CUSTOMER SEGMENTATION USING A DEVELOPED
RFM MODEL: AN APPLICATION IN A RUG&CARPET
MANUFACTURING COMPANY

A THESIS
SUBMITTED TO ABDULLAH GÜL UNIVERSITY
SOCIAL SCIENCES INSTITUTE
IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF MASTER OF SCIENCE

By
Yağmur Gizem İMDAD
August, 2022
Kayseri, TURKEY

CUSTOMER SEGMENTATION USING A DEVELOPED RFM MODEL: AN APPLICATION IN A RUG&CARPET MANUFACTURING COMPANY

A THESIS

SUBMITTED TO ABDULLAH GÜL UNIVERSITY

SOCIAL SCIENCES INSTITUTE

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS

FOR THE DEGREE OF MASTER OF SCIENCE

By

Yağmur Gizem İMDAD

August, 2022

Kayseri, TURKEY

SCIENTIFIC ETHICS COMPLIANCE

I hereby declare that all information in this document has been obtained in accordance with academic rules and ethical conduct. I also declare that, as required by these rules and conduct, I have fully cited and referenced all materials and results that are not original to this work.

Name-Surname: Yağmur Gizem İMDAD

Signature :

REGULATORY COMPLIANCE

M.Sc. thesis Titled Customer Segmentation Using a Developed RFM Model: An Application in a Rug&Carpet Manufacturing Company has been prepared in accordance with the Graduate Thesis Preparation Guidelines of the Abdullah Gül University, Social Sciences Institute.

Prepared By   Advisor

Yağmur Gizem İMDAD   Prof. Dr. Cengiz YILMAZ

Signature   Signature

Head of the Data Science Program

Assoc. Prof. Umut TÜRK

Signature

ACCEPTANCE AND APPROVAL

M.Sc. thesis titled Customer Segmentation Using a Developed RFM Model: An Application in a Rug&Carpet Manufacturing Company and prepared by Yağmur Gizem İMDAD has been accepted by the jury in the Data Science Graduate Program at Abdullah Gül University, Social Sciences Institute.

23 / 08 / 2022

JURY:

Advisor : Prof. Dr. Cengiz YILMAZ

Member : Assoc. Prof. Umut TÜRK

Member : Assoc. Prof. Kumru UYAR

APPROVAL:

The acceptance of this M.Sc thesis has been approved by the decision of the Abdullah Gül University, Social Sciences Institute, Management Board dated ….. /….. / ……….. and numbered …………..……..

……….. /……….. / ………..

Director of Social Sciences Institute
Assoc. Prof. Umut TÜRK

Name Surname	: Yağmur Gizem İMDAD
Program	: Data Science (M.Sc.)
Advisor	: Prof. Dr. Cengiz YILMAZ
Thesis Title	: Customer Segmentation Using A Developed RFM Model:
	An Application in a Rug&Carpet Manufacturing Company
Date and Place	: August, 2022 – Kayseri, Turkey

ABSTRACT

Data science has gained enormous importance by contributing to the in-depth understanding and interpretation of information. Especially companies consult on data analysis to make strategic decisions in the competitive market. Much more important than the decisions taken is a determination of the customer or customer groups to which these decisions will be adapted. For that reason, customer segmentation by identifying similarities and differences between customers becomes crucial. In recent times, the RFM model is preferred mostly for customer segmentation. The RFM model is based on the customer's last purchase date, how often they purchase, and how much money contributes to the company. It is an easy model to understand and interpret results in a clear way. Many researchers prefer to apply the RFM method by adding extra variables to the analysis. Thus, customers are evaluated from a broader perspective. This study aims to present a developed RFM model by adding extra variables which are Loyalty, Dependence, and Expectation which are determined by a broad literature review and as a result of a survey relating to 106 dealers. There are some studies that create a segmentation model by using loyalty and the RFM model. However, this study developed a new model by including the dependence and expectation variables, which are not been used previously with the RFM model, besides loyalty. In the study, dealers are analyzed by the K-means clustering method and the optimum number of clusters is indicated as six. Each cluster has its specific customer behavior and this study guides the company to constitute marketing strategies regarding customers' specifications.

Keywords: RFM, K-Means Clustering, Data Science, Customer Segmentation

Ad Soyad	: Yağmur Gizem İMDAD
Anabilim Dalı, Program	: Veri Bilimi (Yüksek Lisans)
Tez Danışmanı	: Prof. Dr. Cengiz YILMAZ
Tez Başlığı	: Geliştirilmiş RFM Modeli ile Müşteri Segmentasyonu: Bir Halı ve Kilim Üretici Firmasında Uygulama
Tarih ve Yer	: Ağustos, 2022 – Kayseri, Türkiye

ÖZET

Veri bilimi, bilginin derinlemesine anlaşılmasına ve yorumlanmasına katkıda bulunarak büyük önem kazanmıştır. Özellikle şirketler, rekabetçi pazarda stratejik kararlar almak için veri analizine başvurmaktadır. Alınan kararlardan çok daha önemli olan, bu kararların hangi müşteri ya da müşteri gruplarına uyarlanacağının belirlenmesidir. Bu nedenle müşteriler arasındaki benzerlikleri ve farklılıkları belirleyerek müşteri segmentasyonu yapmak çok önemli hale geliyor. Son zamanlarda müşteri segmentasyonu için RFM modeli çoğunlukla tercih edilmektedir. RFM modeli, müşterinin son satın alma tarihine, ne sıklıkla satın aldığına ve şirkete ne kadar para kazandırdığına dayanmaktadır. Sonuçları net bir şekilde anlamak ve yorumlamak için kolay bir modeldir. Birçok araştırmacı analize ekstra değişkenler ekleyerek RFM yöntemini uygulamayı tercih etmektedir. Böylece müşteriler daha geniş bir perspektiften değerlendirilir. Bu çalışma, geniş bir literatür taraması ve 106 bayiye ait anket sonucunda belirlenen Sadakat, Bağımlılık ve Beklenti değişkenleri eklenerek geliştirilmiş bir RFM modeli sunmayı amaçlamaktadır. Sadakat ve RFM modelini kullanarak segmentasyon modeli oluşturan bazı çalışmalar bulunmaktadır. Ancak bu çalışma, sadakatin yanı sıra RFM modeli ile daha önce kullanılmayan bağımlılık ve beklenti değişkenlerini de dahil ederek yeni bir model geliştirmiştir. Çalışmada, bayiler K-ortalama kümeleme yöntemi ile analiz edilmiş ve optimum küme sayısı altı olarak belirtilmiştir. Her kümenin kendine özgü müşteri davranışı vardır ve bu çalışma şirkete, müşterilerin özelliklerine göre pazarlama stratejileri oluşturma konusunda rehberlik eder.

Anahtar kelimeler: RFM, K-Ortalamalı Kümeleme, Veri Bilimi, Müşteri Segmentasyonu

ACKNOWLEDGEMENT

I would like to express my gratitude to my advisor, Prof. Dr. Cengiz Yılmaz for supporting and thoughts that sheds light on my research. Also, it is a pleasure to indicate my thanks to Assoc. Prof. Umut Türk for his suggestions, ideas and motivational speeches throughout the research.

Furthormore, I would like to thank my jury member who is Assoc. Prof. Kumru UYAR for sharing her valuable comments.

Finally, I would like to endless thanks to my family and all my close friends especially Edanur Kılıç for supports and motivating me throughout my thesis process.

TABLE OF CONTENTS

LIST OF ABBREVIATIONS

RFM	Recency, Frequency, Monetary
WCSS	Within Cluster Sum of Squares
ANOVA	Analysis of Variance
RFMLED	Recency, Frequency, Monetary, Loyalty, Expectation, Dependence

LIST OF TABLES

LIST OF FIGURES

LIST OF GRAPHS

1. INTRODUCTION

Data is vital to comprehend people's behavior or operations in many organizations and increases its importance day by day by helping to make good decisions. Therefore, many organizations record and store information in a storage system. Especially, companies desire to know customers' preferences to manage their relationships with customers. Surely, the interpretation of data is as important as being stored. Collecting the right data and analyzing data with the right methods in line with the targets provide crucial benefits to the company.

Understanding the preferences of customers in the rapidly changing market and offering products and services gives the company a huge advantage in the competitive market. Because there are many customers with different demands, it is tremendously important to figure out customers and separate them into segments. Hence, a customer-oriented view becomes more important than a product-oriented view and this situation increases the importance of customer segmentation.

Segmenting customers is a major issue in determining the focus of the company in the carpet industry, as in many other industries. However, there is little work in the literature on this matter. No matter how good a product or service is, it means nothing if it does not reach the right customer. The main purpose of clarifying customer segmentation is to specify the customer group to be focused on truthfully. Thus, maintaining profitability with customer-oriented products and pricing, the realization of different service delivery for each customer group, and effective use of sales channels are provided.

Customers can be divided into groups by choosing different criteria for each organization in customer segmentation. Recently, Recency, Frequency, and Monetary values are used widely and preferred by many scientists (Khajvand & Tarokh, 2011; Cheng & Chen, 2009; Spring et al., 1999). Hence, it is researched to obtain more handy models by adding other variables to these three variables (Wu et al., 2014; Hajiha et al. 2011; Hossein, 2010). Therefore, this study handles RFM values with other variables, which are loyalty, expectation and dependence to take a broad perspective.

Although the importance of customer segmentation in customer relationship management is known, few studies are using a developed RFM model with loyalty and no work uses expectation or dependence as a segmentation criteria with RFM. As a result of that, this study aims to find out a developed RFM model in customer segmentation and specifically segments customers of a rug&carpet company, contributing to the customer segmentation analysis.

2. LITERATURE REVIEW

This section handles the relevant literature to the field of the recency, frequency, and monetary (RFM) model, also a customer segmentation concept.

Customer segmentation has always been a crucial matter for businesses in creating a competitive advantage in the market. It is the process of dividing customers into separated parts based on their similar needs and buying behavior. Thus, companies can create specific services and products and present unique marketing plans for each class.

Customer segmentation can be done with various methods. Notwithstanding, all methods have the common point which analyzes data and specifies the similarities and dissimilarities, and finally categorizes them into different groups. The RFM model is one of the customer segmentation approaches and has been implemented broadly for customer value analysis (Dogan et al., 2018). This model separates into customers using three attributes, namely recency, frequency, and monetary. Recency represents the length of time since the last purchase, frequency is the number of purchases in a given period, and monetary is the amount of money spent in a particular period.

In 2019, a study generated by Jintana and Mori was carried out on the application of data analytics techniques to a courier business which is Green Bus to manage customer segmentation in a better way. Recently, K-means clustering and RFM are used together greatly and provide substantial information about customers. In the study, customers were classified into four groups based on RFM values and a customer pyramid was created. The top of the customer pyramid belongs to the Platinum group which are the most profitable customers and then others in order Gold, Silver, and Lead groups. Based on this pyramid, the company has achieved to make decisions effectively by data science techniques.

Monalisa et al. (2019) highlighted the importance of customer lifetime value analysis by exploring beneficial customers for a company which is LWC. At the same time, transaction data was analyzed to cluster customers with the Fuzzy C-means algorithm via the RFM model. Then, the number of clusters was verified by the PCI

(Partition Coefficient Index) model, and AHP (Analytical Hierarchy Process) weights were multiplied with ranking to explore customer lifetime value. Related to that, the most valuable customers for the company were found out and customers were clustered as superstar customers, typical customers, and dormant customers.

The RFM model serves a purpose with its three critical aspects besides, these values could be used in some studies with other variables. Güçdemir (2013) has analyzed customer segmentation in an International TV Manufacturing Company and used recency, frequency, and monetary values with other characteristics which are loyalty, average annual demand, average annual sales revenue, long-term relationship potential, the average percentage change in annual demand, and average percentage change in annual sales revenue. In the study, data is normalized and the importance weights of criteria are specified using MCDM (Multi-Criteria Decision Making) technique. In addition to that, fuzzy AHP is used in the stage. The study contains two different approaches which are single dimension-based segmentation and multiple dimensions based segmentation. Single dimension-based segmentation uses an overall score of customers via a weighted average of the characteristics. Meanwhile, the second approach segments customers based on the similarities of each characteristic.

The research (Wu et al., 2014) has aggregated LRFM (length, recency, frequency, and monetary) model and cluster analysis in a pediatric dental clinic to examine customer values. Length represents the number of days between the first date visit and the last visit date. To cluster 1,462 patients into twelve groups, a two-stage approach by self-organizing maps and the K-means method was implemented. The results demonstrate that three clusters could be classified into loyal patients with L, R, and F values that are greater than the separately average values. On the other hand, three clusters could be seen as lost patients that have any variable above the average values of L, R, and F.

Hajiha et al. (2011) propose an expanded RFM model that is named RFMDC. RFMDC is a model to cluster customer values with additional two parameters Diversity and Continuousness. Diversity of the purchase represents the variety of products that meet customers' expectations to buy everything they need in a company. If the diversity is high, it means that customers are more loyal. Continuousness

indicates the largest following purchase which is continuous in a particular period. In the research, the proposed RFMDC model was implemented on Kalleh Dairy company which is the producer of ice cream and milk products in Iran. RFMDC model and K-means clustering were implemented for the customers of the Kalleh Dairy dataset. As a result of that, it demonstrates that the proposed model has more effective results than the segmentation of customer value via just using RFM variables and the K-means algorithm.

Companies should understand the customers' data well and know the differences and similarities among customers for good customer management. The research (Dogan et al., 2018) proposes two different clustering methods, which are K-means and two-step clustering, and RFM variables to segment 70.0032 customers. In the study, a retail sports store chain already had customer segmentation based on just the expense of customers. However, recent research has revealed that it is not sufficient to segment customers by their expenses. Therefore, the research suggests using the RFM model and it has proved that customer segmentation gives more useful information about customers by using more parameters.

The study (2010) done by Hosseini et al. gives a convenient methodology that can be used to examine customer loyalty in marketing strategies. Furthermore, it deeply looks over the studies related to loyalty, customer relationship management, the RFM model, and K-means clustering. Meantime, the main purpose of the study is to find out the loyalty degree of the product to figure out a great CRM that assists to maximize profits in the B2B concept. The proposed model was practiced by a supplier company that is named SAPCO Co. SAPCO Co is one of the market-leading supplier companies in Iran providing equipment, and car accessories for various automobile factories in Iran via some distributor agency companies like ISACO Co, which usually offer after-sales services to the customer. In the proposed model, R represents the supply percent of each product for ISACO, F means the frequency of the purchase of each product by ISACO, M symbolizes the monetary value of the order from the ISACO site, and L represents the period of product activity. Unlike other analyzes, two parameters which are the distance and the integrated rate of each cluster are used as a combination in the study's classification and clustering analysis. The result has

proved that the developed method for CRM is a convenient model with a high level of confidence when compared with other common models.

Consequently, all research demonstrates that RFM is a useful method and customer segmentation could be developed by adding parameters to the RFM model. Even, the results express that the developed model is more suitable than just using the RFM model (Wei et al., 2010).

3. CUSTOMER SEGMENTATION

The longevity of companies largely depends on their relationships with their customers. Customers have various characteristics thus their expectations and preferences differ from each other. In a competitive market, firms display customer-oriented behavior to catch an advantage (Güçdemir, 2013). Since companies have limited resources, they need to segment their customers to use these resources in the best way. Also, it is hard to develop marketing strategies that are compatible with all customers.

The segmentation concept was introduced in 1956 and showed that customers have common behavior in the diversified market (Chen et al., 2019). Customer segmentation has many advantages that help to concentrate each group with ad hoc services. It provides special offers to detected beneficial customers, which are evaluated based on their purchasing behaviors (Mousaeirad, 2020). Thereby, it leads to increase customer satisfaction and ensures customer retention.

Companies are aware of the importance of customer segmentation which is crucial to understand their customers and behave to their demands in a competitive market. Hence, they believe that customer segmentation contributes to separating the most valuable customers and don't need to make an effort for all customers. In addition, different marketing methods can be applied to various customer segments (Tsai et al., 2015).

Customer segmentation can be applied through various methods such as classification, clustering, etc. While there are different ways to implement customer segmentation, they all serve a common purpose that determining similarities and dissimilarities and grouping customers by using customer data. Furthermore, a variable for customer segmentation named RFM and clustering as a technique is heavily preferred (Güçdemir and Selim, 2015). Moreover, K-means clustering is seen as the most chosen one among clustering methods.

4. DATASETS

In this study, a carpet&rug company is selected as a sample to apply a developed RFM model. The company has collaborators in the domestic market and also abroad. In Turkey, the company has two types of wholesalers which are furniture and carpet. Furniture wholesalers mean that they have furniture stores and also sell carpet&rug. Hence, they have only one carpet&rug company to work together. On the other hand, carpet wholesalers represent that they have hybrid carpet stores with different brands. Thus, the company is not the only seller for these wholesalers. Moreover, these two types of wholesalers differ in the way they work. For instance, furniture wholesalers manage all their operational processes with SAP while carpet wholesalers use an order platform. Also, furniture stores have worked with the company longer than carpet stores and they are included in this study because of that there is enough data about them while carpet stores are the opposite.

The proposed study includes an analysis of two datasets, the sales dataset as transaction data and the survey dataset which is conducted to identify new variables in the study. Datasets belongs to a carpet&rug company. The company has a distribution system with the from company to wholesalers and to dealers and has more than 1500 sales points in Turkey. The company delivers all products to its dealers by cargo, that is, the products do not go to the wholesaler first and then to the dealer, they reach the dealer directly. However, during the ordering process, the dealer sends an order to the company as long as wholesaler approves its order. All transactions are kept through SAP.

The sales dataset contains the sales rug&carpet from the wholesaler to the dealer between 2021 and the 17th of April, 2022. The raw data is obtained from SAP that is a ERP (Enterprise Resource Planning) software used by the carpet&rug company. Meanwhile, all sales are made from the factory to the wholesaler, from the wholesaler to the dealer and from the dealer to the final consumer are recorded via SAP software. The dataset belongs to 976 dealers with 124.360 transactions, including invoice date, wholesales' names, dealers' names, city where the dealer is located, ordered products' name, total sales amount, total sales quantity, tax amount, VAT rate

and maturity period. Some variables are crucial in this datasets to create RFM values such as wholesalers' names, total sales amount, total sales quantity, and order date. On the other hand, the survey dataset was created after the survey was sent to all wholesalers and wholesalers shared the survey questions with their dealers. All questions except questions which are related to the personal information and sales point information were obtained using the 5-point Likert scale to merge results with sales datasets. A total of 48 questions were posed to the dealers, including questions to understand the general profile of the dealers and the participant. Also, specific questions are included to the survey about the variables determined as a result of an broad literature review. All questions are shared in Appendix A.

4.1 Description of the Datasets

The Carpet&Rug company has sales points at stores of well-known three furniture brands around the Turkey and these stores have SAP software to order products. When a dealer want to order a product, firstly sent the request to its wholesaler. Then, if the order is appropriate, it is sent to the from the wholesaler to the company. All orders made from the dealer to the wholesaler, from the wholesaler to the company are occurred at SAP software. Thus, sales datasets are obtained thanks to the reporting modules of SAP.

For the study, sales of the dealers between 2021 and the 17th of April, 2022 is extracted from SAP. The sales dataset contains some metrics such as invoice date, wholesalers' names, dealers' names, city where the dealer is located, ordered products' name, total sales amount, total sales quantity, tax amount, VAT rate and maturity period. In the dataset, some variables are more crucial like invoice date, wholesalers' names, dealers' names, total sales amount, total sales quantity which are determinants of the Recency, Frequency and Monetary values. In detail, "invoice date" determine the Recency value, "total sales quantity" help to constitute Frequency value and "total sales amount" is an indicator of Monetary value. The selected metrics in the dataset based on the method are explained in Table 1.

Table 1 : Selected Metrics of Sales Dataset

Metrics	Descriptions
Invoice Date	Indicates the date of purchase of dealers from wholesaler
Wholesalers' Name	Represents wholesaler of the dealer
Dealers' Name	Represents dealer name of the carpet&rug company
Total Sales Amount	Specifies the total sales amount for each transaction
Total Sales Quantity	Specifies the total sales quantity for each transaction

Survey datasets includes metrics which are Loyalty, Confidence in Brand, Trust, Dealer Dependence, Supplier Dependence, Product Expectation, and Service Expectation to create new developed RFM model. All metrics in the dataset are explained in Table 2.

Table 2 : Selected Metrics of Survey Dataset

Metrics	Descriptions
Loyalty	Represents how much the dealers are loyal to the carpet&rug company
Confidence in Brand	Represents how much the dealers have confidence to the carpet&rug company
Trust	Indicates how much the dealers trust to the carpet&rug company
Dealer Dependence	Indicates how much the dealers have dependence towards to the carpet&rug company
Supplier Dependence	Indicates how much the suppliers have dependence towards to dealers
Product Expectation	Shows how much the dealers have expectation about the product quality
Service Expectation	Shows how much the dealers have expectation about the service quality

4.2 Determining Variables For Developed RFM Model

Understanding customers and determining their interest in the sector is vital for organizations. Customers could be evaluated and clustered based on many different features. Thus, companies can improve special products or services for crucial customers easily. In this study, firstly seven different characteristics of customers are taken into account and evaluated in terms of similarities and dissimilarities. These scales are detected as a consequence of a broad literature review. After that, questions are assigned for each variable. Based on that, survey questions for dealers are created. However, cronbach's alpha test and factor analysis is applied to allocate the variable's underlying communality and reability. Thus, three variables which are loyalty, dependence and expectation are determined. The detail about factor analysis are referred at analysis part in the study.

In this section of this study, the importance of loyalty, confidence in brand, and trust of a customer and supplier and dealer dependence, product and service expectation of a customer is emphasized. After factor analysis, loyalty, confidence in brand and trust are mentinoned as "loyalty", supplier and dealer dependence as "dependence", product and service expectation as "expectation".

In the survey, questions related to dependence and trust are taken from thanks to Mutlu's (2007) study. Hereby, the study of Celly and Frazier (1996), Kumar and his friends'(1995) and Ganesan (1994) set light to the survey questions. In addition to that, questions about loyalty and confidence in brand are created from the articles (Knox, 1998; Hultén, 2007; Malik et al., 2013; Srivastava and Kaul, 2016; Driggs et al., 2006; Mao, 2010; Laroche et al., 1996; Sheth & Howard, 1989) describing these two concepts. Lastly, studies (Wu and Chou, 2011; Choudhuri & Parida, 2014; William et al., 2016) shaped expectation quesitons. Also, they are clarified in consultation with the managers of the company about whether the related questions measure the product and service expectation.

4.2.1 Loyalty

Customer loyalty is crucial just because it encourages existing customers to spend more thereby endorsing profitable growth. If a customer spends more on a

particular brand despite having an alike product in other brands over time, he or she is more closely associated with the brand. These preferences hold up that brand loyalty exists.

A big part of markets can be segmented using this simple proposition, and customers could be classified according to purchasing characteristics, motivations, and strategies by levels of loyalty (Knox, 1998). Many companies can get the edge by finding out which customers are the 'best 'and distinguishing them in terms of profitability and loyalty (Hultén, 2007). Best customers are critical for the future of the company while requiring less time and effort spent.

Loyalty becomes so important for organizations to get optimum price and hike up sales volume additionally customer retention is much easier than looking for new customers (Malik et al., 2013). Enhancing loyal customers is mandatory for achieving high performance. Studies emphasize the fact that it is much more expensive to take new customers than to retain them. On the other hand, loyal customers contribute financially by making lots of purchases.

Especially in the retailing sector, many companies are aware that they can boost customer loyalty and performance via a better understanding of customers. On the other hand, repeat purchasing behavior could be seen as an indicator to be loyal to the brand (Srivastava and Kaul, 2016). If the customer repeats the purchase under any circumstances, it is proof of the customer's love for the brand.

In a survey of a service company, 81% of customers indicated they will buy from the companies that they are becoming loyal to, and nearly half of them said they will purchase more or converge to specials from an alike company. In a nutshell, customer loyalty finds out pricing strength in the competitive market and also provides superior protection against threats (Driggs et al., 2006). Just because of this, the more loyal customers a firm has, the more it has power in the market.

Loyal customers ensure more trustworthy and comfy working conditions for the company in many respects. One of them is that they are less price-sensitive to the purchase of a product. Related to that, when a product's price is increased, they don't reduce the quantity of buying because they believe that it is the right price of the

product (Mao, 2010). Thus, they show how they believe in the company's decisions. Correspondingly, they intend to focus less on the price.

4.2.2 Confidence In The Brand

The confidence term was first presented as one of the motivations for purchase intentions by Sheth and Howard. Confidence is the degree to which a person is confident that their general judgment about a brand is correct (Laroche et al., 1996).

Sheth and Howard (1989) claimed that confidence has a significant influence on the intention of customers. Concordantly, Bennett and Harrell commit that confidence in the brand is a key aspect to forecast buying intentions. Research, organized by Laroche, Kim, and Zhou in 1996, has also proved that confidence in brand evaluation is one of the determinants of purchase intention. All these evaluations show that confidence in the brand is an important criterion in understanding the customer and examining the purchasing behavior.

4.2.3 Trust

Trust term generally has been defined from a three-dimensional construct that consists of helpfulness, competence, and integrity (Gefen et al., 2003; Pavlou and Fygenson, 2006). Meanwhile, helpfulness is a belief that is cared for from the other side and encouraged to treat in one's interest. And, competence is a belief that someone or something has the power and qualification to do what needs to be done and integrity is a belief in someone or something that acts in good faith. Related to that, integrity entities declare the valid and be aware of the treating ethically. Thus, these entities most likely keep their promises.

Trust is a key variable in understanding and predicting consumer behavior. On the other hand, it is mostly not used as a customer segmentation criterion. Trust or trust-related variables used in a segmentation framework that includes empirical work are extremely scarce. A study has used trust as a variable to segment potential customers in two self-service bank channels (Dimitriadis et al., 2011). As a result of

that, the research has proved a remarkable straight and correlator impact of trust on purchase intention.

On the other hand, research by Yilmaz et al. (2004) gives crucial information on whether and when trust may take place as a key factor in long-term channel dyads. It finds out that the circumstances under which trust-based efforts will succeed in targeted channel member relations are specified via perceived interdependence.

4.2.4 Dependence

In inter-firm relations, dependency is defined as a firm's need to maintain a business relationship with its partner to achieve its goals (Kumar et al., 1995). Generally, the relationship exists as long as the mutual interests of the firms proceed.

According to Emerson (1962), the dependence of a firm on another firm with which it does business is determined by two major factors: (1) the total economic value obtained from the other party (economic importance), and (2) the difficulty in substituting the counterparty and the economic value derived from it (substitutability). As the firm's dependence on the dealer increases, its power, and control ability will decrease. Besides, the reverse of the same situation is valid for the dealers.

When a firm intends to end the business relationship with the current partner, it is necessary to find alternative partners instead. If an alternative is not available to both parties, the higher the degree of interdependence occurs. In a nutshell, substitutability is one of the significant variables affecting the degree of dependency (Celli and Frazier, 1996). Consequently, dependency increases as the substitutability of the partner decrease.

4.2.5 Product Expectation

When a person decides to buy a product, he/she does not only look out for the price of the product but also evaluates the product in terms of many properties. According to Wu and Chou (2011), a consensus among the customer and the company

about the value of the product and good customer service is also important in terms of buying behavior. Related to that, companies should not produce a monotype product because it will not mostly please all customers. In other words, companies should segment their customers to understand the purchasing behavior of customers and present unique products for each segment. Thus, it develops an accomplished marketing strategy and gains a competitive advantage in the market.

4.2.6 Service Expectation

Service quality plays a fundamental role in the business strategy of the companies in the delivery process, improvement of productivity, profitability, and customer satisfaction. Meanwhile, there is an expectation of the service by the customers in line with the parallel of brand identity. In addition, service quality directly influences the repurchase behavior of customers (Choudhuri and Parida, 2014).

William et al. (2016) indicate that providing quality service is observed as a striking factor in a competitive marketplace and has incrementally become a substantial element of business strategy.

On the other hand, companies are now endeavoring to use information technology (IT) and information technology-enabled services (ITES) to acquire a competitive advantage against their competitors (Choudhuri, Parida, 2014). Therefore, companies that offer information technology-supported services are one step ahead now.

5. METHODOLOGY

The proposed research methodology consists of five main steps. First, data cleaning is applied to a survey dataset and a sales dataset. Second, a developed RFM Model is created. RFM values come from the sales dataset and other variables, loyalty, expectation and dependence, are obtained from the survey dataset after factor analysis. Then, the optimum cluster number is determined via the Elbow Method. After the finding best cluster number as 6, data is analyzed by using K-means clustering. Finally, the results are explained by anova and regression analysis.

Table 3: The Steps of the Proposed Methodology

Data Cleaning	→	Creating a Developed RFM Model	→	Factor Analysis	→	K-Means Clustering	→	Anova and Regression Analysis

In the research, primary and secondary data are used. The datasets belong to one of the well-known rug&carpet companies in Turkey. The company produces and offers carpets & rugs to dealers in Turkey and B2B customers abroad. The secondary dataset covers the sales of the dealers from 2021 to the 17th of April, 2022.

On the other hand, a survey is shared with 976 dealers of the company and there are responses from 152 dealers. It is determined that the responses of 106 dealers would be used after data cleaning. The 5-point Likert scale is used in the survey because very complex scale-formatted questions largely go unanswered or result in an "I don't know" answer (Leigh and Martin, 1987). Evaluation options for the statements in the questionnaire that measure the concepts of loyalty, confidence in the brand, quality expectation, service expectation, trust, and dependence are as follows: 1-Strongly Disagree, 2- Disagree, 3-Hesitant, 4-Agree, 5-Strongly Agree.

After that, the sales dataset is arranged according to the survey results. Thereby, a scoring system from 1 to 5 is obtained for each variable. Based on the values below table is created as a summary.

25

Table 4: Descriptive Statistics of Variables

	Recency	Frequency	Monetary	Loyalty	Dependence	Expectation
Mean	2.98	2.98	2.98	4.11	3.35	4.66
Min	1.00	1.00	1.00	1.50	1.43	2.44
Max	5.00	5.00	5.00	5.00	5.00	5.00
St. D.	1.43	1.43	1.43	0.70	0.88	0.48

5.1 RFM

RFM is a model that is based on customer behavior consisting of recency, frequency, and monetary variables and promotes predictions on behavior (Wei et al., 2010). Meantime, recency represents the customer's last purchase date; frequency denotes how frequently customers purchase in a given period, and monetary is the amount of money that customers spend in a given period. RFM model has been preferred in many fields such as government agencies, banking, insurance agencies, telecommunication industries, marketing industries, etc. The RFM model contributes to finding out organizations' top customers that have high RFM scores. Thus, a unique strategy can be applied to this segment.

On the other hand, businesses can use RFM analysis to cluster customers who have similar values instead of using RFM values only to identify top customers. Thereby, they can evaluate each segment individually and implement a specific strategy for each segment. Furthermore, the model can be used to calculate customer value and determine customer lifetime value (CLV).

RFM is a useful and cost-effective model that analyzes customer behavior easily. Also, it is easy to understand and interpret the results. In the model, all customers have a score between 1 and 5 for the variables of recency, frequency, and monetary value based on their purchase history in the customer database. Hence, 5 is the highest score and 1 is the lowest score for each variable. Once scores are revealed for each variable, the three scores are written side by side, and a score is assigned for

each customer. Example: 555, 545, 444, etc. Thus, 555 represents the most important and highly profitable customers, while 111 is the opposite.

Some papers mention that there is an attempt to develop the RFM model by adding some other variables to provide better predictions of customer behavior (Wei et al., 2010). For instance, Yeh et al. (2009) developed a new model by extending the RFM model. The new model is named RFMTC and has two new variables which are time since first purchase and churn probability. In addition, Buckinx and Van (2005) got a retailer dataset and observed past purchase behavior. They added other customer variables to the analysis including RFM values. Hence, the findings showed that past purchase behavior variables, especially RFM variables are substantial predictors of customer behavior. In addition to that, Coussement and Van (2009) analyzed call center e-mails to find whether the extended RFM model (eRFM) or eRFM-EMO which is a model that adds emotionality to variables is more accurate to predict customer behavior. As a result of that, it has been proven that the eRFM-EMO model gives better results than the eRFM model. Based on these papers, in this study, some other variables are included in customer segmentation analysis via a survey.

5.2 K-Means Clustering

Clustering is a tool for exploratory data mining in various applications such as classification and machine learning fields (Inaba et al., 1994). K-Means is one of the clustering algorithms that categorize a particular dataset thanks to a certain number of clusters called k (Kodinariya & Makwana, 2013).

Table 5: Types of Clustering Methods

```
                    ┌─────────────────────┐
                    │  Clustering Methods │
                    └─────────────────────┘
                               │
                 ┌─────────────┴─────────────┐
         ┌───────────────┐          ┌──────────────────┐
         │  Hierarchical │          │ Non-hierarchical │
         └───────────────┘          └──────────────────┘
            ↙         ↘                       │
    Agglomerative    Divisive              K-Means
          │
          ↓
```
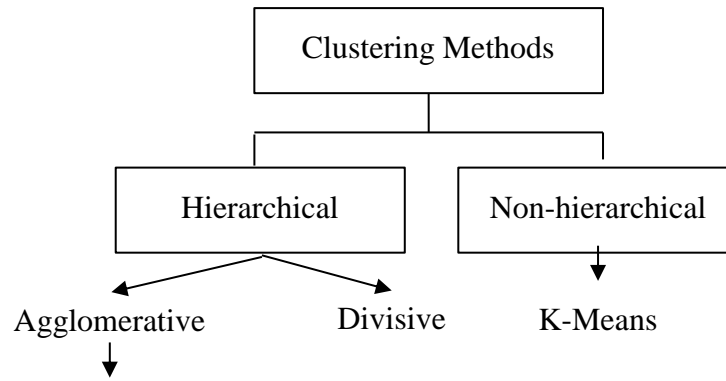
- Single Linkage
- Complete Linkage
- Average Linkage

In K-Means clustering, the main issue is determining the k centroids of each cluster. Then is to calculate each point of a given data set and correlate it with the nearest centroid. If there is no point pending, the first step is finished and an early group is created. In the meantime, k new centroids are re-calculated. After that, a new linking has occurred between the same data points and the nearest new centroid. Thus, a cycle has been set off. Therefore, k centroids change until centroids do not move from their location anymore.

Table 6: The steps of K-Means Clustering

```
                          ┌──────────────┐
                          │    Start     │
                          └──────┬───────┘
                                 │
                                 ▼
          ┌────────────────────────────────────────────┐
          │         Select the Number of Clusters       │
          └───────────────────┬────────────────────────┘
                              │              ◄──────────────┐
                              ▼                             │
          ┌────────────────────────────────────────────┐   │
          │        Set Initial Cluster Center Randomly   │   │
          └───────────────────┬────────────────────────┘   │
                              ▼                             │
          ┌────────────────────────────────────────────┐   │
          │         Put Objects to the Closest Center    │   │
          └───────────────────┬────────────────────────┘   │
                              ▼                             │
          ┌────────────────────────────────────────────┐   │
          │       Recalculate the New Cluster Centers    │   │
          └───────────────────┬────────────────────────┘   │
                              ▼                             │
          ┌────────────────────────────────────────────┐   │
          │    Create Clusters Based on Smallest Distance │   │
          └───────────────────┬────────────────────────┘   │
                              ▼                             │
          ┌────────────────────────────────────────────┐   │
          │            Objects Move to Clusters          ├───┘
          └───────────────────┬────────────────────────┘
                              ▼
          ┌────────────────────────────────────────────┐
          │                  Output                      │
          └────────────────────────────────────────────┘
```
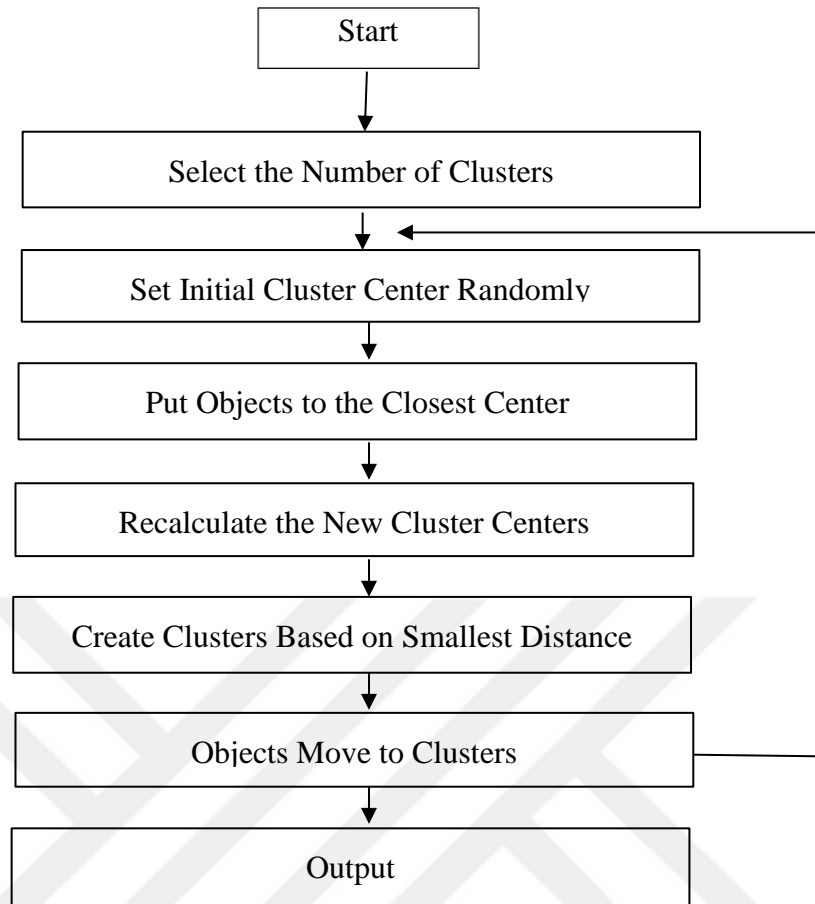
In the k-means clustering, a given set of observations (x1, x2, x3, x4,…, xn) that each observation (xi) indicated a d-dimensional real vector. The purpose of the k-means clustering algorithm is to divide n observations into k observation groups, called clusters with $k \leq n$, and also to minimize the sum of the squares of the distances among the observations in a distinctive cluster (Vattani, 2011). Euclidean and Manhattan distance measures can be used to appoint each record to the nearest cluster. In K-means clustering, the euclidean distance measure is the preferred one.

Equation 1: Euclidean distance measure

Euclidean distance measure: $d = \sqrt{(a1 - b1)2 + (a2 - b2)2 + \cdots + (an - bn)2}$

On the other hand, there is the difficulty of choosing the number of clusters called k. Related to that, there are several studies to determine the right k such as by rule of thumb, information criterion approach, an information-theoretic approach, elbow

method, choosing k using the silhouette, and cross-validation. The Elbow method is the oldest one to indicate the exact number of clusters in a data set (Kodinari and Makwana, 2013). The method starts with determining k=2 and increases the k by 1 in each step. Meanwhile, the clusters and the cost that is because of the training are calculated. Where the cost at some value for K decreases sharply, and then, it reaches a flat while K is increasing, it results in the K value wanted.

Graph 1: Find to Optimal Number of Clusters(k) followed by Elbow Point
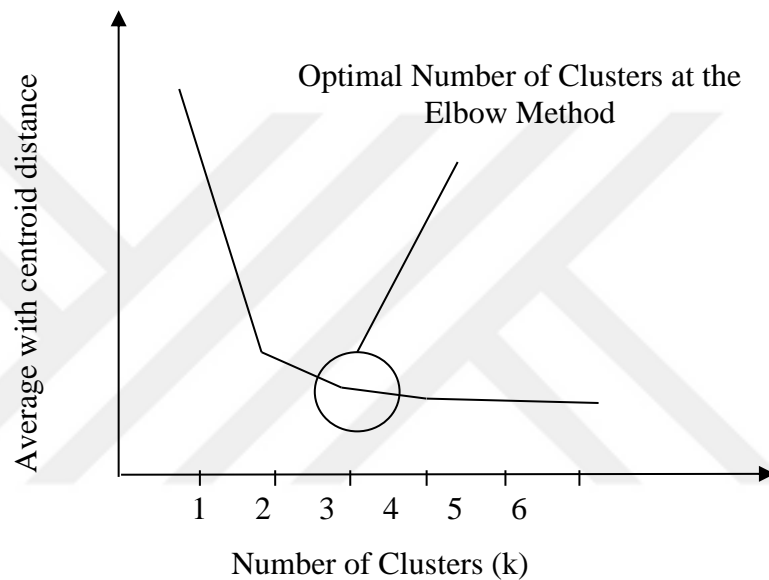
6. ANALYSIS

6.1 Customer Profile

Apart from the questions about the scales, the participants who participated in the survey were asked questions such as gender, age, graduation status, position at the dealer, working time at the dealer, experience in the carpet industry, and how many years he has been working with the company. Of the total 106 respondents, 67 are male and 39 are female. The respondents were between the ages of 21 and 56, with an average age of 37. Furthermore, their seniority duration is between 6 months and 32 years with an average of 11 years. Looking at the educational status of the participants, 2 of them have primary education, 44 people with a high school, 19 people with college, 33 people with undergraduate, and 8 people with postgraduate education. Besides, 35 of the respondents are business owners, 15 of them are middle managers and 32 of them are top managers. This information about the respondents is shown in Table 7.

Table 7: Descriptive Statistics on Survey Respondents

		Numeric Values	Percent Values
		(number of people)	
Gender	Woman	39	37%
	Man	67	63%
Age		(year)	
	Minimum	21	
	Maximum	56	
	Mean	37	
	Standard deviation	8,7	
Seniority Duration		(year)	
	Minimum	0,5	
	Maximum	32	
	Mean	11	
	Standard deviation	8,8	
Education Status		(number of people)	
	Primary Education	2	2%
	High School	44	42%
	College	19	18%
	Undergraduate	33	31%
	Postgraduate	8	8%
Position		(number of people)	
	Business Owner	35	33%

		Middle Manager	15	14%
		Top Manager	32	30%
		Other	24	23%
		TOTAL	106	

On the other hand, in the survey, how long they have been working in the carpet industry and how long they have been working with the current manufacturer were asked to the dealers. The responses are summarized in the table below.

Table 8: Descriptive Statistics on Survey Dealer

		Numeric Values
		(year)
Industry Experience	Minimum	0,5
	Maximum	35
	Mean	10,5
	Standard deviation	8,1
Relationship Duration		(year)
	Minimum	0,7
	Maximum	25
	Mean	5,6
	Standard deviation	5
	Total	106

Participants of the survey has industry experience from six months to 420 months and working time with the company varies between eight to 300 months.

Graph 2: Dealers' Industry Experience and Working Time With the Company
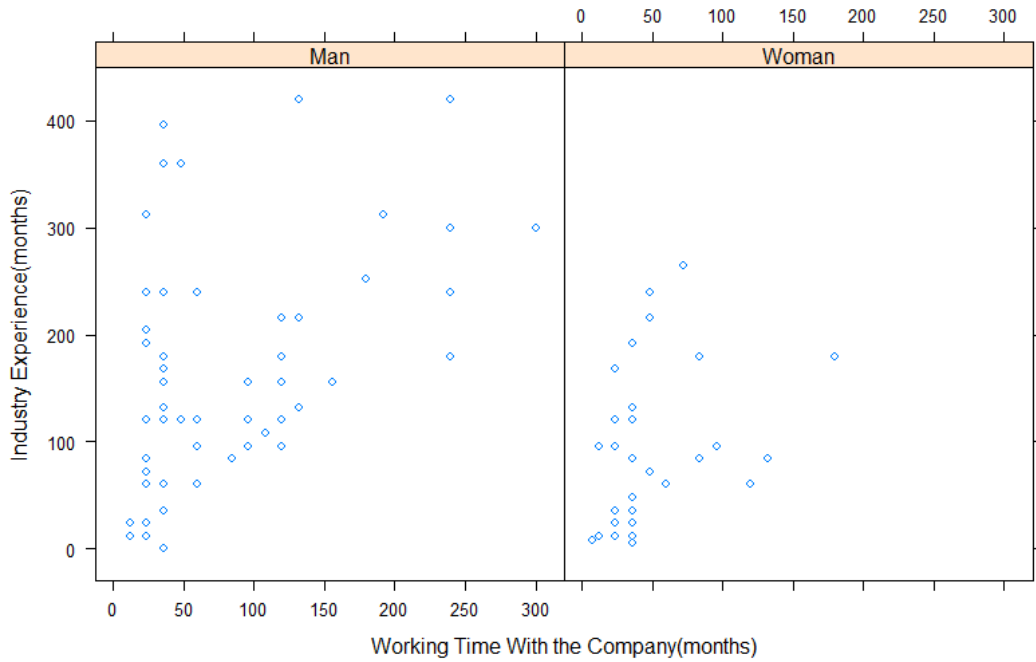
The survey results give a perspective of the tendencies in the loyalty, dependency and expectations of the dealers. Thereby, these variables changes according to the level of education.

Graph 3: Loyalty values based on Education Status
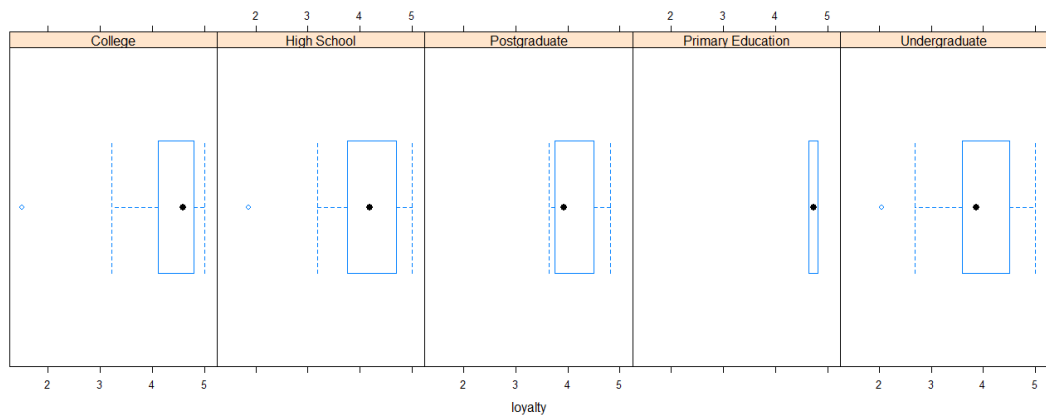
Graph 4: Dependence values based on Education Status
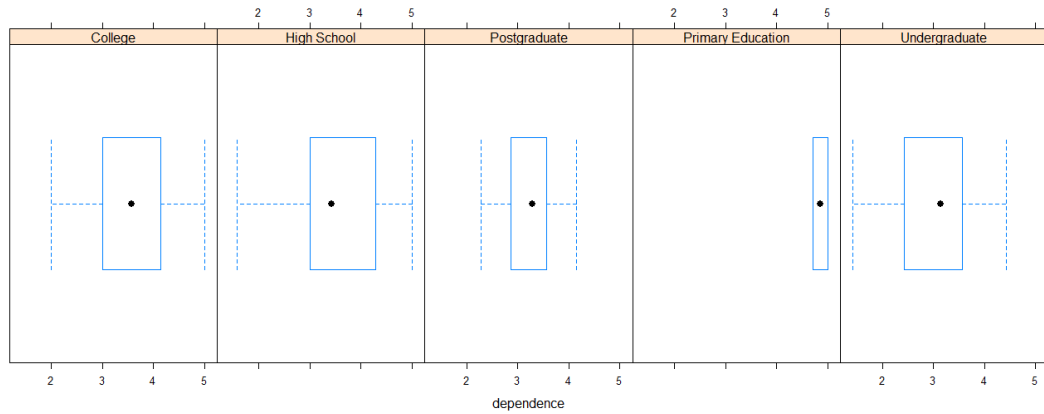
Graph 5: Expectation values based on Education Status
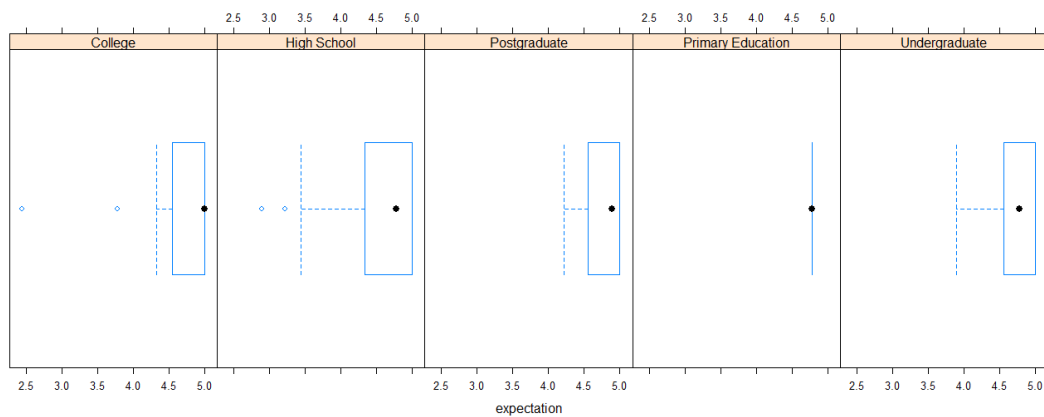
6.2 Reliability

The sudy aims to present a developed model by adding extra variables to RFM model. Hereby, adding variables are determined by survey questions. In the meantime, survey questions are composed via a detailed literature review. First, important criterias for customer segmemtation are determined and then questions are created based on each topic. Absolutely, questions need to have internal consistency to generate a scale. Therefore, the questions which evaluate the same variable, should be in the same questionarre group. As a result of that, Cronbach's alpha is used to test internal consistency.

All three factors has high a= 0.96 and also when we remove an item, it is not changes so much. In addition, values seem as having a high reliability that all have value more than 0.85. The detailed results are shown in the table 9.

Table 9: The Results of Alpha Test

	Cronbach's Alpha	Cronbach's Alpha Based on Standardized Items	Number of Items
Loyalty	0.96	0.96	22
Dependence	0.86	0.86	7
Expectation	0.83	0.86	9

6.3 Factor Analysis

Factor analysis are statistical methods used for examining the relationships between a group of observed variables which are determined through items or questions and finding the underlying variables called factors to make it more understandable. (Beavers, et al.,2013) In this study, 40 questions associated with seven variables are created after a broad literature review. However, there is a need to test whether this relationship between the questions is appropriate. Therefore, an exploratory factor analysis apply the study is planned.

Firstly, survey questions are subjected to proficiency tests for factor analysis. Based on the Kaiser-Meyer-Olkin and Bartlett's Test of Sphericity, questions are indicated that they are suitable for factor analysis. According to Kaiser-Meyer-Olkin factor adequacy, MSA is found as 0.89 that means it can be conducted a factor analysis. In addition to that, p value in Bartlett's Test of Sphericity is emerged as 0 that also support that a factor analysis can be useful for the study because of being $p<0.0.5$.

In the survey, some questions have common content or meaning in terms of containing correlated variable and explatory factor analysis allows to specify the underlying variables of these content. Determining the number of factors is one of the issues to be considered. Therefore, a parallel analysis is applied to the survey questions

and then parallel analysis suggests that the number of factors as three. Related to that, factor analysis is carried out by considering factor as three.

Graph 6: Parallel Analysis Scree Plots

Parallel Analysis Scree Plots
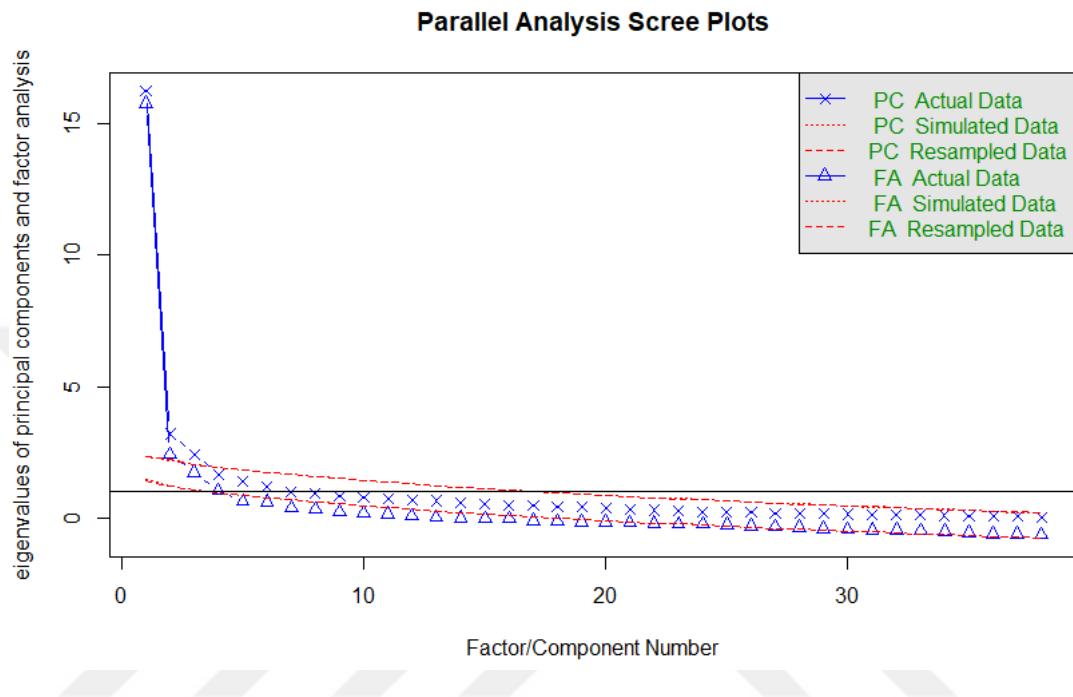
In the survey dataset, each question are numbered based on related variable. For instance, the first question is indicated as "loyalty". After that, an exploratory factor analysis is applied to the survey questions with oblimin rotation. As a result of that, the below table is created and each loading is assigned to a factor.

Table 10: The Results of the Survey Factor Analysis

	Questions	Factor1	Factor2	Factor3
1	We plan to continue with this brand for 5 years.	0.49		0.33
2	The manufacturer produces highly qualified products at the right price.	0.72		
3	When a product is marked up, we don't lessen the quantity of buying because we believe that it is the right price of the product.	0.51		

4	When the manufacturer makes a mistake, we do not immediately break the collaboration because we believe that the problem will be resolved sensibly.	0.65		
5	We can safely recommend the manufacturer to anyone who considers working with the manufacturer.	0.71		
6	We are satisfied with the service that has been received from the manufacturer.	0.77		
7	We are satisfied with the products that have been produced by the manufacturer.	0.74		
8	We often visit the website and review the products or interact with the social media channels of the manufacturer.	0.57		
9	Even when our manufacturer makes an unbelievable statement, we believe that it is true.	0.54		
10	Our manufacturer generally keeps its promises to our company.	0.85		
11	We know that the manufacturer provides the best advice regarding our business.	0.74		
12	We trust the sincerity of our manufacturer company.	0.90		
13	Our manufacturer company often gives us true information.	0.60		
14	When our manufacturer company makes important decisions, it considers our well-being.	0.74		
15	When we share our problems with our manufacturer, we think that they will understand us.	0.71		
16	We trust our manufacturer will consider how its actions will affect us when making a new decision.	0.78		
17	The manufacturer rewards its loyal customers.	0.59		

#	Statement			
18	The manufacturer has an absolutely honest understanding.	0.80		
19	We have full trust in the manufacturer company in every respect.	0.86		
20	We trust it to do the right thing.	0.86		
21	We trust the manufacturer to do the job right.	0.94		
22	The manufacturer is always true-hearted.	0.86		
23	Our success in the business depends heavily on the marketing efforts of the manufacturer.			0.39
24	Despite all efforts, our sales would drop significantly if we no longer represent this manufacturer.			0.68
25	We can't easily find an alternative manufacturer to replace the manufacturer.			0.52
26	The manufacturer is in a key position that affects our overall business performance.			0.66
27	If we no longer represent this manufacturer, its sales in our region will drop significantly.			0.75
28	We are a key selling point in our trading area for the supplier's products.			0.69
29	We create high sales volume for this supplier.			0.61
30	The manufacturer should have the best carpet product groups in the market.		0.49	0.32
31	The manufacturer should have a wide product range in the market.		0.50	0.32
32	The manufacturer should offer its products to the market at competitive prices.		0.54	
33	Innovative products attract the attention of consumers and they should be in the manufacturer company.		0.49	
34	Fast delivery to the customer is very important for the quality of service.		0.86	

35	After-sales services should be provided directly by the manufacturer.		0.76	
36	The manufacturer has to provide broad information about the products.		0.58	
37	It is sufficient for the manufacturer to keep stock itself to serve the consumer.		0.45	
38	It is very crucial to take advantage of the information technology-supported services provided by the manufacturer.		0.66	

Thus, total 50% cumulative variance is explained by the 3 factors. Shortly,

- Factor 1 has high loadings for Loyalty from 1st to 22th questions.
- Factor 2 has high loadings for Expectation from 30th to 38th questions.
- Factor 3 has high loadings for Dependence from 23rd to 29th questions.

On the other hand, there are two questions in the survey which are excluded because they are not belong to corrolated variables. The questions are;

- "The manufacturer's products meet customers' expectations."
- "The supplier can easily switch to another dealer in our trading area."

6.4 RFM

The sales datasets has information to compose Recency, Frequency and Monetary values of each dealer. Related to that, these values are calculated based on the sales transactions. First, the last purchase date of the each dealer is found. As a result of that, RFM scores are assigned via RFM package in R. The main process behind the assignment of these scores is that quartiles are created based on the total scores and each quartile represent a number from 1 to 5. Meanwhile, 5 indicates the most recent date while 1 is the most distant. Similarly to that, Frequency and Monetary scores are calculated by using formulas in R. Frequency is found by summing up how often the dealer purchase. According to the quartiles created, 5 points are assigned to the most frequent purchaser and 1 point to the least purchaser. Lastly, monetary representing the total amount of money purchased by a dealer. If a dealer is in the

quartile which has higher purchase amount, then it gets 5 point. All these values create the RFM values for each dealer.

6.5 K-Means Clustering

To generate the proposed model, the sales dataset has been converted into an RFM model and then each customer held RFM scores from 1 to 5. Also, a score between 1 and 5 is set off by taking the average of the questions for each variable related to each dealer in the survey dataset. Hence, a score is obtained for each dealer for the ten variables described in Table 11.

Table 11: Variables of the Proposed Method

Variable	Description
R	Recency
F	Frequency
M	Monetary
L	Loyalty
D	Dependence
E	Expectation

After that, we centered and scaled the columns of a numeric matrix by using the scale () function in R. Then, we used to estimate the optimal number of clusters via the Elbow Method. In the method, you begin with K=2 and gradually increase it by 1 in each step, calculating the clusters and the training cost. The cost drops sharply at a certain number for K, then plateaus when you increase it more. Finally, it is the K value you're looking for. In the study, k is determined as 6 thanks to the Elbow method and the results are below in Graph 7.

Graph 7: Elbow Point of the Proposed Method

Optimal number of clusters
Elbow Method
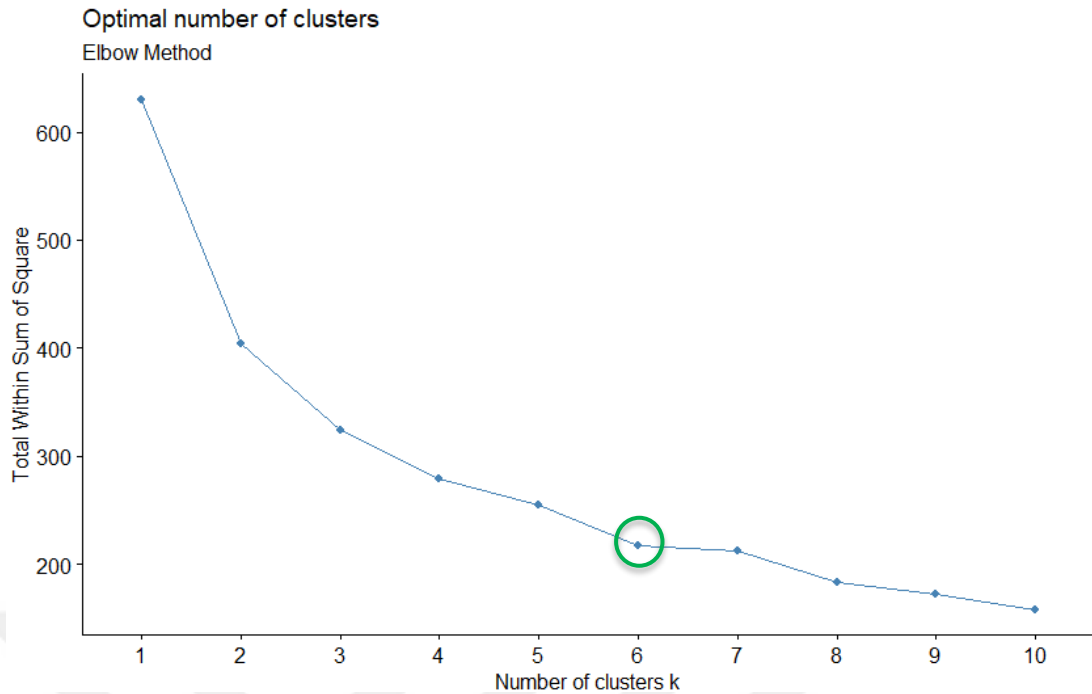
According to the determined number of clusters, each cluster has a group of dealers. The number of each cluster and the within-cluster sum of squares are shown below. Hence, the ratio of the total sum of squares and the between-cluster sum of squares has resulted in 66.5%.

Table 12: Within Clusters Sum of Squares of Clusters

Cluster	Number of Dealers	WCSS
1	30	47,8
2	12	27
3	3	9,1
4	11	28,2
5	28	60,5
6	22	38,3

To visualize clusters, the K-means cluster plot was created via an R programming language. All clusters are shown in the graph below.

Figure 1: Cluster Plot of the Proposed Method

Cluster plot
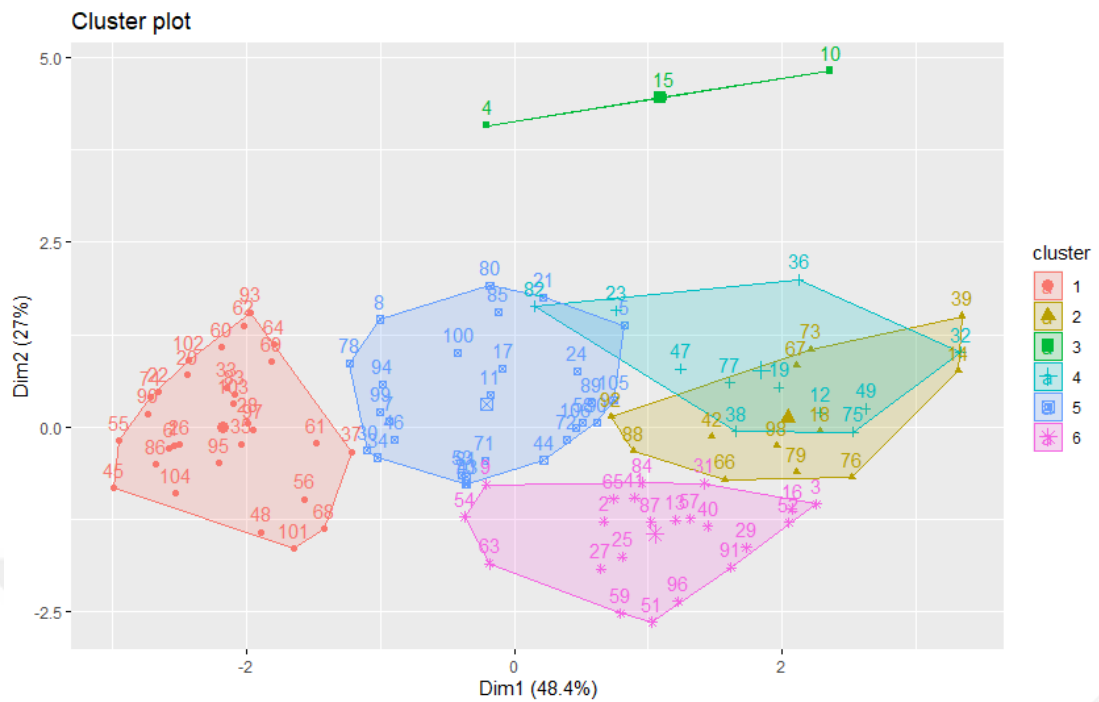
After the clusters are determined, each cluster is examined. In summary, the mean of each cluster is indicated in Table 13.

Table 13: Mean Values for each Cluster

Cluster/ Variables	Recency	Frequency	Monetary	Loyalty	Dependence	Expectation
1	4.37	4.6	4.53	4.52	4.16	4.86
2	2.25	1.42	1.42	3.44	2.20	4.82
3	4	3.67	4	2.18	2.33	2.93
4	1.91	2	1.82	3.67	2.91	3.86
5	3.25	3.29	3.39	4.07	2.96	4.75
6	1.55	1.64	1.64	4.46	3.73	4.81

Table 14: Standard Deviation for each Cluster

Cluster/ Variables	Recency	Frequency	Monetary	Loyalty	Dependence	Expectation

1	0.81	0.62	0.68	0.41	0.54	0.21
2	1.36	0.52	0.52	0.60	0.55	0.20
3	1	1.15	1	0.89	0.44	0.50
4	0.83	1	0.87	0.53	0.72	0.29
5	1.04	0.76	0.74	0.55	0.47	0.24
6	0.67	0.66	0.66	0.44	0.66	0.22

Table 15: Mean Values of Variables

Variables	Mean
R	2.98
F	2.98
M	2.98
L	4.11
D	3.35
E	4.65

Each cluster shows differences in terms of purchasing habits and behavior toward the company. In this case, the company can identify specific strategies for each group and detect the focus point of dealers. Naturally, applying the same strategy to dealers with similar characteristics while developing strategies provide great convenience to the company.

Group 1 contains the "most valuable customers" for the company. Because it has high Frequency, Monetary, Recency, Loyalty and Dependence score. In addition, this group has high expectation from the company in terms of product and service quality. Group 2 and Group 4 have similar characterstics and they can be named ad "loss customers". On the other hand, Group 3 and Group 5 are in critical postition and can be determined as "critical customers". Because their sales volume is high but they have low loyalty and dependence values. Therefore, these dealers may leave the business association with the company. Finally, Group 6 is a "focus group" that has

low sales values but high Loyalty, Dependence and Expectation values. It is really remarkable to examine. Hence, group details for the clusters are referred at Table 16.

Table 16: Description of Clusters

Segments	Interpretations
Group 1	
$F_1 > F$	Purchases were more frequent than in other dealers.
$M_1 > M$	Dealers that make a larger amount of purchases than other dealers.
$R_1 > R$	Purchases were made at a more recent time than in other dealers.
$L_1 > L$	Dealers that are more loyal to the company than other dealers.
$D_1 > D$	Dealers and supplier have high dependency on each other.
$E_1 > E$	Dealers that in high expectations of the service.
Group 2	
$F_2 < F$	Purchases were less frequent than in other dealers.
$M_2 < M$	Dealers that make a less amount of purchases than other dealers.
$R_2 < R$	Purchases were made at a more far time than in other dealers.
$L_2 < L$	Dealers that are less loyal to the company than other dealers.
$D_2 < D$	Dealers and supplier have low dependency on each other.
$E_2 > E$	Dealers that in high expectations of the service.
Group 3	
$F_3 > F$	Purchases were more frequent than in other dealers.
$M_3 > M$	Dealers that make a larger amount of purchases than other dealers.
$R_3 > R$	Purchases were made at a more recent time than in other dealers.
$L_3 < L$	Dealers that are less loyal to the company than other dealers.
$D_3 < D$	Dealers and supplier have low dependency on each other.
$E_3 < E$	Dealers that in low expectations of the service.
Group 4	
$F_4 < F$	Purchases were less frequent than in other dealers.

$M_4<M$	Dealers that make a less amount of purchases than other dealers.
$R_4<R$	Purchases were made at a more far time than in other dealers.
$L_4<L$	Dealers that are less loyal to the company than other dealers.
$D_4<D$	Dealers and supplier have low dependency on each other.
$E_4<E$	Dealers that in low expectations of the service.

Group 5

$F_5>F$	Purchases were more frequent than in other dealers.
$M_5>M$	Dealers that make a larger amount of purchases than other dealers.
$R_5>R$	Purchases were made at a more recent time than in other dealers.
$L_5<L$	Dealers that are less loyal to the company than other dealers.
$D_5<D$	Dealers and supplier have low dependency on each other.
$E_5>E$	Dealers that in high expectations of the service.

Group 6

$F_6<F$	Purchases were less frequent than in other dealers.
$M_6<M$	Dealers that make a less amount of purchases than other dealers.
$R_6<R$	Purchases were made at a more far time than in other dealers.
$L_6>L$	Dealers that are more loyal to the company than other dealers.
$D_6>D$	Dealers and supplier have high dependency on each other.
$E_6>E$	Dealers that in high expectations of the service.

6.6 ANOVA

The company has achieved six groups that have the minimum within-cluster sum of squares and a maximum between the sum of squares. After that, in the study, each variable has evaluated based on the analysis of the variance test and the results shown below. First, covariance among variables is checked, and seen that there is a positive relationship between loyalty and dependence. In addition, loyalty has a positive relationship with expectation. Therefore, if a dealer has high loyalty towards the company, it has more expectation and dependence. On the other hand, dependence seems the variable that contributes to sales mostly. The details are in Figure 2.

Figure 2: Correlation between Variables
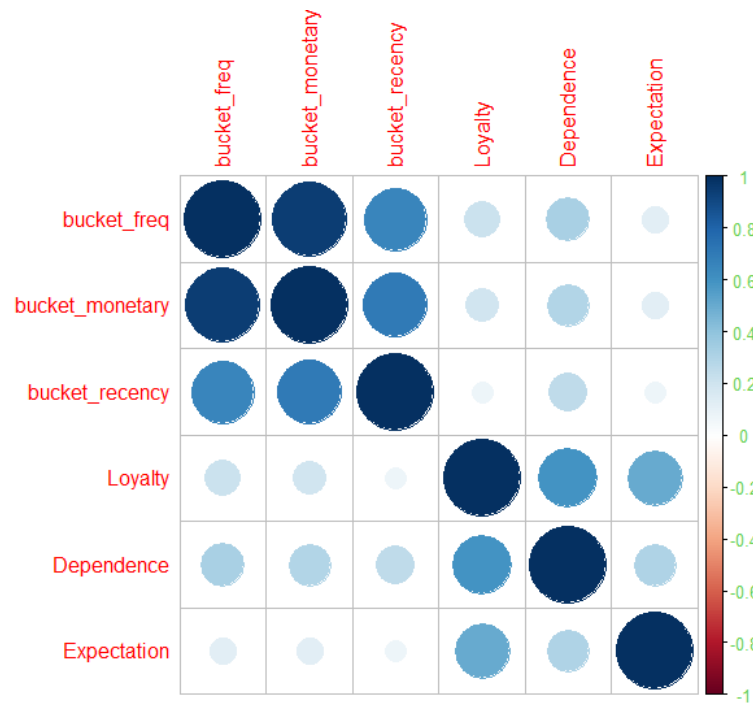
In ANOVA, the null hypothesis (H$_0$) represents that there is no difference in means, and the alternative hypothesis (H$_A$) is that the means are different from one another. If the p-value in our ANOVA table is less than .05 we rejected the null hypothesis. In this study, each variable is included in the ANOVA on a cluster basis. Based on ANOVA analysis differences between means are explained via Tamhane Test.

First, Levene Test is applied to the analysis. However, the results show that all p value is higher than 0.05 that the variance among the three groups are equal. Therefore, Tamhane technique is used because the variances are not homogeneous.

Table 17: ANOVA Test for Frequency

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
Frequıency	5	162.37	32.47	62.95	<2e-16
Residuals	100	51.59	0.52		

Interpretation: P-value (<2e-16) is less than .05, we have sufficient evidence to reject the null hypothesis. Thus, there is a difference in the means of frequency values between clusters.

Table 18: ANOVA Test for Monetary

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
Monetary	5	164.17	32.83	65.95	<2e-16
Residuals	100	49.79	0.50		

Interpretation: P-value (<2e-16) is less than .05, we have sufficient evidence to reject the null hypothesis. Thus, there is a difference in the means of monetary values between clusters.

Table 19: ANOVA Test for Recency

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
Recency	5	127.13	25.426	29.28	<2e-16
Residuals	100	86.83	0.868		

Interpretation: P-value (<2e-16) is less than .05, we have sufficient evidence to reject the null hypothesis. Thus, there is a difference in the means of recency values between clusters.

Table 20: ANOVA Test for Loyalty

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
Loyalty	5	26.49	5.297	20.93	2.91e-14
Residuals	100	25.30	0.253		

Interpretation: P-value (<2e-16) is less than .05, we have sufficient evidence to reject the null hypothesis. Thus, there is a difference in the means of loyalty values between clusters.

Table 21: ANOVA Test for Dependence

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
Dependence	5	48.10	9.621	29.52	<2e-16
Residuals	100	32.59	0.326		

Interpretation: P-value (<2e-16) is less than .05, we have sufficient evidence to reject the null hypothesis. Thus, there is a difference in the means of dependence values between clusters.

Table 22: ANOVA Test for Expectation

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
Expectation	5	18.288	3.658	63.89	<2e-16
Residuals	100	5.725	0.057		

Interpretation: P-value (<2e-16) is less than .05, we have sufficient evidence to reject the null hypothesis. Thus, there is a difference in the means of expectation values between clusters.

Table 23: Levene Test for Frequency

	Df	F value	Pr(>F)
group	5	0.6721	0.6455
	100		

Table 24: Levene Test for Monetary

	Df	F value	Pr(>F)
group	5	0.539	0.7463
	100		

Table 25: Levene Test for Recency

	Df	F value	Pr(>F)
group	5	0.4479	0.8139
	100		

Table 26: Levene Test for Loyalty

	Df	F value	Pr(>F)
group	5	0.7106	0.6168
	100		

Table 27: Levene Test for Dependence

	Df	F value	Pr(>F)
group	5	0.6697	0.6473
	100		

Table 28: Levene Test for Expectation

	Df	F value	Pr(>F)
group	5	1.1095	0.3602
	100		

Table 29: Tamhane T2 Test for Frequency

Clusters	1	2	3	4	5
2	7.2e-14				
3	0.995	0.669			
4	3.1e-05	0.806	0.828		
5	4.0e-08	6.2e-09	1.000	0.024	
6	< 2e-16	0.994	0.745	0.994	1.8e-09

- As a result of the test, there is a significant difference between Group 1 and Group 2; Group 1 and Group 4; Group 1 and Group 5; Group 1 and Group 6; Group 2 and Group 5; Group 4 and Group 5; Group 5 and Group 6 because they are lower than 0.05.

Table 30: Tamhane T2 Test for Monetary

Clusters	1	2	3	4	5
2	4.0e-14				
3	0.9999	0.4500			
4	2.2e-06	0.9668	0.4878		
5	1.7e-06	1.6e-09	0.9996	0.0011	
6	2e-16	0.9943	0.5195	1.0000	1.8e-10

- As a result of the test, there is a significant difference between Group 1 and Group 2; Group 1 and Group 4; Group 1 and Group 5; Group 1 and Group 6; Group 2 and Group 5; Group 4 and Group 5; Group 5 and Group 6.

Table 31: Tamhane T2 Test for Recency

Clusters	1	2	3	4	5
2	0.00249				
3	1.00000	0.63258			
4	2.1e-06	0.99993	0.53397		
5	0.00052	0.42148	0.99708	0.00505	
6	< 2e-16	0.83513	0.49019	0.97811	1.3e-07

- As a result of the test, there is a significant difference between Group 1 and Group 2; Group 1 and Group 4; Group 1 and Group 5; Group 1 and Group 6; Group 4 and Group 5; Group 5 and Group 6.

Table 32: Tamhane T2 Test for Loyalty

Clusters	1	2	3	4	5
2	0.0005				
3	0.4745	0.8557			
4	0.0039	0.9971	0.7492		
5	0.0148	0.0711	0.6078	0.5354	
6	1.0000	0.0009	0.4807	0.0081	0.1096

- As a result of the test, there is a significant difference between Group 1 and Group 2; Group 1 and Group 4; Group 1 and Group 5; Group 2 and Group 6; Group 4 and Group 6.

Table 33: Tamhane T2 Test for Dependence

Clusters	1	2	3	4	5
2	2.0e-08				
3	0.13675	1.00000			
4	0.00184	0.22529	0.89458		
5	2.7e-11	0.00816	0.84571	1.00000	
6	0.24296	1.4e-06	0.16884	0.07406	0.00068

- As a result of the test, there is a significant difference between Group 1 and Group 2; Group 1 and Group 4; Group 1 and Group 5; Group 2 and Group 5; Group 2 and Group 6; Group 5 and Group 6.

Table 34: Tamhane T2 Test for Expectation

Clusters	1	2	3	4	5
2	1.00				
3	0.26	0.25			
4	7.6e-07	5.6e-07	0.68		
5	0.63	1.00	0.28	1.9e-06	
6	1.00	1.00	0.27	7.5e-07	1.00

- As a result of the test, there is a significant difference between Group 1 and Group 4; Group 2 and Group 4; Group 4 and Group 5; Group 4 and Group 6.

6.7 Regression

In this study, regression analyses are made to examine the relationship between Recency, Frequency and Monetary variables and other variables based on clusters. The mathematical formula of the multiple linear regression can be written as,

Equation 2: Multiple Linear Regression Formula

$$y_i = b_1 x_1 + b_2 x_2 + ... + b_n x_n + c$$

$x_1, x_2, .., x_n$: independent variable

y: dependent variable

H_0: There is no relationship between the dependent variable and the independent variable(s)

H_1: There is a relationship between the dependent variable and the independent variable(s)

Table 35: The Relationship between Frequency and other variables for the Cluster 1

Coefficients:

	Estimate	Std. Error	t value	$Pr(>|t|)$
(Intercept)	6.69075	2.56590	2.608	0.0149
Loyalty	-0.02260	0.28965	-0.078	0.9384
Dependence	-0.56337	0.23468	-2.401	0.0238
Expectation	0.07273	0.56036	0.130	0.8977

Residual standard error: 0.5742 on 26 degrees of freedom

Multiple R-squared: 0.2345, Adjusted R-squared: 0.1462

F-statistic: 2.655 on 3 and 26 DF, p-value: 0.06949

- Dependence is the predictor variable which is statistically significant because p value is less than the usual significance level of 0.05.
- One-unit increase in the Dependence changes by -0.56 value of the Frequency.

Table 36: The Relationship between Recency and other variables for the Cluster 1

Coefficients:

	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	12.8971	3.3807	3.815	0.000756
Loyalty	-0.5698	0.3816	-1.493	0.147465
Dependence	0.1912	0.3092	0.618	0.541635
Expectation	-1.3891	0.7383	-1.882	0.071144

Residual standard error: 0.7566 on 26 degrees of freedom
Multiple R-squared: 0.2153, Adjusted R-squared: 0.1247
F-statistic: 2.378 on 3 and 26 DF, p-value: 0.09286

- Expectation is the predictor variable which is statistically significant because p value is less than the usual significance level of 0.05.
- One-unit increase in the Expectation changes by -1,39 value of the Recency.

Table 37: The Relationship between Monetary and other variables for the Cluster 1

Coefficients:

	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	8.3513	2.8638	2.916	0.00721
Loyalty	-0.2776	0.3233	-0.859	0.39831
Dependence	-0.4192	0.2619	-1.600	0.12162
Expectation	-0.1689	0.6254	-0.270	0.78924

Residual standard error: 0.6409 on 26 degrees of freedom

Multiple R-squared: 0.207,   Adjusted R-squared: 0.1155

F-statistic: 2.262 on 3 and 26 DF, p-value: 0.1049

Table 38: The Relationship between Frequency and other variables for the Cluster 2

Coefficients:

	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	-4.1276	3.2861	-1.256	0.2445
Loyalty	0.3490	0.3089	1.130	0.2912
Dependence	-0.5256	0.2769	-1.898	0.0942
Expectation	1.1407	0.7717	1.478	0.1776

Residual standard error: 0.4116 on 8 degrees of freedom

Multiple R-squared: 0.5353, Adjusted R-squared: 0.361

F-statistic: 3.072 on 3 and 8 DF, p-value: 0.0908

- Dependence is the predictor variable which is statistically significant because p value is less than the usual significance level of 0.1.
- One-unit increase in the Dependence changes by -0.52 value of the Frequency.

Table 39: The Relationship between Recency and other variables for the Cluster 2

Coefficients:

	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	-8.3988	10.4416	-0.804	0.444
Loyalty	-1.2164	0.9814	-1.239	0.250
Dependence	1.6220	-0.8799	1.844	0.102
Expectation	2.3332	2.4521	0.952	0.369

Residual standard error: 1.308 on 8 degrees of freedom

Multiple R-squared: 0.3242, Adjusted R-squared: 0.07081

F-statistic: 1.279 on 3 and 8 DF, p-value: 0.3455

Table 40: The Relationship between Monetary and other variables for the Cluster 2

Coefficients:

	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	-4.1276	3.2861	-1.256	0.2445
Loyalty	0.3490	0.3089	1.130	0.2912
Dependence	-0.5256	0.2769	-1.898	0.0942
Expectation	1.1407	0.7717	1.478	0.1776

Residual standard error: 0.4116 on 8 degrees of freedom

Multiple R-squared: 0.5353, Adjusted R-squared: 0.361

F-statistic: 3.072 on 3 and 8 DF, p-value: 0.0908

- Dependence is the predictor variable which is statistically significant because p value is less than the usual significance level of 0.05.
- One-unit increase in the Dependence changes by -0.52 value of the Monetary.

Table 41: The Relationship between Frequency and other variables for the Cluster 3

Coefficients:

	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	-0.3981	NaN	NaN	NaN
Loyalty	1.2816	NaN	NaN	NaN
Dependence	0.5437	NaN	NaN	NaN
Expectation	NA	NA	NA	NA

Residual standard error: NaN on 0 degrees of freedom

Multiple R-squared: 1, Adjusted R-squared: NaN

F-statistic: NaN on 2 and 0 DF, p-value: NA

- There is not enough observation.

Table 42: The Relationship between Recency and other variables for the Cluster 3

Coefficients:

	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	8.7670	NaN	NaN	NaN
Loyalty	0.2136	NaN	NaN	NaN
Dependence	-2.2427	NaN	NaN	NaN
Expectation	NA	NA	NA	NA

Residual standard error: NaN on 0 degrees of freedom

Multiple R-squared: 1, Adjusted R-squared: NaN

F-statistic: NaN on 2 and 0 DF, p-value: NA

- There is not enough observation.

Table 43: The Relationship between Monetary and other variables for the Cluster 3

Coefficients:

	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	3.3350	NaN	NaN	NaN
Loyalty	1.0680	NaN	NaN	NaN
Dependence	-0.7136	NaN	NaN	NaN
Expectation	NA	NA	NA	NA

Residual standard error: NaN on 0 degrees of freedom

Multiple R-squared: 1, Adjusted R-squared: NaN

F-statistic: NaN on 2 and 0 DF, p-value: NA

- There is not enough observation.

Table 44: The Relationship between Frequency and other variables for the Cluster 4

Coefficients:

	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	6.0102	6.7808	0.886	0.405
Loyalty	-0.5980	0.6737	-0.888	0.404
Dependence	0.5912	0.4420	1.338	0.223
Expectation	-0.9157	1.2541	-0.730	0.489

Residual standard error: 0.9904 on 7 degrees of freedom

Multiple R-squared: 0.3134, Adjusted R-squared: 0.01913

F-statistic: 1.065 on 3 and 7 DF, p-value: 0.4229

Table 45: The Relationship between Recency and other variables for the Cluster 4

Coefficients:

	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	10.1455	6.0006	1.691	0.135
Loyalty	-0.7133	0.5962	-1.196	0.270
Dependence	-0.1134	0.3911	-0.290	0.780
Expectation	-1.3700	1.1098	-1.234	0.257

Residual standard error: 0.8764 on 7 degrees of freedom

Multiple R-squared: 0.2218, Adjusted R-squared: -0.1118

F-statistic: 0.6649 on 3 and 7 DF, p-value: 0.5997

Table 46: The Relationship between Monetary and other variables for the Cluster 4

Coefficients:

	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	3.5315	6.7810	0.521	0.619
Loyalty	-0.3267	0.6737	-0.485	0.643
Dependence	0.2957	0.4420	0.669	0.525
Expectation	-0.3559	1.2542	-0.284	0.785

Residual standard error: 0.9904 on 7 degrees of freedom

Multiple R-squared: 0.1008, Adjusted R-squared: -0.2846

F-statistic: 0.2616 on 3 and 7 DF, p-value: 0.851

Table 47: The Relationship between Frequency and other variables for the Cluster 5

Coefficients:

	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	3.38865	2.93599	1.154	0.2598
Loyalty	0.01781	0.30925	0.058	0.9546
Dependence	-0.61996	0.31041	-1.997	0.0573
Expectation	0.34958	0.67736	0.516	0.6105

Residual standard error: 0.7437 on 24 degrees of freedom

Multiple R-squared: 0.1553, Adjusted R-squared: 0.04968

F-statistic: 1.47 on 3 and 24 DF, p-value: 0.2476

Interpretations:

- Dependence is the predictor variable which is statistically significant because p value is less than the usual significance level of 0.05.
- One-unit increase in the Dependence changes by -0.62 value of the Frequency while other variables are constant.

Table 48: The Relationship between Recency and other variables for the Cluster 5

Coefficients:

	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	6.8682	4.2424	1.619	0.119
Loyalty	0.2582	0.4469	0.578	0.569
Dependence	0.1296	0.4485	0.289	0.775
Expectation	-1.0647	0.9788	-1.088	0.287

Residual standard error: 1.075 on 24 degrees of freedom

Multiple R-squared: 0.05243, Adjusted R-squared: -0.06601

F-statistic: 0.4427 on 3 and 24 DF, p-value: 0.7246

Table 49: The Relationship between Monetary and other variables for the Cluster 5

Coefficients:

	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	4.68762	2.75363	1.702	0.1016
Loyalty	0.23309	0.29004	0.804	0.4295
Dependence	-0.71516	0.29113	-2.457	0.0216
Expectation	-0.02693	0.63529	-0.042	0.9665

Residual standard error: 0.6975 on 24 degrees of freedom

Multiple R-squared: 0.2045, Adjusted R-squared: 0.1051

F-statistic: 2.057 on 3 and 24 DF, p-value: 0.1327

Interpretations:

- Dependence is the predictor variable which is statistically significant because p value is less than the usual significance level of 0.05.

- One-unit increase in the Dependence changes by -0.72 value of the Monetary while other variables are constant.

Table 50: The Relationship between Frequency and other variables for the Cluster 6

Coefficients:

	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	-1.68615	4.41560	-0.382	0.707
Loyalty	-0.02420	0.38106	-0.064	0.950
Dependence	-0.05498	0.23262	-0.236	0.816
Expectation	0.75618	0.73618	1.027	0.318

Residual standard error: 0.685 on 18 degrees of freedom

Multiple R-squared: 0.07093,   Adjusted R-squared: -0.08391

F-statistic: 0.4581 on 3 and 18 DF, p-value: 0.7149

Table 51: The Relationship between Recency and other variables for the Cluster 6

Coefficients:

	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	-5.1728	4.3573	-1.187	0.251
Loyalty	0.2950	0.3760	0.785	0.443
Dependence	0.1182	0.2295	0.515	0.613
Expectation	1.0318	0.7265	1.420	0.173

Residual standard error: 0.676 on 18 degrees of freedom

Multiple R-squared: 0.1301, Adjusted R-squared: -0.01487

F-statistic: 0.8974 on 3 and 18 DF, p-value: 0.4617

Table 52: The Relationship between Monetary and other variables for the Cluster 6

Coefficients:

	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	-0.5937	4.3882	-0.135	0.894
Loyalty	-0.1784	0.3787	-0.471	0.643
Dependence	0.1713	0.2312	0.741	0.468
Expectation	0.4963	0.7316	0.678	0.506

Residual standard error: 0.6808 on 18 degrees of freedom

Multiple R-squared: 0.08243,   Adjusted R-squared: -0.0705

F-statistic: 0.539 on 3 and 18 DF, p-value: 0.6616

As a result of that, regression analyses show that

- dependence is the predictor variable of frequency and expectation is the predictor variable of recency in Cluster 1,
- dependence is the predictor variable of frequency and monetary in Cluster 2 and 5.

7. DISCUSSION

This study aims to create a new customer segmentation model by using RFM values with other three variables which are Loyalty, Dependence and Expectation. Specifically, the developed model is applied to a rug&carpet company. The new model has the advantage of providing broad information about customers in terms of buying behavior. The study has also examined how variables have a relationship with each other. All analyses are made via R-Studio.

Primary and secondary datasets are merged in the analysis. First, the sales dataset from 2021 to the 17th of April, 2022 as a secondary dataset covers the sales of dealers. Second, a survey is made toward dealers with 40 questions under seven topics. 152 dealers answer the questions and 106 respondents are included in the analysis after data cleaning.

In this study, the elbow method is used to determine the best customer number to cluster, and the results as 6. Based on two datasets, K-means clustering is applied and achieved 66,5% which is the ratio of the total sum of squares and the between-cluster sum of squares. Surely, the higher the ratio is for the formation of more homogeneous clusters, the better, but 66,5% seems to be sufficient for this study.

As a result of K-means clustering, Group 1 includes the most considerable customers for the company. Because it has high Frequency, Monetary, Recency, Loyalty and Dependence values. On the other hand, Group 6 can be a group to focus on that has low Frequency, Monetary, Recency values but high Loyalty, Dependence and Expectation.

After the clusters are created, correlation and regression between variables are controlled. Loyalty has a positive relationship with expectation and dependence. Thus, if a dealer has high loyalty value, it has more product and service expectation and also high dependency. On the other hand, it is seen that the variable that affects sales the most is dependency.

Hence, ANOVA is applied to each variable to see how means are different in clusters and shown with tables.

8. FUTURE PLANNING FOR CLUSTERS

Each cluster for the rug&carpet company has distinctive features. Therefore, understanding customers is critical to act based on their needs and wants. In this study, six customer groups are created and especially some of them are more substantial than others. At this point, the company can treat with a different strategy for each goup.

Customer retention is as important as gaining them. For this reason, Group 1 is the most significant one which the company should show great respect. In certain periods, different campaigns can be organized for these customers. Thus, they will feel more worthy. On the other hand, the company has two groups which are really low sales values and also low loyalty and dependence with the company. Surely, the company should try to keeping customers. First, it can try to strengthen their loyalty to the company. For that, it can arrange events and organizations for a good relationship among company staff and dealer employees. Hence, Group 3 and Group 5 seems have some problems with the company because they have high sales values but their loyalty and dependence values are low. In this situation, when they find a better manufacturer, it may break the cooperation with the company. Thus, the company should find out problems of the dealers that is the first step for future strategy. Finally, Group 6 behaves differently by showing high loyalty, dependence and expectation from the company despite its low sales volume. In this case, it would not be wrong to say: These dealers want to work with the company, but they need help on how to make sell. As a consequence, the company can plan trainings to give information about the work and control the sales of these dealers on a regular basis.

Not all customers may show loyalty to the company in the same way or the efficiency they receive from the company is not same. Notwithstanding, the company can comprehend its dealers which behaves differently from each other. Related to that, a new strategy can be developed to collobarate better to the customers.

9. CONCLUSION

The importance of customer segmentation gives rise day by day because businesses have returned from a product-oriented view to a customer-oriented view. Businesses especially companies focus on customer groups to get a better understanding of customer behavior and develop strategies uniquely. This situation surely obligates an effective segmentation for the companies. Regarding that, there is an easy method for customer segmentation named RFM and RFM includes three critical points of information about the customers which are recency, frequency, and monetary. Hence, the model could be developed by adding extra variables to make out customers better.

This study apprises the company which segments of its customers are more profitable and whether they are loyal or not. In addition, it informs the company on how can be created an optimum cluster. Thus, a marketing plan for each cluster can be applied specifically.

This study aims to propose a developed model for customer segmentation and the model is created by adding extra criterias to the RFM model. These criterias are Loyalty, Dependence and Expectation. They are important to understand customers in an effective way.

On the other hand, this study recommends K-means clustering with the RFM model to cluster customers. The developed RFM is applied to a rug&carpet company and six different customer groups are achieved. All studies are generated R studio, an analyzing tool in data science. Besides, the model could be applied to other firms and examined how the model is more effective than standard the RFM model.

APPENDIX A: SURVEY QUESTIONNAIRE

This survey aims to contribute to the creation of a new model in customer segmentation. In terms of the scientific nature of the research, it is extremely important that you answer the questions sincerely. Also, it is important that only one person from each of our dealers should participate in the survey. Thank you in advance for your participation.

The options are numbered 1 to 5 and represent the following statements.

Strongly Disagree	1
Disagree	2
Neutral	3
Agree	4
Strongly Agree	5

A) Personal Information:

- Age
- Gender

Woman/ Man

- How long have you been working in the company?
- Education status:

Primary Education / High School / College / Undergraduate/ Postgraduate

- Position in the institution

Business Owner / Middle Manager / Top Manager / Other

- Sector experience

B) Information about the sales point:
- Name of the sales point
- Name of the wholesale
- How long has your company been working with this manufacturer company?

	Loyalty					
1	We plan to continue with this brand for 5 years.	1	2	3	4	5
2	The manufacturer produces highly qualified products at the right price.	1	2	3	4	5
3	When a product is marked up, we don't lessen the quantity of buying because we believe that it is the right price of the product.	1	2	3	4	5
4	When the manufacturer makes a mistake, we do not immediately break the collaboration because we believe that the problem will be resolved sensibly.	1	2	3	4	5
5	We can safely recommend the manufacturer to anyone who considers working with the manufacturer.	1	2	3	4	5
6	We are satisfied with the service that has been received from the manufacturer.	1	2	3	4	5
7	We are satisfied with the products that have been produced by the manufacturer.	1	2	3	4	5
8	We often visit the website and review the products or interact with the social media channels of the manufacturer.	1	2	3	4	5

	Confidence in the Brand					
1	Even when our manufacturer makes an unbelievable statement, we believe that it is true.	1	2	3	4	5
2	Our manufacturer generally keeps its promises to our company.	1	2	3	4	5
3	We know that the manufacturer provides the best advice regarding our business.	1	2	3	4	5
4	We trust the sincerity of our manufacturer company.	1	2	3	4	5

5	Our manufacturer company often gives us true information.	1	2	3	4	5
6	When our manufacturer company makes important decisions, it considers our well-being.	1	2	3	4	5
7	When we share our problems with our manufacturer, we think that they will understand us.	1	2	3	4	5
8	We trust our manufacturer will consider how its actions will affect us when making a new decision.	1	2	3	4	5
9	The manufacturer rewards its loyal customers.	1	2	3	4	5

	Product Expectation					
1	The manufacturer should have the best carpet product groups in the market.	1	2	3	4	5
2	The manufacturer should have a wide product range in the market.	1	2	3	4	5
3	The manufacturer should offer its products to the market at competitive prices.	1	2	3	4	5
4	The manufacturer's products meet customers' expectations.	1	2	3	4	5
5	Innovative products attract the attention of consumers and they should be in the manufacturer company.	1	2	3	4	5

	Service Expectation					
1	Fast delivery to the customer is very important for the quality of service.	1	2	3	4	5
2	After-sales services should be provided directly by the manufacturer.	1	2	3	4	5
3	The manufacturer has to provide broad information about the products.	1	2	3	4	5
4	It is sufficient for the manufacturer to keep stock itself to serve the consumer.	1	2	3	4	5

5	It is very crucial to take advantage of the information technology-supported services provided by the manufacturer.	1	2	3	4	5

	Trust					
1	The manufacturer has an absolutely honest understanding.	1	2	3	4	5
2	We have full trust in the manufacturer company in every respect.	1	2	3	4	5
3	We trust it to do the right thing.	1	2	3	4	5
4	We trust the manufacturer to do the job right.	1	2	3	4	5
5	The manufacturer is always true-hearted.	1	2	3	4	5

	Dependence					
	(a) dealer dependence on the supplier	1	2	3	4	5
1	Our success in the business depends heavily on the marketing efforts of the manufacturer.	1	2	3	4	5
2	Despite all efforts, our sales would drop significantly if we no longer represent this manufacturer.	1	2	3	4	5
3	We can't easily find an alternative manufacturer to replace the manufacturer.	1	2	3	4	5
4	The manufacturer is in a key position that affects our overall business performance.	1	2	3	4	5
	(b) dealer perceptions of supplier dependence	1	2	3	4	5
1	If we no longer represent this manufacturer, its sales in our region will drop significantly.	1	2	3	4	5
2	The supplier can easily switch to another dealer in our trading area.	1	2	3	4	5

| 3 | We are a key selling point in our trading area for the supplier's products. | 1 | 2 | 3 | 4 | 5 |
| 4 | We create high sales volume for this supplier. | 1 | 2 | 3 | 4 | 5 |

REFERENCES

[1] Beavers, A. S., Lounsbury, J. W., Richards, J. K., Huck, S. W., Skolits, G. J., & Esquivel, S. L. (2013). Practical considerations for using exploratory factor analysis in educational research. *Practical Assessment, Research, and Evaluation*, *18*(1), 6.

[2] Buckinx, W., & Van den Poel, D. (2005). Customer base analysis: partial defection of behaviourally loyal clients in a non-contractual FMCG retail setting. *European journal of operational research*, *164*(1), 252-268.

[3] Celly, K. S., & Frazier, G. L. (1996). Outcome-based and behavior-based coordination efforts in channel relationships. *Journal of marketing research*, *33*(2), 200-210.

[4] Chen, Y. L., Kuo, M. H., Wu, S. Y. & Tang, K. (2009). Discovering recency, frequency, and monetary (RFM) sequential patterns from customers' purchasing data. *Electronic Commerce Research and Applications*, 8(5), 241-251.

[5] Chen, H., Zhang, L., Chu, X., & Yan, B. (2019). Smartphone customer segmentation based on the usage pattern. *Advanced Engineering Informatics*, *42*, 101000.

[6] Choudhuri, P. S., & Parida, B. B. (2014). Evaluation of Customers Expectation-Perception Score on Service Quality in Life Insurance Corporation of India. *International Journal of Marketing & Business Communication*, *3*(3).

[7] Coussement, K., & Van den Poel, D. (2009). Improving customer attrition prediction by integrating emotions from client/company interaction emails and evaluating multiple classifiers. *Expert Systems with Applications*, *36*(3), 6127-6134.

[8] Dimitriadis, S., Kouremenos, A., & Kyrezis, N. (2011). Trust-based segmentation: Preliminary evidence from technology-enabled bank channels. *International Journal of Bank Marketing*.

[9] Dogan, O., Ayçin, E., & Bulut, Z. (2018). Customer segmentation by using RFM model and clustering methods: a case study in retail industry. *International Journal of Contemporary Economics and Administrative Sciences*, *8*.

[10] Driggs, W. W., Ramsey, S. S., & Nunes, P. F. (2006). Think your customers are loyal? Think again. *Outlook Journal, September,[Online] Available at:< http://www. accenture. com/Global/Research_and_Insights/Outlook/By_Issue Y*, *2006*.

[11] Emerson, R. M. (1962). Power-dependence relations. *American sociological review*, 31-41.

[12] Ganesan, S. (1994). Determinants of long-term orientation in buyer-seller relationships. *Journal of marketing*, *58*(2), 1-19.

[13] Güçdemir, H. (2013). Customer segmentation using a fuzzy ahp and clustering based approach: An application in an international TV manufacturing company (Doctoral dissertation, DEÜ Fen Bilimleri Enstitüsü).

[14] Güçdemir, H., & Selim, H. (2015). Integrating multi-criteria decision making and clustering for business customer segmentation. *Industrial Management & Data Systems*.

[15] Hajiha, A., Radfar, R., & Malayeri, S. S. (2011, December). Data mining application for customer segmentation based on loyalty: An iranian food industry case study. In *2011 IEEE International Conference on Industrial Engineering and Engineering Management* (pp. 504-508). IEEE.

[16] Hosseini, S. M. S., Maleki, A., & Gholamian, M. R. (2010). Cluster analysis using data mining approach to develop CRM methodology to assess the customer loyalty. *Expert Systems with Applications*, *37*(7), 5259-5264.

[17] Inaba, M., Katoh, N., & Imai, H. (1994, June). Applications of weighted Voronoi diagrams and randomization to variance-based k-clustering. In *Proceedings of the tenth annual symposium on Computational geometry* (pp. 332-339).

[18] Jintana, J., & Mori, T. (2019). Customer clustering for a new method of marketing strategy support within the courier business. *Academia Book Chapter*, *31*(2), 1-19.

[19] Khajvand, M., & Tarokh, M. J. (2011). Estimating customer future value of different customer segments based on adapted RFM model in retail banking context. *Procedia Computer Science*, *3*, 1327-1332.

[20] Kim, T. K. (2017). Understanding one-way ANOVA using conceptual figures. *Korean journal of anesthesiology*, *70*(1), 22.

[21] Kodinariya, T. M., & Makwana, P. R. (2013). Review on determining number of Cluster in K-Means Clustering. *International Journal*, *1*(6), 90-95.

[22] Knox, S. (1998). Loyalty-based segmentation and the customer development process. *European Management Journal*, *16*(6), 729-737.

[23] Kumar, N., Scheer, L. K., & Steenkamp, J. B. E. (1995). The effects of supplier fairness on vulnerable resellers. *Journal of marketing research*, *32*(1), 54-65.

[24] Laroche, M., & Sadokierski, R. (1994). Role of confidence in a multi-brand model of intentions for a high-involvement service. *Journal of Business Research*, *29*(1), 1-12.

[25] Laroche, M., Kim, C., & Zhou, L. (1996). Brand familiarity and confidence as determinants of purchase intention: An empirical test in a multiple brand context. *Journal of business Research*, *37*(2), 115-120.

[26] Leigh, J. H., & Martin Jr, C. R. (1987). "Don't know" item nonresponse in a telephone survey: Effects of question form and respondent characteristics. *Journal of Marketing Research*, *24*(4), 418-424.

[27] Mousaeirad, S. (2020). Intelligent vector-based customer segmentation in the banking industry. arXiv preprint arXiv:2012.11876.

[28] Malik, M. E., Ghafoor, M. M., Hafiz, K. I., Riaz, U., Hassan, N. U., Mustafa, M., & Shahbaz, S. (2013). Importance of brand awareness and brand loyalty in assessing purchase intentions of consumer. *International Journal of business and social science*, *4*(5).

[29] Mao, J. (2010). Customer brand loyalty. *International journal of business and management*, *5*(7), 213.

[30] Monalisa, S., Nadya, P., & Novita, R. (2019). Analysis for customer lifetime value categorization with RFM model. *Procedia Computer Science*, *161*, 834-840.

[31] Pavlou, P.A. and Fygenson, M. (2006). Understanding and predicting electronic commerce adoption: an extension of the theory of planned behavior. *MIS Quarterly,* Vol. 30 No. 1, pp. 115-43.

[32] Spring, P., Leeflang, P. S. & Wansbeek, T. (1999). The combination strategy to optimal target selection and offer segmentation in direct mail. Journal of Market-Focused Management, 4(3), 187-203.

[33] Srivastava, M., & Kaul, D. (2016). Exploring the link between customer experience–loyalty–consumer spend. *Journal of Retailing and Consumer Services*, 31, 277-286.

[34] Tsai, C. F., Hu, Y. H., & Lu, Y. H. (2015). Customer segmentation issues and strategies for an automobile dealership with two clustering techniques. *Expert Systems*, *32*(1), 65-76.

[35] Vattani, A. (2011). K-means requires exponentially many iterations even in the plane. *Discrete & Computational Geometry*, 45(4), 596-616.

[36] Wei, J. T., Lin, S. Y., & Wu, H. H. (2010). A review of the application of RFM model. *African Journal of Business Management*, *4*(19), 4199-4206.

[37] William, O., Appiah, E. E., & Botchway, E. A. (2016). Assessment of customer expectation and perception of service quality delivery in Ghana commercial bank. *Journal of Humanity*, *4*(1), 81-91.

[38] Wu, R. S., & Chou, P. H. (2011). Customer segmentation of multiple category data in e-commerce using a soft-clustering approach. *Electronic Commerce Research and Applications*, *10*(3), 331-341.

[39] Wu, H. H., Lin, S. Y., & Liu, C. W. (2014). Analyzing patients' values by applying cluster analysis and LRFM model in a pediatric dental clinic in Taiwan. *The Scientific World Journal*, *2014*.

[40]   Yeh, I. C., Yang, K. J., & Ting, T. M. (2009). Knowledge discovery on RFM model using Bernoulli sequence. *Expert Systems with Applications*, *36*(3), 5866-5871.