank customers. *Expert Systems with Applications,* **27**(4), pp. 623-633.

JONKER, J.J., PIERSMA, N. and VAN DEN POEL, D., 2004. Joint optimization of customer segmentation and marketing policy to maximize long-term profitability. *Expert Systems with Applications,* **27**(2), pp. 159-168.

KHAJVAND, M., ZOLFAGHAR, K., ASHOORI, S. and ALIZADEH, S., 2011. Estimating customer lifetime value based on RFM analysis of customer purchase behavior: Case study. *Procedia Computer Science,* **3**(0), pp. 57-63.

KIANG, M.Y., HU, M.Y. and FISHER, D.M., 2006. An extended self-organizing map network for market segmentation—a telecommunication example. *Decision Support Systems,* **42**(1), pp. 36-47.

KOHONEN, T., 2014. *MATLAB Implementations and Applications of the Self-Organizing Map.* 1st edn. Helsinki: Unigrafia Oy.

KOHONEN, T., 2013. Essentials of the self-organizing map. *Neural Networks,* **37**, pp. 52-65.

KOTLER, P. and KELLER, K.L., 2011. *Marketing Management.* 14th edn. New Jersey: Prentice Hall.

LI, D., DAI, W. and TSENG, W., 2011. A two-stage clustering method to analyze customer characteristics to build discriminative customer management: A case of textile manufacturing business. *Expert Systems with Applications,* **38**(6), pp. 7186-7191.

LING, R. and YEN, D.C., 2001. Customer Relationship Management: An Analysis Framework and Implementation Strategies. *Journal of Computer Information Systems,* **41**(3),.

LIU, D., LAI, C. and LEE, W., 2009. A hybrid of sequential rules and collaborative filtering for product recommendation. *Information Sciences,* **179**(20), pp. 3505-3519.

NGAI, E.W.T., XIU, L. and CHAU, D.C.K., 2009. Application of data mining techniques in customer relationship management: A literature review and classification. *Expert Systems with Applications,* **36**(2, Part 2), pp. 2592-2602.

OJA, M., KASKI, S. and KOHONEN, T., 2003. Bibliography of Self-Organizing Map (SOM) Papers: 1998-2001 Addendum. *Neural Computing Surveys,* **3**, pp. 1-156.

OLSON, D.L. and CHAE, B., 2012. Direct marketing decision support through predictive customer response modeling. *Decision Support Systems,* **54**(1), pp. 443-451.

PHAN, D.D. and VOGEL, D.R., 2010. A model of customer relationship management and business intelligence systems for catalogue and online retailers. *Information & Management,* **47**(2), pp. 69-77.

RYGIELSKI, C., WANG, J. and YEN, D.C., 2002. Data mining techniques for customer relationship management. *Technology in Society,* **24**(4), pp. 483-502.

SARLIN, P. and YAO, Z., 2013. Clustering of the Self-Organizing Time Map. *Neurocomputing,* **121**, pp. 317-327.

SHIM, B., CHOI, K. and SUH, Y., 2012. CRM strategies for a small-sized online shopping mall based on association rules and sequential patterns. *Expert Systems with Applications,* **39**(9), pp. 7736-7742.

SHIN, H.W. and SOHN, S.Y., 2004. Segmentation of stock trading customers according to potential value. *Expert Systems with Applications,* **27**(1), pp. 27-33.

VESANTO, J. and ALHONIEMI, E., 2000. Clustering of the Self-Organizing Map. *IEEE Transactions on Neural Networks,* **11**(3),.

VESANTO, J., 1999. SOM-based data visualization methods. *Intelligent Data Analysis,* **3**(2), pp. 111-126.

WANG, Y., CHIANG, D., HSU, M., LIN, C. and LIN, I., 2009. A recommender system to avoid customer churn: A case study. *Expert Systems with Applications,* **36**(4), pp. 8071-8075.

WEI, J., LEE, M., CHEN, H. and WU, H., 2013. Customer relationship management in the hairdressing industry: An application of data mining techniques. *Expert Systems with Applications,* **40**(18), pp. 7513-7518.

WEI, J., LIN, S., WENG, C. and WU, H., 2012. A case study of applying LRFM model in market segmentation of a children's dental clinic. *Expert Systems with Applications,* **39**(5), pp. 5529-5533.

YAO, Z., HOLMBOM, A., EKLUND, T. and BACK, B., 2010. Combining unsupervised and supervised data mining techniques for conducting customer portfolio analysis, *In Proceedings of the 10th Industrial Conference on Data Mining (ICDM 2010)*, 12-14 July, 2010 2010, pp. 292-307.

YAO, Z., SARLIN, P. and EKLUND, T., 2014. Combining visual customer segmentation and response modeling. *Neural Computing and Applications,* **25**(1),.