İSTANBUL TEKNİK ÜNİVERSİTESİ ★ FEN BİLİMLERİ ENSTİTÜSÜ

BAYİ DEĞERLENDİRMESİ İÇİN VERİ MADENCİLİĞİ UYGULAMASI

YÜKSEK LİSANS TEZİ

Müh. Ahmet Cüneyt BAYSAL

Anabilim Dalı : İŞLETME MÜHENDİSLİĞİ

Programı : İŞLETME MÜHENDİSLİĞİ

OCAK 2008

İSTANBUL TEKNİK ÜNİVERSİTESİ ★ FEN BİLİMLERİ ENSTİTÜSÜ

BAYİ DEĞERLENDİRMESİ İÇİN VERİ MADENCİLİĞİ UYGULAMASI

YÜKSEK LİSANS TEZİ
Müh. Ahmet Cüneyt BAYSAL
(507021012)

Tezin Enstitüye Verildiği Tarih : 24 Aralık 2007
Tezin Savunulduğu Tarih : 29 Ocak 2008

Tez Danışmanı : Prof.Dr. Demet BAYRAKTAR

Diğer Jüri Üyeleri Doç.Dr. Tufan Vehbi KOÇ

Yrd.Doç.Dr. Hür Bersam BOLAT

OCAK 2008

ÖNSÖZ

Veri madenciliği konusuna olan ilgimin artmasını sağlayan ve tezin hazırlanmasında etkin yönlendirmeleri ile yol gösteren hocam Prof. Dr. Demet BAYRAKTAR'a, yapılan analizler için gerekli olan veri setinin temin edilmesinde gösterdikleri kolaylıklar için uygulama yapılan firmanın yönetici ve çalışanlarına, analizler sırasında yardım ve yorumları için arkadaşım Özgüç AKDAĞ'a içten teşekkürlerimi sunarım.

Ayrıca tüm eğitimim süresince bana her konuda tam destek veren aileme ve bende emeği bulunan tüm hocalarıma saygı ve sevgilerimi sunarım.

Ocak, 2008 **Ahmet Cüneyt BAYSAL**

İÇİNDEKİLER

ÖNSÖZ ii
KISALTMALAR v
TABLO LİSTESİ vi
ŞEKİL LİSTESİ vii
SEMBOL LİSTESİ viii
ÖZET ix
SUMMARY x

1. GİRİŞ 1
1.1. Çalışmanın Kapsamı 2
1.2. Çalışmanın Amacı 2

2. KARAR DESTEK SİSTEMİ 4
2.1. Bilgi Yönetimi 4
2.2. Veri Ambarı 6

3. VERİ MADENCİLİĞİ 9
3.1. Veri Madenciliğinin Tanımı 9
3.2. Veri Madenciliğinin Tarihsel Gelişimi 10
3.3. Veri Madenciliği Uygulama Alanları 12
3.4. Veri Madenciliği Süreci 13
3.4.1. Problemin Tanımlanması 13
3.4.2. Verilerin Hazırlanması 14
3.4.3. Modelin Kurulması ve Değerlendirilmesi 15
3.4.4. Modelin Kullanılması ve İzlenmesi 17
3.5. Veri Madenciliği Modelleri 17
3.5.1. Sınıflandırma ve Regresyon Modelleri 18
3.5.1.1. Karar Ağaçları (Decision Trees) 18
3.5.1.2. Yapay Sinir Ağları (Artificial Neural Network) 19
3.5.1.3. K-En Yakın Komşu (K-Nearest Neighbour) 19
3.5.2. Birliktelik Kuralları ve Ardışık Zamanlı Örüntüler 20
3.5.3. Kümeleme Modelleri 20
3.5.3.1. Hiyerarşik Kümeleme Teknikleri 22
3.5.3.2. Hiyerarşik Olmayan Kümeleme Teknikleri 23

4. MÜŞTERİ İLİŞKİLERİ YÖNETİMİ **25**
4.1. Bölümleme 26
4.1.1. Literatür Araştırması 27
4.1.2. Bölümleme Kriterleri 31
4.1.3. Bölümlemenin Faydaları 32

5. BAYİ YÖNETİM SİSTEMİ MODELLEMESİ **33**
5.1. Modelleme Aşamaları 34
5.1.1. Kümelemeye Esas Olacak Değişkenlerin Seçilmesi 34
5.1.2. Uzaklık Ölçüsüne Karar Verilmesi 37
5.1.3. Kümelemede Kullanılacak Tekniğin Seçilmesi 37
5.1.4. Küme Oluşturma ve Küme Sayısının Belirlenmesi 37
5.1.5. Kümelerin Tanımlanması ve Yorumlanması 38
5.2. Analizler 38
5.2.1. İstatistiksel Analizler 39
5.2.1.1. Faktör Analizi (Anket Değişkenleri) 39
5.2.1.2. Faktör Analizi (Ürün Aileleri Miktar Verileri) 41
5.2.1.3. Diskriminant Analizi (Odeme_Prob) 43
5.2.1.4. Diskriminant Analizi (Per_Yeterlilik) 44
5.2.1.5. Diskriminant Analizi (Isyeri_Buyukluk) 45
5.2.2. Kümeleme Analizleri 48
5.2.2.1. Kümeleme Analizi (Anket Değişkenleri) 48
5.2.2.2. Kümeleme Analizi (Ürün Ailesi Miktar Verileri) 58
5.2.2.3. Kümeleme Analizi (Finansal Veriler) 60
5.2.3. Regresyon Analizi (Top_Ciro ve Teminat_Mik) 61

6. SONUÇ **63**

KAYNAKLAR **66**

EK A - ANKET SORULARI **69**

EK B - FREKANS TABLOLARI **72**

EK C - FAKTÖR ANALİZİ (ANKET) TABLOLARI **75**

EK D - FAKTÖR ANALİZİ (ÜRÜN AİLELERİ) TABLOLARI **78**

EK E - DİSKRİMİNANT ANALİZİ (ODEME_PROB) **79**

ÖZGEÇMİŞ **81**

KISALTMALAR

ERP	: Enterprise Resource Planning
KDS	: Karar Destek Sistemi
MİY	: Müşteri İlişkileri Yönetimi
OLAP	: Online Analytical Processing
SAP	: Systems, Applications, Products in data processing

TABLO LİSTESİ

Sayfa No

Tablo 3.1	Veri Madenciliği Evrimsel Adımları.....	11
Tablo 4.1	Kaynak Künyeleri Tablosu.....	29
Tablo 5.1	Analizde Kullanılan Değişkenler ve Açıklamaları	36
Tablo 5.2	Anketten Elde Edilen Değişkenler ve Açıklamaları.....	40
Tablo 5.3	Grup Merkez Fonksiyonları.....	44
Tablo 5.4	Son Küme Merkezleri.....	50
Tablo 5.5	Kümeleme Analizi Varyans Tablosu.....	51
Tablo 5.6	Standartize Edilmiş Diskriminant Fonksiyon Katsayıları.....	53
Tablo 5.7	Merkez Grup Fonksiyonları.....	54
Tablo 5.8	Küme Üyelerinin Bölgelere Göre Dağılımı.....	56
Tablo 5.9	Küme Merkezleri Tablosu.....	57
Tablo 5.10	Varyans Tablosu.....	58
Tablo 5.11	Küme Merkezleri Tablosu.....	59
Tablo 5.12	Grup Merkezleri Fonksiyonu.....	61
Tablo 5.13	Korelasyon Tablosu.....	62
Tablo 5.14	Model Katsayısı.....	62
Tablo 6.1	Kümeleme Analizleri Özet Tablosu.....	63
Tablo B.1	Bolge_Kodu Değişkeni Frekans Tablosu.....	72
Tablo B.2	Sozlesme Değişkeni Frekans Tablosu.....	72
Tablo B.3	Odeme_Prob Değişkeni Frekans Tablosu.....	73
Tablo B.4	Bayi_Vade Değişkeni Frekans Tablosu.....	73
Tablo B.5	Per_Yeterlilik Değişkeni Frekans Tablosu.....	73
Tablo B.6	Isyeri_Buyukluk Değişkeni Frekans Tablosu.....	74
Tablo B.7	Mulkiyet Değişkeni Frekans Tablosu.....	74
Tablo B.8	Sürekli Değişkenlerin Tanımlayıcı İstatistikleri.....	74
Tablo C.1	Ortak Varyans Tablosu.....	75
Tablo C.2	Faktör Sayısı ve Açıklanan Varyans.....	76
Tablo C.3	Faktör Matrisi.....	76
Tablo C.4	Faktör Skor Katsayı Matrisi.....	77
Tablo D.1	Ortak Varyans Tablosu.....	78
Tablo D.2	Faktör Sayısı ve Açıklanan Varyans.....	78
Tablo D.3	Faktör Matrisi.....	78
Tablo D.4	Faktör Katsayıları Tablosu.....	78
Tablo E.1	Box's M Test Sonuçları.....	79
Tablo E.2	Özdeğer Tablosu.....	79
Tablo E.3	Wilks' Lamda Değer Tablosu.....	79
Tablo E.4	Kanonik Diskriminant Fonksiyon Katsayıları.....	79
Tablo E.5	Grup Merkez Fonksiyonları.....	79
Tablo E.6	Sınıflandırma Başarım Sonuçları.....	80

ŞEKİL LİSTESİ

Sayfa No

Şekil 2.1 **:** Veri, Bilgi ve Kurumsal Bilgi İlişkisi 6
Şekil 2.2 **:** Veri Ambarı Mimarisi 7
Şekil 2.3 **:** Veri Ambarından Elde Edilen Bilgilerin Kullanım Alanları 8
Şekil 3.1 **:** Veri Madenciliği Modelleme Süreci..... 16
Şekil 5.1 **:** Faktör Skorlarına Göre Saçılım Grafiği 42
Şekil 5.2 **:** Kanonik Diskriminant Fonksiyon Grafiği..... 46
Şekil 5.3 **:** Alan Haritası 47
Şekil 5.4 **:** Saçılım Diyagramı..... 61

SEMBOL LİSTESİ

d	**:** İki nesne arasındaki uzaklık
i	**:** Veri matrisinde satır
j	**:** Veri matrisinde sütun
k	**:** Küme sayısı
m	**:** Veri matrisinde boyut
n	**:** Bir değişkenin alabileceği sayı
N	**:** Örnek sayısı
s	**:** Bir denklemin alabileceği değişken değeri
X	**:** Veri matrisinde nesne gösterimi

BAYİ DEĞERLENDİRMESİ İÇİN VERİ MADENCİLİĞİ UYGULAMASI

ÖZET

Yoğun rekabet ortamında müşterilerin tüm süreçlerin odak noktası haline gelmiş olması, gerek mal gerekse hizmet üreten firmaları, müşterilerin istek ve gereksinimlerini doğru tespit eden ve bunları karşılayan bir yapı kazanmaya zorlamaktadır. Bu noktada, müşteri istek ve gereksinimlerini belirleyen, onlarla her türlü iletişimi sağlayan kanalların etkin şekilde yönetilmesi, firmaları rakipleri karşında avantajlı hale getirmektedir.

Bayi yönetimi için bayilerin değerlendirilmesi, müşteri odaklılık ve müşteri ilişkileri yönetimi kapsamında incelenmiştir. Müşteri sayısının her geçen gün artması, onlarla sürekli iletişimde olan bayilerin önemini arttırmakta ve bayi yapısının yönetilmesini zorunlu hale getirmektedir. Müşteri sayısının artması ve firma ile olan iletişimlerinin çeşitlilik arz etmesi incelenmesi gereken veri miktarını da arttırmaktadır. Eş zamanlı analizlerin yapılması ve analizler sonucunda elde edilen bilgiler ile söz konusu ilişkilerin etkinliliğinin arttırılması ancak gelişmiş bilgisayar teknolojileri ve uygun yazılımlar sayesinde olabilmektedir. Bu durum firmaların veri madenciliği konusuna olan ilgisini arttırmıştır. Bununla birlikte firmalar süreçlerinin etkinliğini arttırmak, hızlı ve doğru kararlar almak ve rakipleri karşısında rekabet avantajı sağlamak gibi nedenlerle veri madenciliği tekniklerini uygulamaya başlamışlardır.

Elde edilen bayi profilleri, bayilere ilişkin hedef ve satış kotalarının belirlenmesi, eğitim ihtiyacının belirlenmesi, kampanya yönetimi vb. faaliyetlerinin yürütülmesinde kullanılabileceği gibi daha yakın ilişkilerin tesis edilmesi de sağlanabilecektir. Ayrıca iyi konumda olan bayiler örnek model olarak alınarak diğer bayilerin gelişim planlarının yapılmasında, sahip olunması gereken özelliklerin ve bu özelliklerin seviyesinin tespit edilmesinde, performans kriterlerinin belirlenmesinde kullanılabilecektir. Analizler sonucunda küme başarımlarının testi için uygulanan diskriminant analizi ile elde edilen denklemler, bayilerin belirlenen değişkenler bakımından hangi kümeye üye olacağının tahmin edilmesinde kullanılabilecektir.

Bayilerin öznel değerlendirilmelerinin yapıldığı anket değişkenleri, sipariş edilen ürün miktar değişkenleri ve finansal değişkenlere bağlı olarak en iyi profile sahip olan bayiler 54 ve 58 kodlu bayiler olarak belirlenmiştir. Söz konusu bayilere ait değişkenlerde meydana gelecek değişiklikler genel yapı hakkında bilgi verebilmekte, değişikliklerin dönemsel, mevsimsel veya bölgesel olup olmadığı konusunda karar desteği sağlayabilecektir.

Elde edilen sonuçlar firmanın mevcut veri setinden elde edilmiş olup mevcut durumu ortaya koymaktadır. Veri madenciliği sürekli bir süreç olduğundan yeni bayilikler verilmesi, bayilikten ayrılma ya da mevcut durumda kullanılan değişkenlerde değişikliğe neden olacak durumlarda analizlerin tekrarlanması ve modelin güncellenmesi gerekmektedir.

A DATA MINING APPLICATION IN DEALER EVALUATION

SUMMARY

The fact that the customers have become the focus of all processes due to intensive competition, leads companies of both real and service sector to specify and meet the needs and wants of the customers accurately. At this point companies specifying the needs and wants of the customers and managing the communication channels effectively become advantageous against their rivals.

The evaluation of the branches is conducted in the context of customer orientation and customer relationships management. The ever increasing number of customers makes the branches more important, which are in close contact with the customers and forces the companies to manage their branch structure. The increasing number of customers and various types of their communications with the company leads the enhancement data amount to be analyzed. Conducting simultaneous analysis and increasing the effectiveness of the communication according to the findings of these analyses can only be fulfilled through developed computer technologies and appropriate software. This fact led the increase of interest to data mining. So the companies began applying data mining techniques to increase the effectiveness of their processes, to have competitive advantage against their rivals, making rapid and accurate decisions.

The liaison offices of a leading company, which reaches the final customers through its branches, are evaluated through one of the data mining techniques called cluster analysis, with the data collected from the actualized transactions and questionnaires for obtaining the subjective evaluations of branches in the context of the thesis work, and a model for Branch Management System is constituted. Cluster analysis is a multivariate statistical analysis technique consisting of different algorithms, enables the classification of the objects according to their specified features. Instead of individual selection, cluster analysis is based on group selection which enables evaluation of the objects as a whole. Having a wide range of application area, cluster analysis techniques have an increasing importance. In this study, various variables are used for clustering the most likely branches according to different variable sets and cluster profiles are determined.

The obtained branch profiles can be used for determining the sales targets and quotes, training requirements, and campaign management and also for establishing closer relations. The branches which are found to perform better can be used as benchmark models for the development plans and performance criteria to set. Clusters are being analyzed via discriminant analysis, where the discriminant facto equations are to be used for prediction of a new branch into these clusters.

According to the subjective questionnaire variables and quantitative data such as order quantities and financial variables, the best performing branches are found as no 54 and 58. The changes at the variables of these branches are informative for the

general structure and also may be useful for decision support in case of cyclical, seasonal and regional varieties.

The results are found from the actual data set of the branches, and represent the actual situation. Data mining is a permanent process so results should be updated in case of branch openings, closings, and any differences in the data and variable structure.

1.GİRİŞ

Günümüz yoğun rekabet ortamında müşterilerin tüm süreçlerin odak noktası haline gelmiş olması, gerek mal gerekse hizmet üreten firmaları müşterilerin istek ve gereksinimlerini doğru tespit eden ve bunları karşılayan bir yapı kazanmaya zorlamaktadır. Bu noktada, müşteri istek ve gereksinimlerini belirleyen, onlarla her türlü iletişimi sağlayan kanalların etkin şekilde yönetilmesi, firmaları rakipleri karşında avantajlı hale getirmektedir.

Firmalar, müşteriler başta olmak üzere ilişkili olduğu tüm taraflar ve süreçler ile ilgili verileri toplamak durumundadır. Bu verilerin işlenmesi ile elde edilecek bilgiler, işletmelerin verimliliğini arttırıcı ve rekabette fark yaratıcı unsur olmaktadır. Çünkü verilerin toplanıp işlenmesi ile kârlı, sadık müşteriler belirlenebilmekte ve özelleşmiş pazarlama teknikleri kullanılarak etkin kaynak kullanımı sağlanabilmektedir.

Firmaların ilişkili olduğu taraflar ve süreçlerden elde ettikleri veri miktarı her geçen gün artmakta, bu durum ise tedarikten satış sonrası hizmetlere kadar tüm faaliyetlerde, verilerden kullanılabilir bilgilerin elde edilebilmesi için gerekli teknolojik gelişmeleri hızlandırmaktadır. Bilgisayar ve iletişim teknolojilerindeki gelişmeler, çok daha fazla verinin daha hızlı ve az maliyetli olarak toplanmasına, depolanmasına, işlenmesine ve bilgiye dönüştürülüp gerekli yerlere iletilmesine imkân vermektedir.

Tüm bu gelişmeler firmaların, veri ambarlama ve veri madenciliği teknolojilerine odaklanmalarını ve bu alanlara yatırım yapmalarını hızlandırmıştır. Bu yatırımların etkin sonuç verebilmesi ise, öncelikle işletme amaçlarının net olarak belirlenmesine ve bu noktalara odaklanılmasına bağlıdır. Oldukça önemli bir süreç olan veri madenciliğinin anlaşılması ve bu alana yönelik doğru çözümlerin seçilebilmesi işletmelere, faaliyetlerinde hız kazandırmakta ve gelecekle ilgili daha doğru kararlar ile ciddi rekabet avantajı sağlamaktadır.

1.1. Çalışmanın Kapsamı

Firmaların nihai müşterilerle iletişimi doğrudan olabildiği gibi firmanın bayileri aracılığıyla da olmaktadır. Dolayısıyla göz ardı edilmemesi gereken ve yönetilmesi bir zorunluluk olan bayi kavramı oldukça önemli bir süreçtir. Tüm süreçlerin müşteri odaklılık etrafında kurulduğu bir yapıda, bayi yönetimi, nihai müşteriler hakkında fikir verirken, onlara ulaşmada etkin bayilik sisteminin nasıl olması gerektiği hakkında da bilgi desteği sağlar.

Bu kapsamda, çalışmada teorik olarak veri madenciliği konusu ele alınmış ve buna ek olarak nihai müşterilerine bayileri aracılığı ile ulaşan bir firmanın bayi verileri incelenmiş, bayi profilleri çıkarılmış, belirlenen kriterlere göre bayiler sınıflandırılarak veri madenciliği tekniklerinden Kümeleme Analizi kullanılarak Bayi Yönetim Sistemi için bir model geliştirilmiştir.

Uygulamada kullanılan, bayi sipariş miktarları, ciroları, ödeme vadeleri vb. veriler firmanın kullandığı SAP sisteminden elde edilmiştir. SAP sisteminin 01.01.2007 tarihinde hayata geçmesi nedeniyle bu veriler 01.01.2007–30.09.2007 tarihleri arasındaki dokuz aylık dönemi kapsamaktadır. Bu verilere ek olarak sayısal olmayan, öznel nitelikli verilerin konu ile ilgili, bilgisi olan çalışanlardan elde edilmesi için, gerek literatürde yer alan, gerekse konu ile ilgili soruları kapsayan anket hazırlanmış ve analiz bu anket sonuçlarıyla genişletilmiştir.

1.2. Çalışmanın Amacı

Firma ile bayileri arasında gerçekleşen çeşitli faaliyetler sonucunda oluşan verilerin toplanıp, işlenmesi ve analiz edilerek çeşitli bilgiler elde edilmesi ve bunların yine firma ile bayi ilişkilerinin etkinliğini arttırmak için kullanılmasına yönelik bir model oluşturulması amaçlanmıştır. Bu kapsamda bayilerin satın alma verileri, sipariş yapıları, satın aldıkları ürünler, firmaya sağladıkları ciro vb. çeşitli verilerin yanı sıra öznel değerlendirmeleri içeren anket hazırlanmıştır.

Söz konusu değişkenlerin analiz edilmesiyle, mevcut durum hakkında bilgi sağlanması ve değişkenler arasındaki ilişkilerden hareketle model oluşturulması amaçlanmıştır.

Bayi profili geliştirilmesi için kendi sektöründe lider ve Türkiye'de geniş bayi ağına sahip bir firmanın bayi verileri incelenmiştir. Uygulama yapılan firmanın pazarlama ve satış temsilcileri, müşteri ilişkileri sorumlusu, lojistik sorumlusu ve muhasebe personellerinden elde edilen bilgiler ışığında aşağıdaki katkıların sağlanması amaçlanmıştır:

- Bayi verilerinin incelenmesi ile firmanın bayi profilinin belirlenmesi
- Bayileri değişik kriterlere göre sınıflandırarak etkin bayi modeli geliştirilmesi
- Başarılı bayi profilinin elde edilmesi ile performans kriterlerinin tespit edilmesi
- Etkin bayi yönetim sistemi geliştirilmesine katkı sağlanması
- Üretim planlamaya katkı sağlanması
- Orta ve üst kademe yöneticilere karar desteği oluşturulması

2. KARAR DESTEK SİSTEMİ

İletişim teknolojilerindeki büyük gelişmeler mesafelerin kısalmasına, işletmelerin pazarlarının gelişmesine, dolayısıyla daha fazla müşteriye ulaşmalarına yardımcı olmuştur. Ancak bununla birlikte rekabet şartları da daha farklı boyutlar kazanmıştır. Gerek müşteri sayısının, gerekse firmaların faaliyetlerini yürütürken göz önünde bulundurması gereken faktörlerin artmış olması karar verme durumunda olan kişilerin işlerini zorlaştırmaktadır. Bilgi teknolojilerindeki gelişmeler ise, karar vericilere buna karşı kullanabilecekleri araçlar sağlamaktadır. Bu sayede, kısa sürede çok fazla faktörü göz önünde bulundurarak faydalı faaliyet bilgilerinin elde edilmesi mümkün olmaktadır. Bunun için firmaların bünyelerinde Karar Destek Sistemleri (KDS) oluşturmaları gerekmektedir.

Karar Destek Sistemleri (KDS), toplanan verinin sistem, araç ve tekniklerinin yazılım ve donanımı destekleyici şekilde koordine edilmesidir. Organizasyon bu bilgileri gerek kendi bünyesinde pazarlama, finans, üretim vb. süreçlerinden gerekse iş çevresinden toplayarak yorumlayarak tüm süreçlerinde kullanılacak hale getirebilmektedir [8]. Karar destek sistemlerinin ana amacı, yönetimsel problemlerin çeşitli açılardan anlaşılması ve zaman, ekonomik, malzeme vb. kısıtlara bağlı olarak en iyi çözümlerin bulunmasına yardımcı olmasıdır [19]. KDS'ler doğrudan çözüm sağlamamakla birlikte, çözüm elde etmek için araç durumundadır [36].

Karar destek sistemlerinde, veri madenciliği teknikleri, çok çeşitli alanlarda karar verme durumunda olanlar için gerekli, gizli ve bazı durumlarda çıkarımı oldukça zor olan bilgilerin elde edilmesi amacıyla yoğun olarak kullanılmaktadır.

2.1. Bilgi Yönetimi

Karar destek sistemi kapsamında, konu bütünlüğü sağlamak amacıyla Bilgi Yönetimi kavramının açıklanması gerekli görülmüş ve kısaca açıklanmıştır.

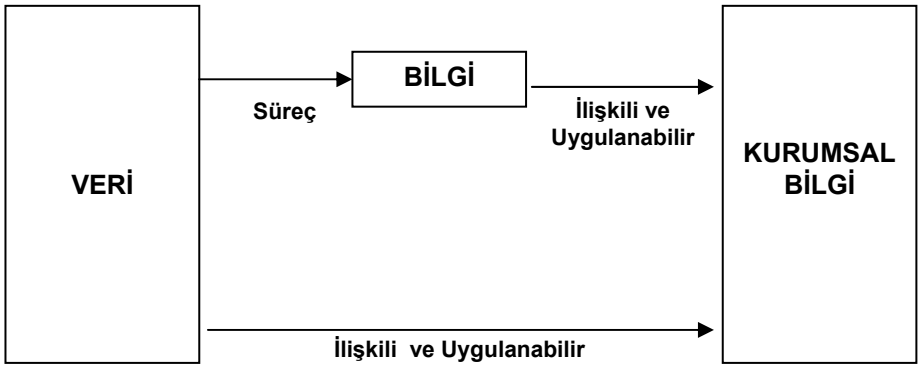
Bilgi Yönetimi: doğru karar vermek için doğru formda, doğru kişiye, doğru maliyetle, doğru zamanda, doğru yerde, doğru bilgiyi sağlamaktır [35].

Bilgi yönetimi ve Bilgi teknolojisi kavramları firmaların ürün odaklı yaklaşımdan müşteri odaklı yaklaşıma geçmesini mümkün kılmıştır. Bu doğrultuda Bilgi Yönetimi'nin hedefi; şirketin, müşterinin ihtiyaç ve beklentilerini hızla algılayabilecek, çözümler oluşturabilecek ve müşteriye değer katabilecek bir bilgi yapısına kavuşmasını sağlamaktır. Ancak, bilginin stratejik değer olabilmesi için, veriden bilgiye dönüşme sürecinde, bilginin doğru, güncel, tam ve anında erişilebilir olması gerekmektedir.

Bilgi Yönetimi konusunda birbiriyle karıştırılan kavramlar aşağıda verilmiştir [34].

- Veri (data): Bir bilgi sistemine girilen, yapısal olmayan, işlenmemiş girdilerin tümüne denir. Bilginin ana kaynağı durumundaki veriler, tek başlarına bir anlam taşımazlar. Bunların bilgi haline gelebilmesi için işlenerek ve diğer verilerle ilişkilendirilerek bir anlam taşıması gereklidir.
- Bilgi (information): Karar alma aşamalarında, verilerin işlenerek kararlara ilişkin faktörlerdeki belirsizlikleri azaltılıp anlamlı hale getirilerek, kullanıcıya sunulmuş halidir. Verilerin belirli sonuçlara ulaşmak üzere işlenmesi ve anlam kazanması sonucu elde edilen bilgiler, yönetim süreçlerinin ve karar alma mekanizmalarının temel girdisi olmaktadır. Bu nedenle bilgi, organizasyonel öğrenmeyi geliştirmektedir.
- Kurumsal Bilgi - Çıkarım (knowledge): Belirli bir amaca yönelik olarak bilginin çeşitli analiz, sınıflama ve gruplama işlemlerinden geçirilerek, gerektiği zamanlarda potansiyel olarak kullanıma hazır hale getirilmesidir. Veri Ambarı (Data Warehouse), Veri Madenciliği (Data Mining), İş Zekâsı (Business Intelligence) vb. uygulamalar bilginin kurumsal bilgiye dönüştürülmesine yönelik uygulamalardır. Kurumsal bilgi işletmenin faaliyet alanındaki bilgisidir. Ayrıca kurumsal bilgi, kullanıldığı zaman tüketilmez, müşteriler başta olmak üzere firmanın ilişkili olduğu tüm taraflar ona katkıda bulunur, bu nedenle değeri artar, zamanla gelişir [34].

Veriden, doğrudan ilişkiler ve uygulanabilirlik elde edilerek kurumsal bilgiye ulaşılabileceği gibi, başta bilgi yapısı kazandıktan sonra kurumsal bilgi de elde edilebilir. Veri, bilgi ve kurumsal bilgi ilişkisi Şekil 2.1'de verilmiştir.

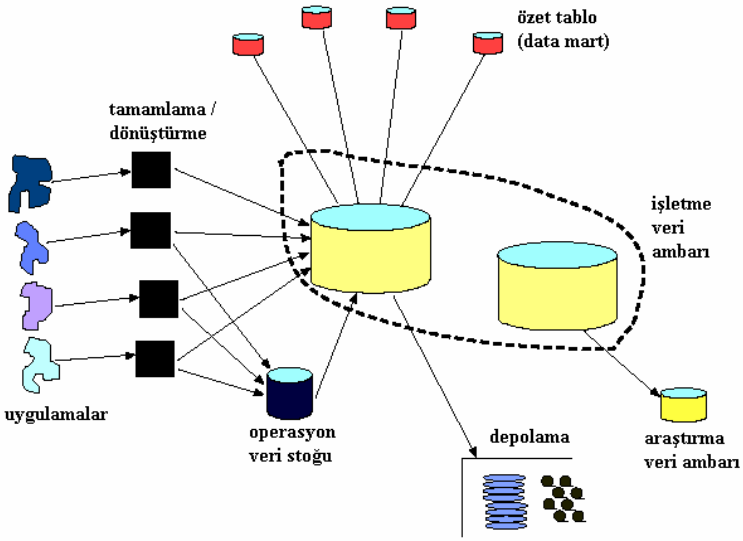


Şekil 2.1: Veri, Bilgi ve Kurumsal Bilgi İlişkisi [34]

2.2. Veri Ambarı

Veri ambarı, yönetimin karar verme sürecini destekleyen, konu tabanlı, tümleşik, zamanla değişen, kalıcı verilerin topluluğudur. Veri ambarının temel hedefi, organizasyon için iş gerçeklerinin tek bir görüntüsünü ortaya koyabilmektir. Bu amaca ulaşmak için, çeşitli programlar aracılığıyla operasyonel ortamdan, veri ambarı verisini barındıran veri tabanından ve kullanıcılara veri sağlayan sistemlerden veriler çıkarılır.

Veri ambarları, stratejik bilgilerin hızlı ve kolay elde edilmesi yoluyla rekabet avantajı sağlamayı hedefleyen kuruluşlar için gittikçe artan bir önem kazanmaktadır. Veri ambarı teknolojisi, hızlı ve etkin iş kararlarının verilmesini desteklemek üzere, verinin yararlı ve güvenilir bilgiye dönüştürülmesi için mükemmel bir yaklaşım sağlamaktadır. Bununla birlikte, bir veri ambarının oluşturulması etkin planlama, kaynak tahsisi, kuruluşa özgü geliştirme yaklaşımı gerektiren karmaşık bir işlemdir. Tipik bir veri ambarı mimarisi Şekil 2.2'de verilmiştir.



Şekil 2.2: Veri Ambarı Mimarisi

Veri ambarlamada en temel parça ilişkili veri tabanıdır. Bu veri tabanı verilerin depolandığı yerdir. İlişkisel veri tabanları yeni verilerin girilmesi ve var olan verilerin standart hale getirilmesini verimli ve etkin bir şekilde gerçekleştirecek şekilde tasarlanmalıdır.

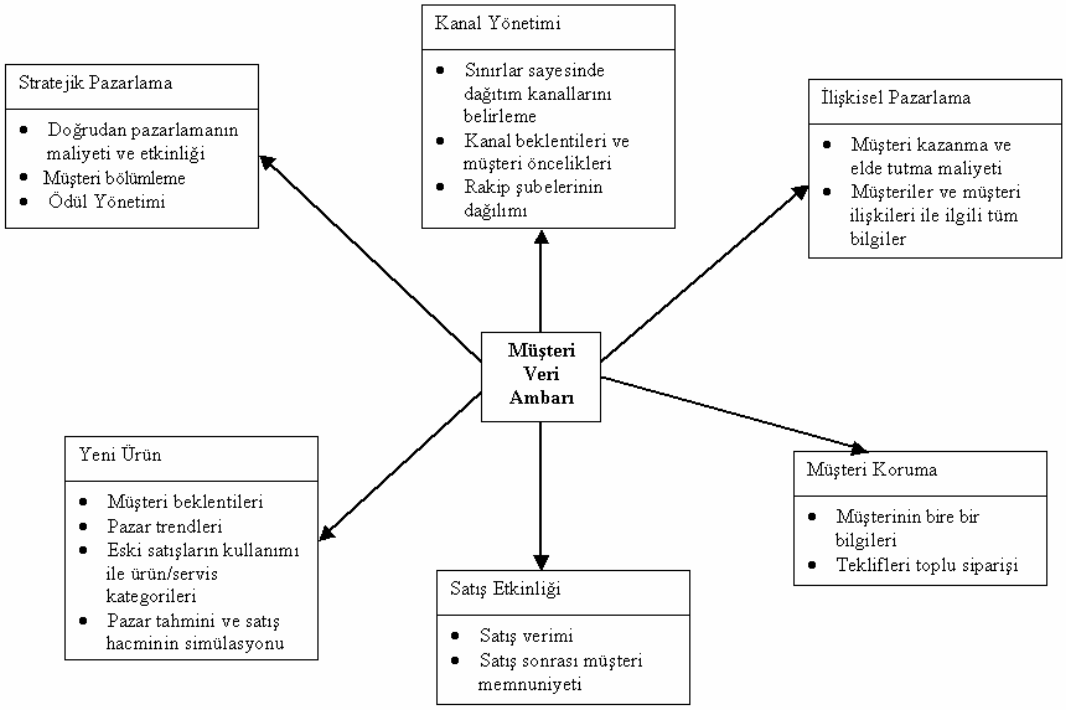
Veri mimarisinin tasarımı, veri ambarlama projesinin belki de en kritik bölümüdür. Bütün verilerin ve farklı parçaların birbirleriyle nasıl ilişkili olduğunun anlaşılmasını içerir. Veri mimarisi tasarımı yapıldıktan sonra, istenilen raporlar elde edilebilir. Bu raporlar kişinin ihtiyacına göre, müşterilerin ödeme bilgileri, işlem zamanları vb. çeşitli verileri kapsayabilir. Veri ambarlamanın gücü firmanın farklı departmanlarının veri birleştirme bağlantılarında ortaya çıkmaktadır [5].

Veri ambarları ve özet tablo yapılarının kullanılması, bilginin iş süreçlerine dönüştürülmesini kolaylaştırmaktadır. Bu ise firma bütçelerinin yıllık %50'ye varan oranlarda iyileşmesini sağlayabilmektedir. Veri ambarı ve özet tablolar çeşitli bilgileri ve heterojen üretim sistemlerini genel veri tabanında tanımlamaya imkân veren yapılardır. Güçlü analitik uygulamalarla bilgiler organize edilebilir, özet bilgiler oluşturulabilir. Böylece de karar vericiler hızlı ve doğru karar verebilirler.

Veri ambarları, karar destek sistemlerinin temeli konumundadır. OLAP (Online Analytical Processing), Veri Madenciliği ve istatistiksel araçlar veri ambarlarından çıkarım yapmak için kullanılırlar. Bu analiz sonuçları ve profillerin belirlenmesi ile promosyon kampanyaları ve uygulama stratejileri belirlenebilmektedir [14].

Veriler, veri ambarına girildiği andan itibaren karar destek sistemi için hazır olarak beklerler. Veri ambarı özet tablolardan verileri toplar ve sahip olduğu verilerle operasyon veri stoklarını besler. Operasyon veri stokları karar destek sistemi ve operasyonel ihtiyaçların karışımı niteliğindedir. Bazı durumlarda özet tablolar veri ambarı konumunda olur. Çünkü bu yapılarda son kullanıcı doğrudan veri ambarını etkiler. Özet tablolar genelde departmanların istek ve gereksinimlerinden doğar ve bu doğrultuda şekillenir.

Veri ambarı, firma verilerinin depolandığı ve istenilen bilgilerin elde edilmesi için kullanılan yapılardır. Bu nedenle veri ambarı yapısının yönetimi oldukça önemlidir. Şekil 2.3'de bir işletmedeki tipik veri ambarı yapısı verilmiştir.



Şekil 2.3: Veri Ambarından Elde Edilen Bilgilerin Kullanım Alanları [34]

Veri ambarı, yeni karar ve sonuçlarla büyüdükçe organizasyon gelecek tahminleriyle ilgili daha doğru kararlar verebilmekte bu da veri ambarının önemini arttırmaktadır. Her ne kadar veri madenciliği için veri ambarı olması zorunlu olan bir yapı değilse de veri madenciliği uygulamalarının verimliliğini arttırmaktadır [3].

3. VERİ MADENCİLİĞİ

Bu bölümde, önemi her geçen gün artmakta olan veri madenciliğinin tanımı, tarihsel gelişimi, kullanım alanları, süreci ve modelleri açıklanmıştır.

3.1. Veri Madenciliğinin Tanımı

Veri madenciliği büyük veri yığınlarından, değişkenlerin ve onlar arasındaki faydalı ilişkilerin çıkarılması sürecidir. Veri tabanı teknolojisi, uzman sistemler, makine öğrenimi, yapay sinir ağları, bilgi çıkarımı, örüntü tanıma, verilerin görselleştirilmesi ve istatistik gibi değişik disiplinler ile yakından ilişkilidir [3], [25].

Diğer bir tanıma göre ise, geniş veri tabanlarında bilgisayar analizleri sayesinde veri tabanı kullanıcılarının ilgilendiği, onlar için değerli olan ancak tahmin edilemeyen ilişkilerin ve kuralların çıkarımı sürecidir [4]. Ayrıca cevabı bilinen bir modeli cevabı bilinmeyen bir modele uygulayarak gerekli sonuçların elde edilmesini sağlayan bir tekniktir [5].

Veri ambarında tutulan çeşitli verilere dayanarak daha önce keşfedilmemiş bilgileri ortaya çıkarmak, bunlara dayalı model ve kurallar oluşturarak karar vermek ve eylem planı gerçekleştirmek veri madenciliğinin kullanılma amacıdır. Veri madenciliği için anahtar kavramlar; "çok fazla" veriden "anlamlı" bilgilerin elde edilmesidir [34].

Veri madenciliği araçları, yorumlanacak örneklerle, veriyi tanımlama ve bazı değişkenleri kullanarak diğer değişkenlerin bilinmeyen değer ve gelecek değerlerinin tespit edilmesini sağlar [7].

Her hangi bir ERP (Enterprise Resourse Planning) sistemi kullanan firma, müşterileri hakkında geniş ve heterojen veri tabanına da sahip olmaktadır. Ancak buradaki problem firmaların çok miktarda veriye sahip olmasına karşın, bilgi bakımından oldukça fakir durumda olmalarıdır. Bu noktada veri madenciliği, verilerden firma faaliyetlerinde kullanabilecekleri stratejik bilgiler elde etmesini sağladığı için firmanın tüm süreçlerine katma değer sağlamaktadır.

Veri madenciliği unsurları, verilerin toplanması, sınıflandırılması, özetlenmesi, grafik ve tablolarla sunulması, analiz edilerek ana kütle hakkında anlamlı bilgiler elde edilmesi ve yorumlar yapılması bakımından istatistikle paralellik göstermektedir. Ancak klasik istatistik ve veri madenciliği bir noktada ayrılmaktadır; gerçek hayatta bir olaya etki eden faktörler birbirleriyle az ya da çok ilişkili olmalarına rağmen klasik istatistik anlayışı bu faktörleri ayrı ayrı ele alarak aralarındaki bu ilişkileri göz ardı etmiş ve gerçekçi olmayan sonuçların elde edilmesine ve sonuç ile ilgili karar verecek kişilerin yanlış yönlendirilmesine neden olmuştur. Buna karşın, ileri istatistik teknikleri ve bilgisayar teknolojisindeki gelişmeler veri madenciliğinin, birçok olaya etki eden çok fazla değişkeni izole etmeden aralarındaki ilişkileri de değerlendirmeye alarak analiz etmesine imkân sağlamıştır [27].

Veri madenciliği kendi başına bir çözüm olmamakla birlikte, çözüme ulaşmak için verilecek karar sürecini destekleyen, problemi çözmek için gerekli olan bilgileri sağlamaya yarayan bir araçtır [28]. Firmalara, veri ambarlarındaki en önemli bilgilere odaklanmalarını sağlaması nedeniyle oldukça büyük katkılar sağlamaktadır.

Veri madenciliği teknoloji ile organizasyonda güçlü tahminsel modeller oluşturulabilir. Daha doğru tahminler ile iş kararları optimum hale getirilebilir, hizmetler iyileştirilebilir, fazla stoklamanın önüne geçilebilir, gelirler arttırılabilir, süreçlerin verimlilikleri maksimum seviyeye getirilebilir ve maliyetler düşürülebilir. Ayrıca müşterilerle yakın ilişkiler kurulmasında kullanıldıkları için eldeki müşterilerden daha fazla kazanç sağlanırken yeni müşterilerin kazanımında ve müşterilerin satın aldıkları mal ya da hizmetlerden daha fazla memnun kalmalarında da büyük katkılar sağlamaktadır.

3.2. Veri Madenciliğinin Tarihsel Gelişimi

1960'larda veri toplama sistemleri oldukça yaygın kullanılmasına karşın verilerin kullanımı ve faydalı sonuçların elde edilmesi sınırlı seviyede kalmıştır. Firmalar faaliyetleri ile ilgili verileri faydalı olacağı düşüncesiyle toplamış ve sadece geçmişle ilgili bilgilere ulaşabilmişlerdir.

1970'lerde ilişkisel veri tabanı modelleri kullanılmaya ve veriler arasında bağıntılar araştırılmaya başlanmış ve bu yapı 1980'lerde yaygın hale gelmiştir. Geleneksel sorguların varsayıma dayalı olması kullanıcıların doğru soru sormasını zorunlu kılmıştır.

1990'larda ve 2000'li yıllarda ise veri madenciliği, veri ambarları, çoklu ortam (mültimedya) ve web veri tabanları kullanımı artmıştır. Bu gelişim sırasında yöntemler çok farklı olmamasına rağmen fark yaratan unsur bilgisayar ve bilişim teknolojilerindeki hızlı gelişmenin çok miktardaki veriden analiz yaparak işletme operasyonları için faydalı bilgilerin elde edilmesine imkân sağlamasıdır. Tablo 3.1'de bu evrim ve her döneme ait özellikler verilmiştir.

Tablo 3.1: Veri Madenciliği Evrimsel Adımları [5]

Evrimsel Adım	**İş Sorusu**	**İmkân Veren Teknolojiler**	**Karakteristikler**
Veri Toplama (1960'lar)	"Son beş yıldaki toplam gelirim ne kadardır?"	Bilgisayarlar, Teypler, Diskler	Geçmişle ilgili statik veri sağlama
Veri Girişi (1980'ler)	"Geçen Ekim ayında Türkiye'de birim satış miktarı ne kadardır?"	İlişkisel Veritabanları, Yapısal Sorgu Dilleri	Geçmişle ilgili, kayıt düzeyinden dinamik veri sağlama
Veri Ambarlama ve Karar Destek (1990'lar)	"Geçen Ekim ayında Türkiye'de birim satış miktarı ne kadardır?"	On-line Analitik Proses (OLAP), Çok boyutlu Veri Tabanları, Veri Ambarlama	Geçmişle ilgili, çok çeşitli düzeylerden dinamik verilerin elde edilmesi
Veri Madenciliği (Bugün)	"Türkiye'de gelecek ay olası birim satış miktarı ne olacaktır? Neden?	İleri Algoritmalar, Çok işlemcili bil., çok büyük veri tabanları	Geçmişle ilgili, ilişkisel bilgi sağlama

Son zamanlarda kullanımı artan, makale ve kitaplara konu olan veri madenciliğinin bu aşamalara gelmesi uzun zaman almış ve çeşitli evrelerden geçmiştir. Bu evrelerdeki teknikler aşağıda açıklanmıştır;

- Veri madenciliği istatistik üzerine kurulmuştur. İstatistik yöntemler, veri ve veriler arasındaki ilişkileri ortaya koyar ki bu ileri istatistik analizlere imkân sağlar.
- Yapay zekâ, istatistiğin aksine höristik temele inşa edilmiş ve istatistik problemlere insan düşüncesini katmayı denemiştir. Ancak bu yüksek işleme kapasitesi gerektirdiğinden ve 1980'lerin başına kadar bu tür bilgisayarların maliyetlerinin yüksek olması nedenleriyle mümkün olmamıştır.

- Makine öğrenimi, istatistik ve yapay zekânın birleşimi olarak tanımlanabilir. Yapay zekânın ticari başarısı yokken, teknikleri makine öğreniminde kullanılmış ve başlangıç maliyeti yapay zekâdan daha az olduğu için kullanım avantajı sağlamıştır. Makine öğrenimi çalıştığı verilerden öğrenerek veri kalitesine bağlı olarak farklı kararlar oluşturabilmektedir.

3.3. Veri Madenciliği Uygulama Alanları

Veri madenciliği teknikleri her geçen gün önemi artan ve her alanda kullanım imkânı bulan ileri istatistik teknikleri sunmaktadır. Özellikle çok fazla verinin işlenmesinin gerekli olduğu durumlarda kullanımları artmaktadır. Aşağıda veri madenciliği için bazı uygulama alanları ve bu alanlardaki özel uygulamalar örnek olarak verilmiştir.

Perakende

- Sepet Analizi: Müşterilerin hangi ürünleri beraber almaya eğilimli olduklarını ifade eder. Bu bilgi stoklamaya, stok yönetim stratejisine ve kampanya faaliyetlerine bilgi desteği sağlar.
- Satış Tahmini: Zaman tabanlı ilişkiler stok yönetimine girdi sağlar. Müşterinin bugün bir ürün alması sonucu, tamamlayıcı ürünü ne zaman alacağı tahmin edilebilir.
- Veri Tabanlı Pazarlama: Müşteri davranışlarına göre müşteri profilleri belirlenebilir. Bu ise reklam faaliyetlerinin etkin yapılmasını sağlar.

Bankacılık

- Kredi Kartı Pazarlama: Müşteri bölümlerinin belirlenmesiyle, daha etkin müşteri kazanımı ve elde tutma politikaları oluşturulur. Hedef ürün geliştirilir ve özelleşmiş fiyatlama yapılabilir.
- Kart Tutma Maliyeti ve Fiyatlama: Kârı maksimum yaparken müşteri kaybını minimum yapacak risk tabanlı fiyatlama yapılabilir.
- Kredi Batığı: Kredinin batmasının çok maliyetli olması nedeniyle eski verilerden yararlanarak bu süreç etkin yönetilebilir.
- Tahmine Dayalı Yaşam Çevrimi Yönetimi: Her müşterinin yaşam değerinin belirlenmesi ve her bölüme uygun servisin sağlanması

Telekomünikasyon

- Arama Detayları Kayıt Analizi: Telekomünikasyon firmalarının detaylı arama kayıtlarına sahip olmaları nedeniyle benzer kullanım alışkanlıklarına sahip olanlar (müşteri bölümleri) belirlenerek etkin fiyatlama ve uygun kampanyalar düzenlenebilir.
- Müşteri Sadakati: Bazı müşteriler firma değiştirerek avantaj elde ederler. Veri madenciliği ile sadık olması muhtemel müşteriler belirlenerek (aynı zamanda kârlı) onlara yatırım yapılabilir.

Diğer Uygulamalar

- Müşteri Bölümleme: Tüm sektörlerdeki firmalar ek değişkenler düşünerek farklı müşteri bölümleri belirleyerek avantaj kazanabilirler.
- Üretim: Üreticiler artık müşteriye özelleşmiş üretim yapmak durumunda olup veri madenciliği uygulamaları müşteri taleplerinin ne yönde olduğunun belirlenmesinde yardımcı olmaktadır.
- Garanti: Üreticiler olması muhtemel şikâyet oranını ve bunun maliyetinin ne olacağını belirleme imkânı bulurlar [29].

3.4. Veri Madenciliği Süreci

Veri madenciliği algoritmasının başarısı inceleme yapılan işin ve verilerin özelliklerinin bilinmesine, anlaşılmasına ve aşağıda yer alan veri tabanında bilgi keşfi süreci adımlarının etkin bir şekilde gerçekleştirilmesine bağlıdır.

3.4.1. Problemin Tanımlanması

Uygulamanın hangi işletme amacı için yapılacağının açık bir şekilde tanımlanmasıdır. İlgili işletme amacı işletme problemi üzerine odaklanmış ve açık bir dille ifade edilmiş olmalı, elde edilecek sonuçların başarı düzeylerinin nasıl ölçüleceği tanımlanmalıdır. Ayrıca yanlış tahminlerde katlanılacak olan maliyetlere ve doğru tahminlerde kazanılacak faydalara ilişkin tahminlere de bu aşamada yer verilmelidir. Problemin yanlış belirlenmesi yapılan analizlerin de yanlış ve faydalı olmayan sonuçlara ulaşmasına neden olur.

3.4.2. Verilerin Hazırlanması

Kullanılacak verilerin kalitesi sonuçları da etkileyeceğinden kullanılacak verilerin ön işlemden geçirilmesi çok önemlidir. Zira modelin kurulması aşamasında ortaya çıkacak sorunlar, bu aşamaya sık sık geri dönülmesine ve verilerin yeniden düzenlenmesine neden olacaktır. Bu durum verilerin hazırlanması ve modelin kurulması aşamaları için, bir analistin veri keşfi sürecinin toplamı içerisinde enerji ve zamanının % 50 - % 85'ini harcamasına neden olmaktadır. Verilerin hazırlanması 5 bölümde ele alınabilir [1].

- Toplama: Tanımlanan problem için gerekli olduğu düşünülen verilerin ve bu verilerin toplanacağı veri kaynaklarının belirlenmesi adımıdır. Kullanılacak veriler kuruluşun kendi iç süreçlerinden elde edilebileceği gibi, harici veri tabanlarından hatta veri pazarlayan kuruluşlardan elde edilebilir.
- Değer Biçme: Veri madenciliğinde kullanılacak verilerin farklı kaynaklardan toplanması, veri uyumsuzluklarına neden olabilmektedir. Bu uyumsuzluklar: farklı zamanlara ait olmaları, kodlama farklılıkları ve farklı ölçü birimleridir. Ayrıca verilerin nasıl, nerede ve hangi koşullar altında toplandığı da önem taşımaktadır. Bu nedenlerle, toplanan verilerin ne ölçüde uyumlu oldukları bu adımda incelenerek değerlendirilmelidir.
- Birleştirme ve Temizleme: Bu adımda farklı kaynaklardan toplanan verilerde bulunan ve bir önceki adımda belirlenen sorunlar mümkün olduğu ölçüde giderilerek veriler tek bir veri tabanında toplanmalıdır.
- Seçim: Kurulacak modele bağlı olarak veri seçimi yapılır. Örneğin tahmin edici bir model için, bu adım bağımlı ve bağımsız değişkenlerin ve modelin eğitiminde kullanılacak veri kümesinin seçilmesi anlamını taşımaktadır. Anlamlı olmayan ve diğer değişkenlerin modeldeki ağırlığının azalmasına da neden olabilecek değişkenlerin modele girmemesi gerekmektedir. Değişken sayısı, her değişken arasındaki ilişkinin daha iyi analiz edilebileceği ve daha iyi yönetilebileceği şekilde faktör analizi vb. tekniklerle azaltılabilir [7]. Bazı veri madenciliği algoritmaları konu ile ilgisi olmayan bu tip değişkenleri otomatik olarak elese de, uygulamada bu işlemin kullanılan yazılıma bırakılmaması daha akılcı olacaktır. Verilerin görselleştirilmesine olanak sağlayan grafik araçlar ve bunların sunduğu ilişkiler, bağımsız değişkenlerin seçilmesinde önemli yararlar sağlayabilmektedir.

Genellikle yanlış veri girişinden veya bir kereye özgü bir olayın gerçekleşmesinden kaynaklanan verilerin önemli bir uyarıcı bilgi içerip içermediği kontrol edildikten sonra veri yığınından atılması tercih edilir.

- Dönüştürme: Modelde iki veya daha fazla veriden türetilmiş değişkenler yerine ham veriler kullanılmalıdır. Ayrıca modelde kullanılan algoritma, verilerin gösteriminde önemli rol oynayacaktır. Örneğin bir uygulamada bir yapay sinir ağı algoritmasının kullanılması durumunda kategorik değişken değerlerinin evet/hayır olması; bir karar ağacı algoritmasının kullanılması durumunda ise örneğin gelir değişken değerlerinin yüksek/orta/düşük olarak gruplanmış olması modelin etkinliğini arttıracaktır.

3.4.3. Modelin Kurulması ve Değerlendirilmesi

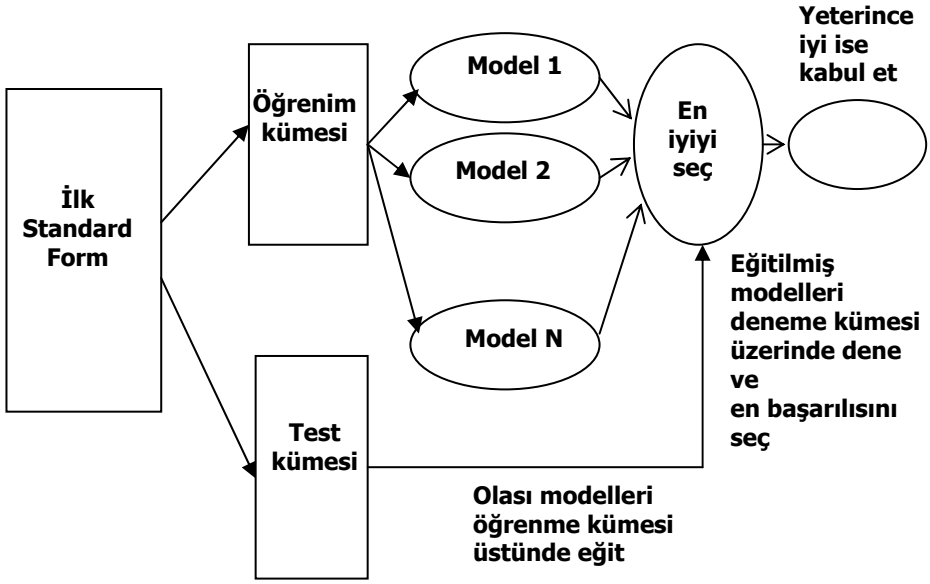
Tanımlanan problem için en uygun modelin bulunabilmesi, olabildiğince çok sayıda modelin kurularak denenmesi ile mümkündür. Bu nedenle veri hazırlama ve model kurma aşamaları, en iyi olduğu düşünülen modele varılıncaya kadar yinelenen bir süreçtir. Bu süreç Şekil 3.1'de verilmiştir. Model kuruluş süreci denetimli (Supervised) ve denetimsiz (Unsupervised) öğrenimin kullanıldığı modellere göre farklılık göstermektedir.

Model kurmanın çeşitli adımları aşağıda verilmiştir.

- Denetimli Öğrenme; Bir denetçi tarafından ilgili sınıflar önceden belirlenen bir kritere göre ayrılarak, her sınıf için çeşitli örnekler verilir. Sistemin amacı verilen örneklerden hareket ederek her bir sınıfa ilişkin özelliklerin bulunması ve bu özelliklerin kural cümleleri ile ifade edilmesidir. Öğrenme süreci tamamlandığında, tanımlanan kural cümleleri verilen yeni örneklere uygulanır ve yeni örneklerin hangi sınıfa ait olduğu kurulan model tarafından belirlenir.

- Denetimsiz öğrenme; kümeleme analizinde olduğu gibi ilgili örneklerin gözlenmesi ve bu örneklerin özellikleri arasındaki benzerliklerden hareket ederek sınıfların tanımlanması amaçlanmaktadır.

Denetimli öğrenimde seçilen algoritmaya uygun olarak ilgili veriler hazırlandıktan sonra, ilk aşamada verinin bir kısmı modelin öğrenimi, diğer kısmı ise modelin geçerliliğinin test edilmesi için ayrılır. Modelin öğrenimi, öğrenim kümesi kullanılarak gerçekleştirildikten sonra, test kümesi ile modelin doğruluk derecesi belirlenir.



Şekil 3.1: Veri Madenciliği Modelleme Süreci

Bir modelin doğruluğunun test edilmesinde kullanılan yöntemler:

- Basit geçerlilik (Simple validation) testi: Verilerin %5 ila %33'ü test verileri olarak ayrılır ve kalan kısım üzerinde modelin öğrenimi gerçekleştirildikten sonra, bu veriler üzerinde test işlemi yapılır. Bir sınıflama modelinde yanlış olarak sınıflanan olay sayısının, tüm olay sayısına bölünmesi ile hata oranı, doğru olarak sınıflanan olay sayısının tüm olay sayısına bölünmesi ile ise doğruluk oranı hesaplanır.
- Çapraz geçerlilik (Cross validation) testi: Sınırlı miktarda veriye sahip olunması durumunda, veri kümesi tesadüfî olarak iki eşit parçaya ayrılır. İlk aşamada "a" parçası üzerinde model eğitimi ve "b" parçası üzerinde test işlemi; ikinci aşamada ise "b" parçası üzerinde model eğitimi ve "a" parçası üzerinde test işlemi yapılarak elde edilen hata oranlarının ortalaması kullanılır.

- "N" Katlı çapraz geçerlilik (N-Fold cross validation) testi*:* Sınırlı miktarda verinin "n" gruba ayrıldığı ve ilk aşamada birinci grubun test, diğer grupların öğrenim için kullanıldığı ve bu sürecin tüm "n" grup için tekrarlandığı testtir. Sonuçta elde edilen "n" sayıda hatanın ortalaması, modelin hata oranını vermektedir.

Model kuruluşu çalışmalarının sonucuna bağlı olarak, aynı teknikle farklı parametrelerin kullanıldığı veya başka algoritma ve araçların denendiği değişik modeller kurulabilmektedir. Farklı modeller kurarak, doğruluk derecelerine göre en uygun modeli bulmak üzere sayısız deneme yapılmasında yarar bulunmaktadır.

3.4.4. Modelin Kullanılması ve İzlenmesi

Kurulan ve geçerliliği kabul edilen model doğrudan bir uygulama olabileceği gibi, bir başka uygulamanın alt parçası olarak kullanılabilir. Zaman içerisinde bütün sistemlerin özelliklerinde ve dolayısıyla ürettikleri verilerde ortaya çıkan değişiklikler, kurulan modellerin sürekli olarak izlenmesini ve gerekiyorsa yeniden düzenlenmesini gerektirecektir. Tahmin edilen ve gözlenen değişkenler arasındaki farklılığı gösteren grafikler model sonuçlarının izlenmesinde kullanılan yararlı bir yöntemdir [1].

3.5. Veri Madenciliği Modelleri

Veri madenciliğinde kullanılan modeller, tahmin edici (Predictive) ve tanımlayıcı (Descriptive) olmak üzere iki ana başlık altında incelenmektedir.

- Tahmin Edici (Predictive) Modeller: Sonuçları bilinen verilerden hareket edilerek bir model geliştirilmesi ve kurulan bu modelden yararlanılarak sonuçları bilinmeyen veri kümeleri için sonuç değerlerinin tahmin edilmesi için kullanılırlar. Modelde bağımsız değişkenler bağımlı değişkenin tahmin edilmesi için kullanılır.
- Tanımlayıcı (Descriptive*)* Modeller: Veri setindeki bağıntı ve profilleri belirleyen ve yorumu kullanıcıya bırakan modellerdir. Dolayısıyla karar vermeye rehberlik etmede kullanılabilecek mevcut verilerdeki bağıntıların tanımlanması sağlanmaktadır.

Tanımlayıcı model, verilerin arasındaki ilişkilerin bulunması, tahmin edici modeller ise bulunan bu ilişkilerin açıklanmasını kapsamaktadır [26].

Veri madenciliği modelleri işlevlerine göre üç başlık altında incelenir.

- Sınıflama ve Regresyon,
- Birliktelik Kuralları ve Ardışık Zamanlı Örüntüler,
- Kümeleme,

Sınıflama ve Regresyon modelleri tahmin edici, kümeleme, birliktelik kuralları ve ardışık zamanlı örüntü modelleri tanımlayıcı modellerdir.

3.5.1. Sınıflama ve Regresyon Modelleri

Sınıflandırma ve regresyon, önemli veri sınıflarını ortaya koyan veya gelecek veri eğilimlerini tahmin eden modelleri kurabilen veri analiz modelleridir. Mevcut verilerden hareket ederek geleceğin tahmin edilmesinde faydalanılan ve veri madenciliği teknikleri içerisinde en yaygın kullanıma sahip olan sınıflama ve regresyon modelleri arasındaki temel fark, tahmin edilen bağımlı değişkenin kategorik (kesikli) veya süreklilik gösteren bir değere sahip olmasıdır. Kesikli değerler sınıflandırma, sürekli değerler ise regresyon yöntemleriyle kullanılırlar. Sınıflandırma iki adımdan oluşur; eğitim veri setinden model kurulması ve bu modelin sınıfları bilinmeyen diğer veri setinde uygulanmasıdır. Sınıflandırma ve regresyon modellerinde kullanılabilen başlıca teknikler aşağıda açıklanmıştır.

3.5.1.1. Karar Ağaçları (Decision Trees)

Karar ağacı, karar düğümü, dal ve yaprak yapılarından oluşan ağaca benzer bir yapıdır. Kategorik değerli hedef fonksiyonlar için görsel olarak bazı kurallar, şartlar tanımlayarak, kararın verilmesine imkân sağlamaktadır. Karar ağaçları birçok test uygular ve hedefin tahmini için en uygun sıraya ulaşmayı dener. Her test, yaprak düğümüne ulaşana kadar daha fazla test yapmayı sağlayan dalları oluşturur. Kökten hedef yaprağa doğru giden yol, kuralı ifade eder ve bu hedefi sınıflandırır.

Kuruluşlarının ucuz, yorumlanmalarının kolay olması, veri tabanları sistemleri ile kolayca bütünleştirilmeleri ve güvenirliklerinin iyi olması nedenleriyle yaygın kullanılmaktadırlar. Karar ağacı, karar vericiye hangi faktörlerin önemli olduğunu ve bu faktörlerin kararın farklı çıktılarıyla nasıl ilişkilendirileceğini anlamasında yardımcı olur [7].

3.5.1.2. Yapay Sinir Ağları (Artificial Neural Networks)

Yapay sinir ağları sınıflandırma ve regresyon algoritmalarının en karmaşık olanıdır. Genellikle kara kutu olarak düşünülür. Çünkü ilişkilerin bulunmasını açıklayamaz. Sinir ağı eğitimi için çok sayıda veri gerektirir, bu da zaman kaybına neden olur. Ancak bir kere eğitildikten sonra yeni olaylar için tahminleri çok hızlı, neredeyse gerçek zamanlı gerçekleştirir. Sinir ağları doğrudan sayılarla işlem yaptığı için bağımlı ya da bağımsız tüm nümerik olmayan veriler sayılara dönüştürülmelidir. Bu sistemin sinir ağları olarak nitelendirilmesinin nedeni, çalışma mekanizmasının insan beynine benzemesidir.

3.5.1.3. K-En Yakın Komşu (K-Nearest Neighbour)

Bir nesnenin niteliklerini inceleyerek bir sınıf veya gruba yerleştirerek ona en yakın olan nitelikleri aynı gruba dâhil etme prensibine göre çalışır. K, olay ya da gruptaki benzer unsurları ifade etmekte olup, rakamsal verilerdeki ilişkileri ve dizileri tespit etmek için kullanılır. Diğer tahmin modellerinden farklı olarak eğitim seti modelin kendisidir. K-NN, uzaklık kavramına bağlıdır ve uzaklıkların belirlenmesi için ölçü gerektirir. Uygun ölçünün seçimi en hassas noktadır, çünkü eğitim verisinde kullanılan farklı ölçüler tamamen farklı tahminlere neden olabilir [7].

K-NN algoritmasında iki temel parametre vardır:

- Kullanılan en yakın olayın sayısı
- Benzerliği ölçecek bir ölçü

Bu teknolojilerin çoğu uzun zamandır kullanımda olmakla beraber küçük veri hacimleri için uygulanmışlardır. Ancak, veri ambarları ve OLAP uygulamaları ile analiz edilen veri hacimleri arttırılmıştır [5].

3.5.2. Birliktelik Kuralları ve Ardışık Zamanlı Örüntüler

Büyük veri kümeleri arasında birliktelik ilişkileri bulunması için kullanılırlar. En tipik örnek sepet analizidir. Bu yöntemle müşterilerin yaptıkları alışverişlerdeki ürünler arasındaki birliktelikleri bularak müşterilerin satın alma alışkanlıkları analiz edilir.

- Hangi ürün veya servisler aynı zamanda satın alınma eğilimindedir?
- Daha sonra takip eden tamamlayıcı satışlar nelerdir?

Bu teknikler perakende sektörü kaynaklı olmakla birlikte hedef pazarlama kampanyaları veya genel uygulamaların oluşturulmasında eşit şekilde uygulanabilirler. Finans sektöründe de müşteri hesap portföylerinin analiz edilmesinde, müşterilerin beraber aldıkları finansal setlerin neler olduğunun tanımlanmasında, satış kampanyalarında ve tıp alanında kullanılabilirler. Genel olarak farklı olayların birbirleri ile ilişkili olduğunun belirlenmesi sonucunda değerli bilgi kazanımının söz konusu olduğu ortamlarda da önem taşımaktadırlar [7].

Ardışık zamanlı örüntüler ise, birbirleri ile ilişkisi olan ancak birbirini izleyen dönemlerde gerçekleşen ilişkilerin tanımlanmasında kullanılırlar [1].

3.5.3. Kümeleme Modelleri

İlk defa 1939 yılında Tryon [15] tarafından kullanılan Kümeleme Analizi; belirli kategoriye göre nesnelerin gruplanmasına imkân veren farklı algoritmaları içeren çok değişkenli istatistik analiz tekniğidir. Farklı nesnelerin ikili olarak iyi tanımlanmış benzerlik kurallarına bağlı olarak analiz edilmesi ve benzerlik derecesi yüksek olması durumunda aynı, düşük olması durumunda ise farklı gruplara atanması esasına dayanır [15] , [30]. Verilen veri setinde n adet veri (örnek) m boyut (nitelik) uyarınca k adet kümeye ayrılır. Bu kümeleri belirlemede anahtar rol “benzerlik”tir [30]. Analiz sonucunda elde edilen kümelerin kendi içlerinde homojen, birbirinden farklı kümelerin ise heterojen olması amaçlanır [27]. Dolayısıyla algoritmalar gruplar arası varyansı maksimum kılmaya, aynı grup içindeki varyansı da minimum yapmaya çalışır [23].

Kaufman ve Rousseeuw (1990) Kümeleme Analizi'ni "veri içinden grupların bulunması sanatı" olarak tanımlarlar. Bu grupların bulunmasında ise benzerlik (farklılık) veya nesneler arasındaki mesafeler grupların ayrılması veya birleştirilmesi için kuralların çıkarımı için kullanılır [15].

Kümeleme analizi, veri madenciliğinin temeli niteliğindedir ve veri tabanlarında toplanan veri miktarının artması ile veri madenciliği araştırmalarında aktif bir konu haline gelmiştir. Diğer istatistiksel analizlerin aksine başlangıçta hipotez olmaması durumunda kullanılır ve olması muhtemel en belirleyici sonuca ulaşır [15]. Kümeleme analizi ile küçük homojen grupların oluşması ile veri yapısını ortaya koyması ve büyük heterojen gruplarla araştırma yapmak zorunluluğunu ortadan kaldırması nedeniyle [38] mümkün olan en anlamlı çözümlere daha kısa sürede ulaşmayı sağlar. Ayrıca eğitim seti modelin kendisidir.

Belirli bir topluluktan grupların tanınması esasına dayanan kümeleme analizi olmaksızın bölümleme ve hedef pazarlama yapılması mümkün değildir [13]. Kümeleme analizi, birçok alanda uygulanabilen en etkili ve kolay yorumlanabilen bir yöntemdir. Pazarlama, coğrafik uygulamalar, bilgi sistemleri, biyoloji, genetik alanlarında, sosyal bilimler, antropoloji, karakter tespiti [9] ilaç sektörü, hastalıkların tedavisi, hastalık belirtilerinin kümelenmesi, belirtilerden hastalıkların ve tedavilerinin kümelenmesi, arkeolojide eşyaların ait olduğu dönemlerin tespiti vb. alan ve konularda uygulanmaktadır. Genel olarak ise çok fazla verinin anlamlı bilgilere dönüştürülmesinde kullanılır [15]. Kısaca diğer veri madenciliği tekniklerinde olduğu gibi bilgi yığınından anlamlı çıkarımlar yapılması gerektiğinde kullanılmaktadır. Etkili ve kolay yorumlanabilen bir teknik olması kullanımını daha da yaygınlaştırmaktadır.

Kümeleme analizi kullanım alanı oldukça geniş bir tekniktir. Özellikle firmaların hedef pazar ve müşterilerini belirlemek için pazar ve müşteri bölümlendirme süreçlerinde, müşterilerini kullandıkları ürün, ürün miktarı, sadakat düzeyleri bakımından tanımlamada, bayi düzeni ve satış performanslarının izlenmesinde kullanılabilmektedir.

Kümeleme analizinde, tek üye seçiminden ziyade grup seçimi söz konusu olduğundan zaman kaybı önlenmekte ve tüm özellikler bakımından toplu değerlendirmeye imkân sağladığından diğer yöntemlere üstünlük sağlamaktadır [23].

3.5.3.1. Hiyerarşik Kümeleme Teknikleri

Başlangıçta tüm örnekleri tek kümede toplayan ve ilerleyen adımlarda örnekleri kümelere bölen bir metottur. Bu yöntem kaç küme olacağının başlangıçta bilinmesini gerektirmez ki bu hiyerarşik olmayan yönteme göre avantajdır [30]. Sınıflama n tane gözlem yani n tane kümenin belirli aşamalarla kümelenmesi ve tek bir kümede toplanması şeklinde gerçekleşmektedir.

Hiyerarşik kümeleme teknikleri aralarında birleştirici kümeleme teknikleri ve ayırıcı kümeleme teknikleri olmak üzere ikiye ayrılmaktadır. Birleştirici kümeleme teknikleri, n tane gözlemi tek bir kümede toplarken; ayırıcı teknikler n gözlemli bir kümeyi n tane ferdi birime ayıran tekniklerdir. Hangi hiyerarşik teknik kullanılırsa kullanılsın, birleştirilmiş ve bir küme haline getirilmiş iki birim bir daha ayrılamaz veya birbirinden ayrılan iki birim tekrar birleştirilemez. Bu hiyerarşik kümeleme tekniklerinin bir dezavantajıdır. Diğer dezavantajı ise küme sayısının başlangıçta bilinmemesidir. Bu nedenle hangi aşamada durulacağına karar verilmesi gerekmektedir. Kümeleme, dendrogram adı verilen grafikler aracılığıyla gösterilmektedir.

Hiyerarşik kümeleme adımları;

1. n tane gözlem, n tane küme olmak üzere işleme başlanır.
2. En yakın iki küme birleştirilir.
3. Küme sayısı bir indirgenerek yinelenmiş uzaklıklar matrisi bulunur.
4. 2 ve 3 nolu adımlar (n-1) kez tekrarlanır.

Algoritma ve birimler arasındaki yakınlığı ölçen yöntemlere dayalı altı farklı hiyerarşik kümeleme tekniği vardır. Bu yöntemler gözlemler arasındaki uzaklık veya benzerliğin hesaplanması şekline göre farklılık göstermektedir.

- Tek Bağlantı Yöntemi (single linkage method): Birbirine en yakın iki gözlemi bulur ve bir küme oluşturur. Daha sonra birbirine en yakın iki başka gözlemi ya da ilk kümeye en yakın başka bir gözlemi bulur ve kümeye birleştirir.
- Tam Bağlantı Yöntemi (complete linkage method): Birimler arasında en uzak değerler esas alındığından tek bağlantı yönteminin tersi şeklindedir.

- Ortalama Bağlantı Yöntemi (average linkage method): İki küme arasındaki uzaklık, iki küme oluşturacak her bir birim çiftleri arasındaki uzaklığın ortalaması olarak hesaplanmaktadır. SPSS programında ortalamalar gruplariçi ve gruplararası olmak üzere iki şekilde hesaplanır.

Yukarıda bahsedilen üç bağlantı yönteminde hesaplamalar doğrudan uzaklık matrisi üzerinden yapılmakta, verilerin gerçek değerlerine ihtiyaç duyulmamaktadır.

- Merkezi Yöntem (centroid method): Elde edilen gruplar, her bir değişken için hesaplanan ortalama değerlerine göre kümelenmektedir.
- Medyan Yöntemi (median method): Birleştirilecek olan iki grubun içerdiği gözlem sayısı birbirinden çok farklı olduğu durumlarda, hesaplanan merkez noktası birim sayısı fazla olan gruba yakın olacağından medyan yöntemi birleştirilecek grupların aynı sayıda birime sahip olduğunu varsayarak bu işlemi yapar.
- Ward Yöntemi: Bir kümenin ortasına düşen gözlemin, aynı kümenin içinde bulunan gözlemlerden ortalama uzaklığını esas alır. Diğer kümeleme yöntemlerinin aksine kümeler arasındaki uzaklıklar hesaplanmamaktadır. Bunun yerine, kümeler içi, homojenliği maksimum yapacak kümeler oluşturulmaya çalışılmaktadır. Yani, kümeler içi hata kareleri toplamı minimum yapılmaya çalışılmaktadır.

3.5.3.2. Hiyerarşik Olmayan Kümeleme Teknikleri

Küme sayısı konusunda ön bilgi var ise ya da küme sayısına önceden karar verilmiş ise, zaman alan hiyerarşik yöntemler yerine, hiyerarşik olmayan teknikler tercih edilmektedir. Hiyerarşik olmayan kümeleme tekniklerinde aşağıdaki temel adımlar izlenmelidir:

1. k tane küme merkezi seçilir. Burada k oluşması istenen küme sayısıdır.
2. Küme merkezine en yakın olan her bir gözlem en yakın olduğu kümeyle birleştirilir.

3. Belirli bir durma noktasına gelinceye kadar (gözlemler k tane küme oluşturuncaya kadar) birleştirme işlemine devam edilir.
4. Belirli durma kriterine göre k tane küme oluşturulmuş olur ya da 2. aşamaya geri dönülür [10].

4. MÜŞTERİ İLİŞKİLERİ YÖNETİMİ

Tedarik zinciri; tedarikçiler, üreticiler, dağıtıcılar, toptancılar, perakendeciler ve müşterilerden oluşan bir yapıyı ifade etmektedir. Tedarik zinciri yönetimi ise; ham madde temininden nihai müşteriye ulaşıncaya kadar tüm süreçleri içine alan ve söz konusu yapıda mal, bilgi, para akışını etkin bir şekilde yönetirken tüm süreçlerin uyum içinde çalışmasını temin edecek yönetim sistemidir.

Tedarik zinciri yönetim süreçleri kapsamında önemli bir yere sahip olan müşteri ilişkileri yönetimi (MİY) ise, tüm süreçleri müşteri odak noktası etrafında, iş stratejilerinin kârlılığını, geliri ve müşteri tatminini optimize etmeye yönelik tasarlanmış ve müşterilerden tedarikçilere kadar tüm tarafları kapsayan yönetim stratejisidir [11].

Müşteri ilişkileri yönetimi, müşteri ile sürekli ve bireysel ilişkiyi gerektiren ve bu nedenle de tüm müşteriler hakkında en ince ayrıntısına kadar bilgi elde etmeyi zorunlu kılan, bu bilgiler bazında da bire bir pazarlama ve bire bir ilişkiye zemin oluşturan bir stratejik yaklaşımdır. Bu stratejinin özü; müşteriyi sürekli öğrenen bir veri toplama, bu verileri depolama, filtre etme, derinlemesine inceleyip ilişkiler yakalama ve tahminler oluşturmaktır [18]. Burada en önemli nokta, müşterileri doğru analiz edebilmek, eğilimleri konusunda doğru sonuçlara varmak ve tüm bunları ortaya koyduktan sonra onlara ulaşmaya çalışmaktır. Temel mekanizma, müşterilerin nasıl farklılıklar gösterdiğinin anlaşılmasını ve bu farklılıkların her müşteriye göre firmanın nasıl davranması gerektiği konusunda bir strateji oluşturulmasını içermektedir [18].

MİY, karar destek sistemleri için, müşteri bilgilerini inceleyerek, ilişkiler bularak ve geleceğe yönelik tahminler yaparak pazarlama fonksiyonunu etkin hale getirir [21]. Bu faaliyetleri gerçekleştirmek için kullandığı araç ise veri madenciliği teknikleridir. Veri madenciliği tekniklerinden kümeleme analizi, müşteri bilgilerinin kümelenmesi ve davranış şekillerinin belirlenmesinde MİY'in temelini oluşturur.

Veri ambarı ve veri madenciliği, MİY kapsamında, müşteri yaşam değerlerinin anlaşılması, müşteri işlemleri, müşteri analizi, müşteri profillerinin oluşturulmasında kullanılır [30], [24]. Böylece firma odaklanması gereken hedef kitleyi belirleyerek kaynaklarını etkin kullanma imkânı bulur.

4.1. Bölümleme

Tüm müşteriler aynı ihtiyaçlara sahip ve firma için aynı değerde olsaydı, kitle pazarlama ile müşterilere aynı önceliklerle ulaşmak mantıklı olabilir ve kitlesel pazarlama ile tek ürün, tek pazarlama karması ve tek reklamla müşterilere ulaşıldığı için tasarruf sağlanabilirdi [31]. Ancak özellikle tüketici istek ve gereksinimlerinin farklılık gösterdiği günümüzde farklı müşterilere aynı ürün ve aynı önceliklerle ulaşmaya çalışmak firmaların rekabet gücünü zayıflatmaktadır. Bunu aşmak için ise Smith'in (1956) gündeme getirdiği ve çok kısa sürede pazarlama teori ve pratiğinin odak noktası haline gelmiş olan bölümleme, firmalar için faydalı bir yöntem olmuştur. Smith, ürün ve servis taleplerindeki farklılığı anlamış ve bu ihtiyaçları benzerliklerine göre gruplayarak özelleşmiş öneriler sunmak gerektiğini söylemiştir [38].

Bölümleme konusu tüm sektörlerde çok sık kullanılan veri madenciliği uygulama konularından birisidir ve veri madenciliğinin gücü müşteri verilerinden müşteri profillerinin çıkarılmasında yatmaktadır [29]. Müşteri profili, müşterilerin modelidir ve istek ve gereksinimlerinin doğru tespit edilerek karşılanmasına yardımcı olur [33]. İstek ve gereksinimleri bakımından kendi içinde homojen, diğer müşteri gruplarına göre heterojen özellikler içeren grupların belirlenmesi ile firmalara, hitap edecekleri müşterileri daha yakından tanıma ve benzer müşterilere aynı pazarlama karması ile ulaşmayı sağlaması bakımından firma kaynaklarının etkin ve verimli kullanılmasına imkân sağlar [2].

Ürün ve hizmetlerin her geçen gün birbirine benzediği, çeşitli yeniliklerin ve teknolojik gelişmelerin çok kısa sürede taklit edildiği rekabet ortamında, karar destek sistemlerinin oluşturulması, stratejilerin belirlenmesi için müşteri değerinin belirlenmesi ve gruplara ayrılması gerekmektedir. Çünkü müşteri değeri artık oluşturulan ve yönetilen bir kavram olup, yönetilmesi için ölçülmesi gerekmektedir [20].

Müşteri değerinin tespit edilmesi ve arttırılması için ise çeşitli faaliyetlerde bulunulması zorunlu hale gelmiştir. Bu da müşteri ile ilgili tüm verilerin toplanmasını, analiz edilmesini ve müşteri değerini arttırıcı politikaların geliştirilmesini gerekli kılmaktadır [6].

MİY kapsamında yer alan bölümleme ve profil çalışmaları hedef kitleyi benzer tercih, satın alma davranışı ve karakteristiklerine, ürün geliştirme, pazarlama programlarına en uygun alt müşteri gruplarına ayrılmasıdır. Bölümleme, tüketici davranışları açısından benzer özellikler taşıyan homojen gruplar oluşturmak ve bu grupları detaylı olarak tanımlamak amacıyla yapılır. Buradaki asıl hareket noktası aynı grupta yer alan tüketiciler arasındaki farklılıkları en aza indirgeyerek, gruplar arasındaki farklılıkları ise maksimize ederek farklı homojen yapıları tanımlayabilmektir [17].

Pazarın bölümlenmesinden ziyade müşterilerin bölümlenmesi daha verimlidir. Bu noktada hangi bölümün daha kârlı olduğu ve hangisinin uzun vadeli getirisinin fazla olduğu önem kazanmaktadır. Eğer bir firma kuruluş aşamasında ise pazarı bölümleyerek fırsatları belirlemesi oldukça önemlidir. Ancak mevcut bir firmanın böyle bir lüksü yoktur. Bu nedenle mevcut müşterilerine odaklanmalı ve doğru müşterileri tespit etmelidir [24]. Bölümleme sayesinde mevcut ve potansiyel müşteri ihtiyaçları daha iyi anlaşılabilir, bu ihtiyaçların tatmin edici şekilde karşılanması durumunda ise müşteriler elde tutulabilmektedir. [24].

4.1.1. Literatür Araştırması

Bölümleme konusu, tarihsel gelişimi içinde, Smith'in (1956) kavram olarak ortaya koymasının ardından pazarlama alanında hemen kabul görmüş ve kısa sürede yaygınlaşmıştır. Daha sonra değişik modeller ve bilgi teknolojilerindeki gelişmelere bağlı olarak veri toplama ve analiz yöntemleri geliştirilmiştir. Aşağıda konunun gelişim aşamaları ve katkı sağlayan kişiler yer almaktadır.

1. Bölümleme temel ve modellerinde gelişmeler

   Hummel (1960), Haley (1968), Wind ve Cardozo (1974), Bonoma ve Shapiro (1983), Moriatry ve Reibstein (1986), File ve Prince (1996)

2. Veri gerekliliği ve veri toplama metodolojisi

   Webster (1978), Silk ve Kalvani (1982), Flodhammer (1988), Mitchell (1994)

3. İstatistiksel analiz araçlarındaki gelişmeler

   Frank ve Green (1968), Green ve Carmone (1977), Rao ve Winter (1977), Acito ve Jain (1980), Klastorin (1983), Green ve Krieger (1991), Fish (1995), Balakrishnan (1996)

4. Bölümlemenin strateji haline gelmesi

   Beik ve Buzby (1973), Mahajan ve Jain (1978), De Kluyver ve Whitlark (1986), Piercy ve Morgan (1995) [12].

Tablo 4.1'de tez kapsamında, konu ile ilgili incelenen bazı kaynaklar kısa açıklamalarıyla tablo halinde verilmiştir.

Tablo 4.1: Kaynak Künyeleri Tablosu

Kaynak Künyesi	Kullanılan Yöntem	Açıklama
Rygielski, C., Wang, J. C. ve Yen, D. C., 2002. Data mining techniques for customer relationship management, *Technology in Society* **24**, 483-502	Sinir ağları	Müşteri ilişkileri yönetimi kapsamında veri ambarlama ve veri madenciliği teknikleri ve bu tekniklerden iki tanesi örnek üzerinden açıklanmaktadır.
Shaw, M. J., Subramaniam, C., Tan, G. W. ve Welge, M. E., 2001. Knowledge management and data mining for marketing, *Decision Support Systems* **31**, 127-137.	Veri madenciliği, bilgi yönetimi, bilgi temelli pazarlama	Pazarlama karar destek sistemi kapsamında veri madenciliğinin ve bilgi yönetiminin önemi açıklanmaktadır.
Ertay, T., Çekyay, B. ve Wets, G. (editör), 2005. Integrated clustering modeling with backpropagation neural network for efficient customer relationship management, Intelligent Data Mining Techniques and Applications, Springer, Berlin.	SOM (Self-Organizing Map) ve sinir ağları	Müşteri İlişkileri Yönetimi kapsamında bölümleme için SOM ve sinir ağları metotları müşteri değeri bağlamında kullanılmıştır.
Lee, J. H. ve Park, S. C., 2005. Intelligent profitable customers segmentation system based on business intelligence tools, *Expert Systems with Applications* **29**, 145-152.	Veri zarflama analizi, SOM, sinir ağları	Müşteri ilişkileri yönetiminde kârlı müşterilerin bölümlendirilmesi açıklanmıştır.
Kim, Y. S. ve Street, W. N., 2004. An intelligent system for customer targeting: A data mining approach, *Decision Support Systems*, **37**, 215-228.	Yapay sinir ağları ve genetik algoritma	Doğrudan postalama etkinliğinin arttırılması ve kazancın maksimum yapılması için genetik algoritma ve sinir ağları ile cevap verme potansiyeli yüksek müşterilerin belirlenmesi gerçekleştirilmiştir.
VanRaaji, E. M., Vernooij, M. J. A. ve VanTriest, S., 2003. The implementation of customer profitability analysis: A case study, *Industrial Marketing Management* **32,** 573-583.	Faaliyet tabanlı maliyet	Müşteri kârlılık analizlerine göre müşteri bölümleri oluşturulur ve bu bölümlere uygun ulaşma şekilleri geliştirilmesinde kullanılır.
Liu, D. R. ve Shih, Y. Y., 2005. Integrating AHP and data mining for product recommendation based on customer lifetime value, *Information & Management,* **42,** 387-400.	Kümeleme, Birliktelik kuralları	Müşteri yaşam değeri, için değişkenler oluşturulmuş, bu değişkenler ağırlıklandırılmış ve kümeleme analizinde kullanılmıştır. Her gruba ürün önerisi için birliktelik kuralları kullanılmıştır.
Hwang, H., Jung, T. ve Suh, E., 2004. An LTV model and customer segmentation based on customer value: A case study on the wireless telecommunication industry, *Expert Systems with Applications*, **26,** 181-188.	Karar ağacı	Müşteri yaşam değerine göre kümeleme yapılmış ve her sınıfa uygun stratejiler geliştirilmiştir.

Erdoğan, Ş. Z., 2004. Veri madenciliği ve veri madenciliğinde kullanılan K-means algoritmasının öğrenci veritabanında kullanılması, *Yüksek Lisans Tezi*, İ.Ü Sosyal Bilimler Enstitüsü, İstanbul.	Kümeleme analizi, K-ortalamalar	Kümeleme analizi (K-ortalamalar) yöntemi ile öğrenci bölümleme uygulaması gerçekleştirilmiştir.
Menteş, G. T., 2000. Faktör ve kümeleme analizi yardımıyla bankacılık ürün ve hizmetlerinin araştırılması üzerine bir uygulama, *Doktora Tezi*, İ.Ü Sosyal Bilimler Enstitüsü, İstanbul.	Faktör analizi, kümeleme analizi	Bankacılık hizmetleri bakımından faktör ve kümeleme analizi yöntemleri ile bankalar bölümlere ayrılmıştır.
Min, H., Min, H. Ve Emam, A., 2002. A data mining approach to developing the profiles of hotel customers, International Journal of Contemporary Hospitality Management **14/6,** 274-285.	Karar ağacı	Müşteri ilişkileri yönetimi kapsamında veri madenciliği ile kârlı ve sadık müşterilerin elde tutulmasına yönelik politikaların geliştirilmesi için karar ağacı algoritması kullanılmıştır.
Machauer, A. ve Morgner, S., 2001. Segmentation of bank customers by expected benefits and attitudes, *International Journal of Bank Marketing*, **19/1,** 6-17.	Kümeleme analizi	Kümeleme analizi müşteri bölümleme ve profillerinin belirlenmesinde kullanılmıştır.
Suh, E. H., Noh, K. C. ve Suh C. K., 1999. Customer list segmentation using the combined response model, *Expert Systems with Applications*, **17**, 89-97.	Lojistik regresyon, sinir ağları	Bölümleme ve müşteri hedefleme için lojistik regresyon ve sinir ağları hedef pazarlama ile entegre olarak kullanılmıştır.
Jonker, J. J., Piersma N. ve Pool, D. V., 2004. Joint optimization of customer segmentation and marketing policy to maximize long-term profitability, *Expert Systems with Applications*, **27,** 159-168.	Markov karar süreci	Doğrudan pazarlama kampanyaları alanında karar ağacı yöntemi kullanılmıştır.
Erdoğan, Ş. Z. ve Timor, M., 2005. A data mining application in a student database, *Journal of Abronautics and Space Technologies*, **2/2,** 53-57.	Kümeleme analizi, K-ortalamalar	Üniversite giriş sınavı verilerine göre kümeleme analizi ile öğrencilerin bölümlere ayrılmasında kullanılmıştır.
Sağlam, B., Salman, S. F., Sayın, S. ve Türkay, M., 2006. A mixed-integer programming approach to te clustering problem with an application in customer segmentation, *European Journal of Operational Research*, **173,** 866-879.	Kümeleme analizi	Müşteri bölümlenmesi için kümeleme analizi yöntemi kullanılmıştır.

4.1.2. Bölümleme Kriterleri

Bölümleme sonucunda elde edilen bölümlerin anlamlı olabilmesi, üzerine stratejiler geliştirilmesi için bazı niteliklere sahip olmaları gerekmektedir. Bunlar;

- Tanımlama: Bölümler farklı ihtiyaçlar, coğrafik faktörler, farklı tutumlar açısından ayırt edilebilir olmalıdır.
- Cevap Verme: Seçilen bölüm üyeleri ürüne olumlu tepki göstermelidir. Müşteri grubunu belirli ürün ya da servis için gruplamak ilk adımdır. Daha sonra uygun reklam ve kampanya ile olumlu tepki alınabilecektir [31].
- Uygun Pazar Potansiyeli: Seçilen bölümler kârlı ve olabildiği kadar geniş olmalıdır.
- Ulaşılabilirlik: Seçilen bölümlere uygun iletişim, reklam araçlarıyla ekonomik olarak ulaşılabilmelidir.
- Sabitlik / Gelişme: Seçilen bölümler boyut, yapı ve boyut gelişiminde istenen nitelikte olmalıdır [31].

Odaklanılması gereken bölüm anlaşıldıktan sonra atılması gereken diğer adım, müşteri sadakati ve elde tutmayı sağlamaktır. Artık firmalar mevcut müşterileri elde tutmanın, yeni müşteri kazanımlarından daha önemli olduğunu anlamışlardır. Çünkü firmalar mevcut müşteriler hakkında bilgi sahibi iken, yeni müşteriler hakkında bilgi sahibi değildir [16]. Bu durum da firmaların müşterilerini tanıması için zaman ve maliyet bakımından bazı külfetlere katlanmasını gerektirmektedir. Buna karşın elde tutulabilen müşterilere yönelik daha etkin çapraz satış, gelişmiş doğrudan pazarlama kampanyaları ve bire bir pazarlama kampanyaları gibi değişik stratejiler izlenerek daha fazla kazanç sağlanabilmektedir. Böylece firmalar mevcut müşterilerini daha iyi algılayarak kârlarını arttırabileceklerdir [24].

Son yıllarda firmalar müşteri istek ve ihtiyaçlarını anlamaya yönelik faaliyetlere daha fazla önem vermeye başlamışlardır. Yoğun rekabet firmaları maliyetlerini minimum yapmaya zorlarken, müşteri farklılıkları ise istek ve ihtiyaçların fazlalığını gerektirmektedir.

Bu çelişkili durumda optimum ürün çeşitliliği için veri madenciliği ve onun kümeleme analizi yöntemi firmalara müşterilerini benzer ihtiyaçlara göre homojen gruplara ayırarak özelleşmiş ürün yada hizmet sunmalarına imkân sağlamaktadır [30].

Firmaların müşteri\bayi yönetimleri tüm sektörler için gerekli hale gelmiştir. Firmalar artık sadece müşterilere mal veya hizmet satmak değil daha yakın ilişkiler geliştirerek uzun süreli bağ kurmayı amaçlamaktadır. Bu da firmaların MİY'i gerçekleştirmelerini gerektirmektedir. Bu ise etkin karar destek sistemine bağlıdır. Bu kapsamda, tezde bayi yönetimi karar destek sistemi oluşturmak için veri madenciliği yöntemlerinden kümeleme analizi yöntemleri uygulanmıştır.

4.1.3. Bölümlemenin Faydaları

- Müşteri istek ve gereksinimlerinin daha iyi belirlenmesi ve bunların sağlanmasına imkân tanıması
- Rekabetin anlaşılması ve rekabette fark yaratacak unsurların takibi
- Daha etkin pazarlama planlarının geliştirilmesi
- Faaliyet sırasında ve sonrasında performansların daha iyi değerlendirilmesi
- Firma kaynaklarının daha etkin kullanımı [22]

5. BAYİ YÖNETİM SİSTEMİ MODELLEMESİ

Bayi yönetim sistemi modeli oluşturulması için kümeleme analizi yöntemleri kullanılmıştır. Modele katkı sağlaması ve bazı değişkenlerin türetilmesi için faktör analizleri ve kümeleme başarımlarının tespiti için diskriminant analizi uygulanmıştır. Belirlenen işletme amaçlarını gerçekleştirmek için gerekli ve mümkün olan veriler belirlenmiş, temin edilip işlenerek analize hazırlamıştır. Analizlerde kümeleme analizi teknikleri uygulanmasında uzaklık ölçüsü olarak kareli Euclide ölçüsü kullanılmıştır. Elde edilen küme doğruluklarının sınanması için, farklı tekniklerin sonuçları karşılaştırılmış, diskriminant analizi ile kümeleme başarımları belirlenmiştir. Son olarak küme profilleri belirlenerek firmanın mevcut bayi yapısı hakkında bilgiler elde edilmiş ve geleceğe yönelik stratejileri için katkı sağlanmaya çalışılmıştır.

Analizlerin gerçekleştirilmesi için SPSS 13.0 paket programı kullanılmıştır. Öncelikle veri yapısının anlaşılması için kategorik ve sürekli değişkenlerin tanımlayıcı istatistik bilgileri verilmiştir.

Uygulama yapılan firmanın Türkiye genelinde 76 bayisine ait verileri incelenmiştir. Firmanın geniş bir bayi ağı olmasına rağmen etkin bir bayi yönetim sistematiği yoktur. Mevcut durumda, daha önceki satış miktarlarına bağlı olarak gelecek yılın satış kotaları belirlenmekte ve bayi başarımları daha çok satış hedeflerinin gerçekleştirilip gerçekleştirilmediğine bağlı olarak ölçülmektedir. Bayilere ilişkin öznel değerlendirmeler satış personellerinde kalmakta, tüm firmaya mal edilememekte, dolayısıyla planlama yapma noktasında kullanılamamaktadır. Bayilerle ilgili olarak daha çok, ciro, sipariş miktarı ve aldığı ürünler vb. veriler mevcut olup, bunlar veri tabanında yer almakta üretim planlamaya girdi sağlayacak bilgiler oluşturulmamaktadır. Müşteri ilişkileri yönetimi kapsamında bazı faaliyetlere başlanmış olmasına rağmen henüz istenen seviyede değildir.

Firmanın yeni kurulmaya çalışılan bayilik yapısı, sadece firma markasını satan bayilikler olması yönündedir. Ancak mevcut durumda rakip markaları da firma bayileri satabilmektedir. Bu durum firma tarafından değiştirilmek istenen bir durumdur.

Bayi yönetim modeli için aşağıdaki işletme amaçları belirlenmiştir.

- Bayilerin kümelenmesi ve bayi profillerinin belirlenmesi
- Bayi niteliklerinin belirlenmesi ile özelleşmiş hizmet sunma ve kaynak tahsisinin mümkün olması
- Bayilik sistemi için performans kriterlerinin belirlenmesi, en iyi bayi profilinden hareketle gelişim planlarının oluşturulması
- Yeni bayilikler için stratejik bir sistem geliştirilmesi

5.1. Modelleme Aşamaları

Bayi değerlendirmesine yönelik yapılan bu çalışmada, verilerin hazırlanması ve analizleri için veri madenciliği adımları uygulanmıştır. Söz konusu adımlar ve yapılan işlemler aşağıda belirtilmiştir.

5.1.1. Kümelemeye Esas Olacak Değişkenlerin Seçilmesi

Veri madenciliği sonuçlarının stratejik kararlara girdi oluşturabilmesi için kümelerin etkin şekilde tanımlanmasını, bu da uygun değişkenlerin modelin oluşturulmasında kullanılmasını gerektirmektedir. Kullanılacak değişkenlerin aynı birim cinsinden olması gerekmektedir. Aksi halde yapılan analiz sonuçları yanıltıcı olabilmektedir. Değişkenler yapısına bağlı olarak nominal, ordinal, aralıklı veya oransal ölçekli olabilmektedir.

- Nominal ölçekli veriler, örneklem içerisindeki sıklık bilgilerinin bulunabileceği bir veri türüdür. Bu tür ölçme ile elde edilen rakamlar yalnız temsil ettikleri birimleri, olayları ve durumları gösterirler. Bu tür verilerle ortalama, standart sapma ve varyans gibi parametrik açıklayıcı istatistiksel bilgiler bulunamaz. Örneğin, çalışmada kullanılan Bolge_Kodu değişkeni sadece bayilerin coğrafik olarak hangi bölgede yer aldıklarını ifade etmektedir.

- Ordinal ölçek, nominal ölçeğe göre kısmen daha çok bilgi içermekte ve sıralama sözkonusu olduğundan dağılım hakkında genel bir fikir vermektedir. Ordinal ölçekli verilere örnek olarak kullanılan Isyeri_Buyukluk ve Bayi_Vade değişkenleri verilebilmektedir.
- Aralıklı ölçme, birçok istatistik tekniğin uygulanmasının mümkün olduğu ölçü yöntemidir. Bu ölçekli verilere örnek Toplanti_Katilim, Vitrin_Yeterlilik değişkenleridir.
- Oransal ölçekli verilerde mutlak bir sıfır değeri bulunmakta ve her noktayı temsil eden veri bir diğerine göre aritmetiksel olarak konumlandırılabilmektedir [27]. Bu tür verilere Top_Sip_Sayi ve Top_Ciro değişkenleri örnek olarak verilebilmektedir.

Tez kapsamında yukarıdaki unsurlar dikkate alınarak bayilerin değişik kriterler bakımından ele alınması ve benzer olanların gruplanması için kullanılacak değişkenler, yapılacak analiz ve elde edilmesi beklenen sonuçlara bağlı olarak belirlenmiştir. Tez kapsamında bayilerin sayısal değişkenleri (sipariş sayısı, elde edilen ciro, teminat miktarı vb.) firmanın 01.01.2007–30.09.2007 tarihleri arasındaki SAP kayıtlarından elde edilmiştir. Analizin kapsamını genişletmek için değişik kaynaklardan ve çeşitli değerlendirmeler esas alınarak Ek A'da verilen anket hazırlanmış ve bayiler ile bire bir ilişkide olan pazarlama, satış ve lojistik departmanlarında görevli kişilerden her bayi için öznel değerlendirmeleri alınmıştır. Tablo 5.1'de tez kapsamında, gerek SAP kayıtlarından gerekse hazırlanan anketten elde edilen değişkenler ve açıklamaları verilmiştir.

Tablo 5.1: Analizde Kullanılan Değişkenler ve Açıklamaları

	Değişken Adı	**Açıklama**	**Birim / Değer**
1	Bayi_Kodu	Bayi Tanımlayıcı	
2	Bolge_Kod	Bölge Tanımlayıcı	
3	Top_Ciro	Bayiden incelenen dönem içinde elde edilen ciro	(YTL)
4	Miktar_1	Bayinin 1 nolu üründen sipariş ettiği miktar	(KG)
5	Ciro_1	Bayiden 1 nolu üründen elde edilen ciro	(YTL)
6	Miktar_2	Bayinin 2 nolu üründen sipariş ettiği miktar	(KG)
7	Ciro_2	Bayiden 2 nolu üründen elde edilen ciro	(YTL)
8	Miktar_3	Bayinin 3 nolu üründen sipariş ettiği miktar	(KG)
9	Ciro_3	Bayiden 3 nolu üründen elde edilen ciro	(YTL)
10	Miktar_4	Bayinin 4 nolu üründen sipariş ettiği miktar	(KG)
11	Ciro_4	Bayiden 4 nolu üründen elde edilen ciro	(YTL)
12	Miktar_5	Bayinin 5 nolu üründen sipariş ettiği miktar	(KG)
13	Ciro_5	Bayiden 5 nolu üründen elde edilen ciro	(YTL)
14	Bayi_Kredi	Bayi kredibilitesi	(YTL)
15	Teminat_Mik	Bayinin teminat miktarı	(YTL)
16	Odenmemis	Bayinin ödenmemiş teminat miktarı	(YTL)
17	Sozlesme	Bayinin sözleşme durumu	(0, 1, 2)
18	Odeme_Prob	Bayinin ödeme problemi	(0, 1)
19	Bayi_Vade	Bayinin genel ödeme vadesi	(1, 2, 3, 4)
20	Toplanti_Katilim	Bayinin toplantılara katılımının değerlendirmesi	(10'lu likert)
21	Iletisim_Duzeyi	Bayinin firma ile olan iletişiminin değerlendirmesi	(10'lu likert)
22	Vitrin_Yeterlilik	Bayinin vitrin, görsel argümanlar vb. değerlendirmesi	(10'lu likert)
23	Urun_Gorsel_Oncelik	Bayinin firma ürününü rakip ürünlerinden görsel olarak önceliklendirmesi	(10'lu likert)
24	Stok_Tut_Egilimi	Bayinin stok tutma eğilimi	(10'lu likert)
25	Stok_Duzen	Bayi stok düzeninin değerlendirilmesi	(10'lu likert)
26	Indirim_Sip_Degisim	Bayinin indirim durumlarında sipariş değişim durumu	(10'lu likert)
27	Kriz_Sip_Degisim	Bayinin kriz durumlarında sipariş değişim durumu	(10'lu likert)
28	Potansiyeli	Bayinin gelecek dönemler için potansiyelinin değerlendirilmesi	(10'lu likert)
29	Kanal_Beceri	Bayinin kanalları kullanım becerisi	(10'lu likert)
30	Sip_Yaklasimi	Bayinin sipariş yaklaşımı	(10'lu likert)
31	Dagitim_Agi	Bayinin dağıtım ağı yapısı	(10'lu likert)
32	Reklam_Faaliyeti	Bayinin bireysel reklam faaliyetleri	(10'lu likert)
33	Basarim_Duzey	Bayinin başarı düzeyi	(10'lu likert)
34	Hedef_Tutturma	Bayinin hedef tutturma seviyesi	(10'lu likert)
35	Mulkiyet	Bayinin iş yeri mülkiyet durumu	(1, 2)
36	Isyeri_Buyukluk	Bayinin iş yeri büyüklüğü	(1, 2, 3)
37	Musteri_Kazanma	Bayinin bireysel müşteri kazanma çabası	(10'lu likert)
38	Musteri_Yakinlik	Bayinin müşteriye yakınlık durumu	(10'lu likert)
39	Stand_Gorunum	Bayinin stand görünümünün değerlendirmesi	(10'lu likert)
40	Musteri_Memnuniyet	Bayinin müşteri memnuniyet değerlendirmesi	(10'lu likert)
41	Geri_Bildirim	Bayinin geri bildirim seviyesi	(10'lu likert)
42	Sikayet_Cozum	Bayinin şikayet çözüm yeterliliği	(10'lu likert)
43	Musteri_Odak	Bayinin müşteri odaklılığının değerlendirmesi	(10'lu likert)
44	Bilgi_Düzey	Bayinin ürün bilgi seviyesi	(10'lu likert)
45	Per_Yeterlilik	Bayinin peronel yeterlilik durumu	(1, 2)

5.1.2. Uzaklık Ölçüsüne Karar Verilmesi

İki nesnenin benzerliğinden söz edebilmek için, bu iki nesnenin ortak özelliklere sahip olması gerekmektedir. Dolayısıyla, uygun bir kümeleme çözümüne ulaşılması için nesneler arasındaki yakınlık veya benzerliklerin net olarak ortaya konulması gerekmektedir [27]. Literatürde yer alan farklı uzaklık ölçüleri içerisinden birim değişikliklerinden etkilenmeyen kareli Euclide uzaklık ölçüsü kullanılmıştır.

Kareli Euclide Uzaklık Ölçüsü: $$d(i,j)=\sum_{s=1}^{n}\left(X_{is}-X_{js}\right)^2 \qquad (5.1)$$

5.1.3. Kümelemede Kullanılacak Tekniğin Seçilmesi

Veri madenciliği tekniklerinden kümeleme analizinin gerçekleştirilmesi için kullanılan SPSS paket programının, kümeleme analizine ilişkin olan,

- Hiyerarşik kümeleme yöntemi
- Hiyerarşik olmayan kümeleme yöntemi (k-ortalamalar)
- İki adım (twostep) yöntemleri kullanılmıştır.

Analizlerde farklı değişkenlerle her üç yöntem de kullanılmış ve sonuçları karşılaştırılmıştır.

5.1.4. Küme Oluşturma ve Küme Sayısının Belirlenmesi

Küme sayısının belirlenmesinde hiyerarşik kümeleme analizleri sonucunda elde edilen ağaç diyagramlarında nesneler arasında normal seyreden uzaklıkların bir anda arttığı nokta, analizin durma noktası olarak belirlenerek küme sayısı tespit edilebilir. Uzaklıkların bir anda arttığı nokta, artık oluşturulmakta olan kümeye benzerliği yüksek nesne kalmamasından dolayı, uzak nesnelerin kümeye dâhil edildiğini ifade etmektedir. Küme sayısına karar verme noktasında diğer bir yöntem ise, hiyerarşik kümeleme yöntemleri ile elde edilen küme sayısının, hiyerarşik olmayan kümeleme tekniklerinde uygulanması ve sonuçların karşılaştırılmasıdır.

Uygulama sırasında yapılan kümeleme analizlerinde hiyerarşik kümeleme analizlerinin farklı metotları ile elde edilen küme sayıları, hiyerarşik olmayan kümeleme analizlerinde girdi olarak kullanılmıştır. Dolayısıyla analizler sırasında nesnel değerlendirmelere yer verilmiştir.

5.1.5. Kümelerin Tanımlanması ve Yorumlanması

Analizler sonucunda elde edilen kümelerin, değişkenlere göre orta noktaları dikkate alınarak küme profilleri belirlenmiş, küme üyesi olan bayilerin genel yapısı hakkında bilgi verilmiştir. Buna bağlı olarak oluşan kümelemeden elde edilen bilginin hangi işletme amaçları ve hangi kararlarda kullanılabileceği yorumlanmaya çalışılmıştır.

5.2. Analizler

Tez kapsamında, Türkiye genelinde 76 bayisi değerlendirilen bir firmanın, bayilerine ilişkin elde edilen ciro, sipariş verilen ürünler ve miktarları, bayi kredibilitesi gibi verileri, firmanın 01.01.2007–30.09.2007 tarihleri arasındaki SAP kayıtlarından elde edilmiştir. Firma kaynaklarından ve hazırlanan anket sonucu elde edilen değişkenlere ilişkin istatistiksel açıklamalar Ek B'de verilmiştir.

Firma Türkiye genelinde kendi yapılanmasına göre toplam 12 bölgede bayilik ağına sahiptir. Bayi sayıları bölgelere göre eşit olarak dağılmamakta, coğrafik büyüklük, müşteri potansiyeli gibi nedenlere bağlı olarak farklılık göstermektedir. Bayilerin bölgelere göre dağılımları Tablo B.1'de verilmiştir.

Firmanın bayilik politikası gereği bayilerin sözleşmeli olması gerekmektedir. Ancak bazı bayilerin uzun yıllardır firmanın bayisi olması veya bazı özel nedenlerle sözleşmesi olmayan bayiler de mevcuttur. Bununla birlikte bayilik süreci başlamış ancak sözleşmeli olmayıp, sözleşme konusunda çekinceleri olmasından dolayı henüz sözleşme imzalamamış bayiler de mevcuttur. Bu veriye ilişkin bayi dağılımları Tablo B.2'de verilmiştir.

Ayrıca firma bayileri arasında, geç ödeme, hiç ödememe vb. şekillerde ödeme problemi olan bayiler mevcuttur. Bayilerin ödeme problemlerine göre dağılımları Tablo B.3'de verilmiştir.

Bayinin siparişlerine karşılık ödemesi gereken süreler farklılık gösterebilmektedir. Bu durum, bazı ürünlere has olabileceği gibi bayinin firma için konumundan da kaynaklanabilmektedir. Buna ilişkin bayilerin ödeme süreleri 45 gün ve daha az, 45 – 90 gün arasında, 90 gün ve 90 günden fazla olmak üzere genel olarak 4 farklı ödeme vadesi söz konusudur. Bu değişkene ilişkin dağılım Tablo B.4'de verilmiştir.

Bayiler faaliyetlerini sürdürme sırasında, ürün satışı, sevkiyatı ve teknik servisi konusunda personel bakımından yeterli olup olmadığına göre de farklılık göstermektedir. Personel yeterliliğine göre bayilerin yapısı Tablo B.5'de verilmiştir.

Bayiliklerde, bayinin faaliyet gösterdiği işyeri büyüklükleri de farklılık göstermektedir. Buna ilişkin değişken küçük, orta ve büyük şeklinde değerlendirilmiştir. Bununla birlikte işyerinin bayi mülkiyetinde olması veya kira olması gibi farklı iki durum da söz konusudur. İşyeri büyüklüklerine göre bayilerin dağılımları Tablo B.6'da, mülkiyet değişkenine göre dağılımları ise Tablo B.7'de verilmiştir.

Firmanın, bayilere ilişkin toplam sipariş sayısı, toplam ciro, bayi kredibilitesi, teminat miktarı gibi değişkenlerinin istatistiksel yapısı Tablo B.8'de verilmiştir.

5.2.1. İstatistiksel Analizler

Bu bölümde, Kümeleme analizi öncesi faktör analizi kullanılarak sürekli olmayan verilerin faktör skorları ile sürekli değişkenler türetilmesi ve bu değişkenler ile diskriminant analizi uygulanması sonucunda oluşturulan modele katkı sağlanması amaçlanmıştır.

5.2.1.1. Faktör Analizi (Anket Değişkenleri)

Tez kapsamında hazırlanmış ve Ek A'da verilmiş olan anketten elde edilen ve bayiler ile ilgili öznel değerlendirmeleri içeren değişkenler ve açıklamaları Tablo 5.2'de verilmiştir.

Tablo 5.2: Anketten Elde Edilen Değişkenler ve Açıklamaları

Değişken Adı	**Açıklama**
Toplanti_Katilim	Bayinin toplantılara katılımının değerlendirmesi
Iletisim_Duzeyi	Bayinin firma ile olan iletişiminin değerlendirmesi
Vitrin_Yeterlilik	Bayinin vitrin, görsel argüman vb. değerlendirmesi
Urun_Gorsel_Oncelik	Bayinin firma ürününü görsel olarak rakip markalara göre önceliklendirmesi
Stok_Tut_Egilimi	Bayinin stok tutma eğilimi
Stok_Duzen	Bayi stok düzeninin değerlendirilmesi
Indirim_Sip_Degisim	Bayinin indirim durumlarında sipariş değişim durumu
Kriz_Sip_Degisim	Bayinin kriz durumlarında sipariş değişim durumu
Potansiyeli	Bayinin gelecek dönemler için potansiyelinin değerlendirilmesi
Kanal_Beceri	Bayinin kanalları kullanım becerisi
Sip_Yaklasimi	Bayinin sipariş yaklaşımı
Dagitim_Agi	Bayinin dağıtım ağı yapısı
Reklam_Faaliyeti	Bayinin bireysel reklam faaliyetleri
Basarim_Duzey	Bayinin başarı düzeyi
Hedef_Tutturma	Bayinin hedef tutturma seviyesi
Musteri_Kazanma	Bayinin bireysel müşteri kazanma çabası
Musteri_Yakinlik	Bayinin müşteriye yakınlık durumu
Stand_Gorunum	Bayinin stand görünüm değerlendirmesi
Musteri_Memnuniyet	Bayinin müşteri memnuniyet değerlendirmesi
Geri_Bildirim	Bayinin geri bildirim seviyesi
Sikayet_Cozum	Bayinin şikâyet çözüm yeterliliği
Musteri_Odak	Bayinin müşteri odaklılığının değerlendirmesi
Bilgi_Düzey	Bayinin ürün bilgi seviyesi

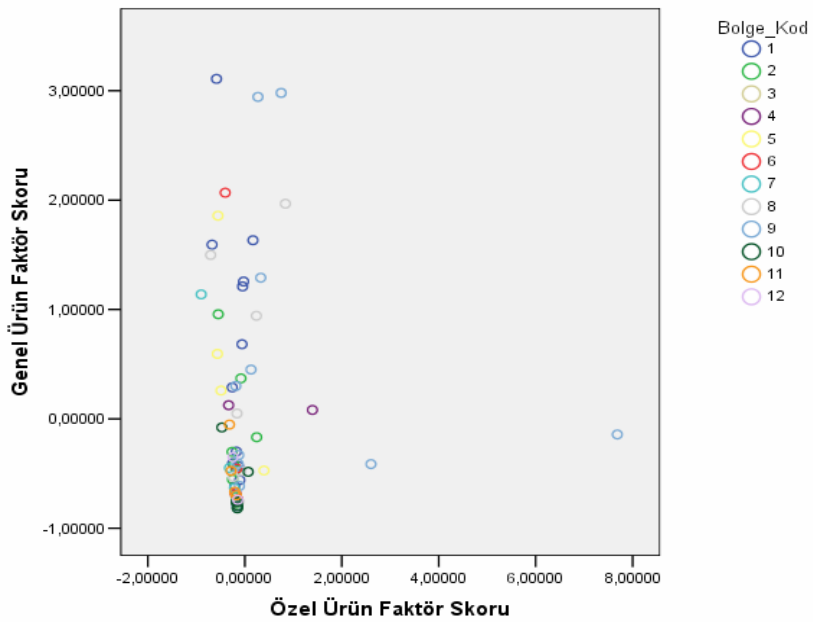
Anket soruları 10'lu likert ölçeği ile ölçülmüş olup faktör analizi uygulandığında Tablo C.1'de verilen, 23 değişken için ilgili faktörlerce açıklanan varyanslar elde edilmiştir. Tablo incelendiğinde her bir değişken için ilgili faktörlerce açıklanan varyansın %40 ile %92,5 arasında değiştiği gözlemlenmektedir. Tüm faktörler göz önünde bulundurulduğunda toplam varyansın %72,867'si açıklanmaktadır. Tablo C.2'de faktörlerce açıklanan varyansın yüzdelerine bakıldığında 1. faktör mevcut varyansın %35 gibi büyük bir kısmını tek başına açıklamaktadır, onu %8,8'le 2.faktör ve %7,41'le 3. faktör izlemektedir. Diğer faktörler de tabloda verildiği şekilde azalan yüzdelerde toplam varyansı açıklamaktadırlar. Tablo C.3'de ise hangi değişkenin faktör analizi sonrası hangi faktörce ağırlıklı olarak açıklandığı bilgisi verilmektedir. Tabloda faktör yükü 0,5'den küçük olan faktörler gösterilmemiştir. Dağılım incelendiğinde birinci faktörün tek başına Indirim_Sip_Degisim, Potansiyeli, Kanal_Beceri, Sip_Yaklasimi, Dagitim_Agi, Reklam_Faaliyeti, Basarım_Duzey, Hedef_Tutturma, Musteri_Kazanma, Geri_Bildirim, Sikayet_Cozum, Musteri_Odak, Bilgi_Duzey gibi pek çok değişkeni tek başına

büyük ölçüde açıkladığı görülmektedir. Buradan hareketle 1. faktör bayi başarısına ilişkin "Genel Başarı Faktörü" olarak adlandırılmıştır. Benzer şekilde 2. faktör incelendiğinde Stok_Tutma_Egilimi, Musteri_Yakinlik ve Kriz_Sip_Degisim değişkenleri açıklanmaktadır. Bu faktör ise "Müşteriye Ulaşma Faktörü" olarak isimlendirilmiştir. 3. faktör Stand_Gorunum ve Musteri_Memnuniyet değişkenleri ile ters ilişkili olup bu değişkenleri açıklayan faktör "Görsel Memnuniyetsizlik Faktörü" olarak ifade edilmiştir. 4. faktör Vitrin_Yeterlilik değişkeni ile doğru, Stok_Duzen değişkeni ile ters ilişkilidir. Dolayısıyla bu faktör "Görsel Etkililik Faktörü" olarak tanımlanmıştır. 5. faktör, Toplanti_Katılım değişkeni ile ters, Iletisim_Duzey değişkeni ile doğru ilişkilidir. Bu nedenle 5. faktör "Uzaktan İletişim Faktörü" olarak adlandırılmıştır. 6. faktör sadece Urun_Gorsel_Oncelik değişkenini 0,5'den fazla faktör yükü ile ifade etmektedir. Bu nedenle 6. faktör "Ürün Öncelik Faktörü" olarak isimlendirilmiştir. Son olarak 7. faktör tek başına Stok_Düzen değişkenini açıkladığından bu faktör "Stok Düzen Faktörü" olarak ifade edilmiştir. Tablo C.4, her bayi için, elde edilen 7 faktöre ait, söz konusu 23 değişken katsayılarını vermektedir. Bu katsayılar kullanılarak bayiler için her faktöre ait faktör skorları hesaplanabilmektedir. Faktör analizi, kullanılan değişken sayısını azaltmak için kullanılmaktadır. Burada elde edilen 7 değişken kısa değerlendirmeler için bayilerin kıyaslanmasında ya da bir bayinin değişiminin izlenmesinde kullanılabilecektir.

5.2.1.2. Faktör Analizi (Ürün Aileleri Miktar Verileri)

Firma üretimini ve ticaretini yaptığı onlarca ürün ailesine sahiptir. Ancak analiz konusu olarak firmanın üretimini yaptığı ve tüm satışları içinde yüksek orana sahip ürün ailelerinden 5'i için bayilerin sipariş miktarlarına ilişkin veriler analiz edilmiştir. Tablo D.1'de verilen ortak varyans tablosunda Miktar_1'e ilişkin varyansın toplam %93,7'si, Miktar_2'ye ilişkin varyansın toplam %66,4'ü, Miktar_3'e ilişkin varyansın toplam %58,3'ü, Miktar_4'e ilişkin varyansın toplam %47,3'ü ve Miktar_5'e ilişkin varyansın toplam %49'u açıklanmaktadır. Tablo D.2'de verilen, açıklanan varyans yüzdeleri tablosunda faktör seçiminde öz değerleri 1'den büyük olan faktörler kullanıldığından modelde yalnızca birinci ve ikinci faktörler yer almaktadır. Bu sebeple beş bileşene ilişkin toplam varyansın %62,936'sı [(2,13+1,016)/5] açıklanmaktadır. Tablo D.3 incelendiğinde, hangi değişkenin faktör analizi sonrası hangi faktörce ağırlıklı olarak açıklandığı görülmektedir. Dağılım incelendiğinde birinci faktörün Miktar_2, Miktar_3,

Miktar_4 ve Miktar_5 değişkenlerini tek başına büyük ölçüde açıkladığı görülmektedir. Buradan 1.faktör "Genel Ürün Faktörü" olarak, 2. faktör "Özel Ürün Faktörü" olarak adlandırılmıştır. Özel Ürün Faktörü (Miktar_1) ve Genel Ürün Faktörü (Miktar_2, Miktar_3, Miktar_4 ve Miktar_5) farklı ürünleri ifade ettiklerinden sipariş miktarları arasındaki ilişkiye bağlı olarak farklı faktörlerce açıklanmaktadırlar. Özel Ürün Faktör'ü ve Genel Ürün Faktörü'ne bağlı Şekil 5.1 incelendiğinde bayilerin satış değerlerine ilişkin varyansın büyük ölçüde "Genel Ürün Faktörü"nce kısmen de "Özel Ürün Faktörü"nce açıklandığı görülmektedir.



Şekil 5.1: Faktör Skorlarına Göre Saçılım Grafiği

Tablo D.4'de verilen faktör katsayılarına bağlı olarak aşağıda verilen denklemler elde edilmektedir.

Özel Ürün Faktör Skoru = 0,945*Miktar_1 + 0,148*Miktar_2 - 0,130*Miktar_3

+ 0,151*Miktar_4 - 0,151*Miktar_5 (5.2)

Genel Ürün Faktör Skoru = 0,039*Miktar_1 + 0,389*Miktar_2 + 0,344*Miktar_3

+ 0,328*Miktar_4 + 0,310*Miktar_5 (5.3)

Buna göre Özel ürün faktör skoru ve Genel ürün faktör skoru sırasıyla (5.2) ve (5.3) denklemleriyle hesaplanabilmektedir. Bu denklemler, bir bayinin her 5 ürün ailesinden verdiği sipariş miktarına bağlı olarak faktör skorları elde edilmesinde kullanılarak, bayilerin mevcut durumu, diğer bayilerle karşılaştırılması veya dönemsel olarak değerlendirilerek kendi içinde veya bulunduğu bölgelerde değişimlerinin incelenmesinde kullanılabilecektir.

5.2.1.3. Diskriminant Analizi (Odeme_Prob)

Faktör analizleri ile anket değişkenlerinden elde edilen 7 ve ürün aileleri miktar verilerinden elde edilen 2 türetilmiş değişken kullanılarak firmanın ödeme problemi olan bayilerinin tahmin edilmesine yönelik olarak diskriminant analizi uygulanmıştır. Firmanın incelenen 76 bayisinden, 64 bayisinin ödeme problemi olmayıp, 12 bayinin ödeme konusunda problemleri vardır. Tablo E.1'de verilen anlamlılık değerinin (0,037) %5'den küçük olması nedeniyle anlamlıdır. Tablo E.2'de verildiği gibi, fonksiyon kullanılan bağımlı değişken varyansının küçük bir kısmını ($(0{,}463)^2 = 0{,}21$) açıklamaktadır. Burada ilgili fonksiyona ilişkin Tablo E.3'de verilen Wilk's Lambda değerleri üzerinden hesaplanan anlamlılık düzeyleri %5'e göre anlamlı çıkmamaktadır. Çünkü toplam varyansın %78,6'sı açıklanamamaktadır. Tablo E.4'de verilen fonksiyon katsayıları modeldeki bağımsız değişkenlerin bağımlı değişkeni tahmin etmede göreceli önem değerlerini göstermektedir. Elde edilen fonksiyon için 0,719 Görsel Etkililik Faktör skoru göreceli olarak en önemli değişkendir. Buna göre (5.4) denklemi elde edilmektedir.

Skor (Odeme_Prob) = 0,519 * Genel Başarı Faktörü + 0,228 * Müşteriye Ulaşma Faktörü + 0,385 * Görsel Memnuniyetsizlik Faktörü + 0,719 * Görsel Etkililik Faktörü + 0,064 * Uzaktan İletişim Faktörü + 0,403 * Ürün Öncelik Faktörü + 0,096 * Stok Düzen Faktörü + 0,023 * Genel Ürün Faktörü - 0,018 * Özel Ürün Faktörü (5.4)

Bu denklemden elde edilecek Odeme_Prob değişkenine ait diskriminant skoru, Tablo E.5'de verilen grup merkezleri için fonksiyon değerleri ile karşılaştırılarak hangi kümeye ait merkez değerine yakın ise, değerlendirilen bayi o kümeye ait olarak tahmin edilmektedir.

Tablo E.6'de elde edilen denklemin, bayileri Odeme_Prob değişkenine göre doğru tahmin etme başarımı verilmiştir. Buna göre denklem, mevcut grubu %80,3 oranında doğru tahmin etmiştir. Elde edilen denklem ile her hangi bir bayinin 9 değişken (türetilen faktörler) değeri girildiğinde, söz konusu bayinin ödeme problemi olup olmadığı tahmin edilebilmektedir. Firma, bayilerini bu denklemle değerlendirerek, bayilerini ödeme problemleri olup olmaması konusunda tahmin etmek için kullanabilecektir. Ayrıca yeni bayilikler için bir ön sınama olarak kullanılabilmesi mümkündür. Elbette bu denklem mevcut veri setinden elde edildiği için, bu veri setinde her hangi bir değişiklik, yapılan analizlerin tekrar yapılmasını ve modelin yenilenmesini gerektirmektedir. Yeni bayiliklerin başlangıcı, mevcut bir bayinin bayilikten ayrılması veya söz konusu değişkenleri etkileyecek her hangi bir durum modelin başarımını azaltacaktır. Bu nedenle tüm veri madenciliği modelleri gibi yeniden düzenlenmesi ve analiz edilmesi gerekmektedir. Ancak buradan ödeme konusunda problemli olabilecek veya hiç problem çıkarmayacak bayilik profili belirlenebilecektir.

5.2.1.4. Diskriminant Analizi (Per_Yeterlilik)

Per_Yeterlilik değişkenine bağlı olarak firma bayilerinin yeterli (1) ve yeterli değil (2) şeklinde tahmin edilmesi için diskriminant analizi uygulandığında elde edilen diskriminant denklemi (5.5) verilmiştir.

Skor (Per_Yeterlilik) = 0,66 * Genel Başarı Faktörü + 0,00 * Müşteriye Ulaşma Faktörü - 0,222 * Görsel Memnuniyetsizlik Faktörü + 0,614 * Görsel Etkililik Faktörü + 0,316 * Uzaktan İletişim Faktörü + 0,523 * Ürün Öncelik Faktörü - 0,568 * Stok Düzen Faktörü - 0,885 * Genel Ürün Faktörü - 0,179 * Özel Ürün Faktörü (5.5)

Tablo 5.3: Grup Merkez fonksiyonları

	Fonksiyon
Per_Yeterlilik	1
1	,437
2	-,227

Bu fonksiyonun bayiler için uygulanması ile elde edilen fonksiyon değeri Tablo 5.3'de verilen grup merkezleri fonksiyon değerlerine yakınlığına bağlı olarak, bayiler personel yeterliliği açısından tahmin edilebilecektir.

Yapılan sınıflandırma başarımı, Per_Yeterlilik değişkeninin "1" değerine karşılık gelen grubu için %53,8, "2" grubu için %68'dir. Elde edilen denklemin personel yeterliliği açısından tüm veri setini doğru tahmin etme başarımı ise %63,2'dir. Elde edilen denklem, bayilerin personel yeterliliklerinin tahmin edilmesinde kullanılabilecek bir denklemdir.

Bayilik verilmesi noktasında göz önünde bulundurulmayan noktalardan biri olan personel yeterliliği çalışmada kişi sayısı olarak değerlendirilmiştir. Bayilerin yüklendikleri bir takım sorumlulukların yerine getirilmesi, müşteriye çok hızlı bir şekilde ulaşılabilmesi, şikâyetleri çözme noktasında yeterlilikleri değerlendirilmesi gerekli unsurlardır. Özellikle firma merkezinden uzakta olan bayi personellerinin tüm müşterilere ulaşabilme ve problemleri yardım almadan çözebilmeleri gerekmektedir. Bu nedenle firma, bayilerinin personel sayılarının yeterliliklerini değerlendirmelidir.

5.2.1.5. Diskriminant Analizi (Isyeri_Buyukluk)

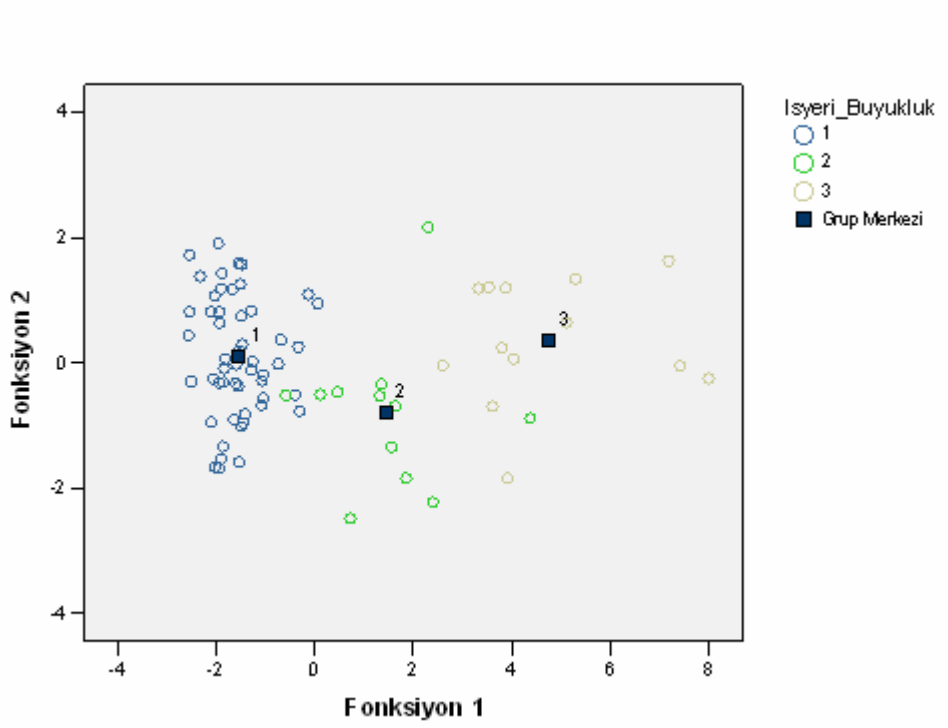
Isyeri_Buyukluk (işyeri büyüklüğü) değişkenine bağlı olarak, bayiler küçük (1), orta (2) ve büyük (3) olmak üzere 3 farklı değer almaktadır. Analiz yapılan veri setinde yer alan bayilerin 51 tanesi küçük, 12 tanesi orta ve 13 tanesi de büyük iş yerine sahiptir. Mevcut veri setinden türetilen 9 değişken ile diskriminant analizi uygulandığında elde edilen diskriminant denklemleri (5.6) ve (5.7) verilmiştir.

Skor (Isyeri_Buyuklugu 1) = 0,198 * Genel Başarı Faktörü + 0,187 * Müşteriye Ulaşma Faktörü - 0,024 * Görsel Memnuniyetsizlik Faktörü + 0,224 * Görsel Etkililik Faktörü - 0,201 * Uzaktan İletişim Faktörü - 0,024 * Ürün Öncelik Faktörü - 0,127 * Stok Düzen Faktörü + 0,991 * Genel Ürün Faktörü - 0,199 * Özel Ürün Faktörü (5.6)

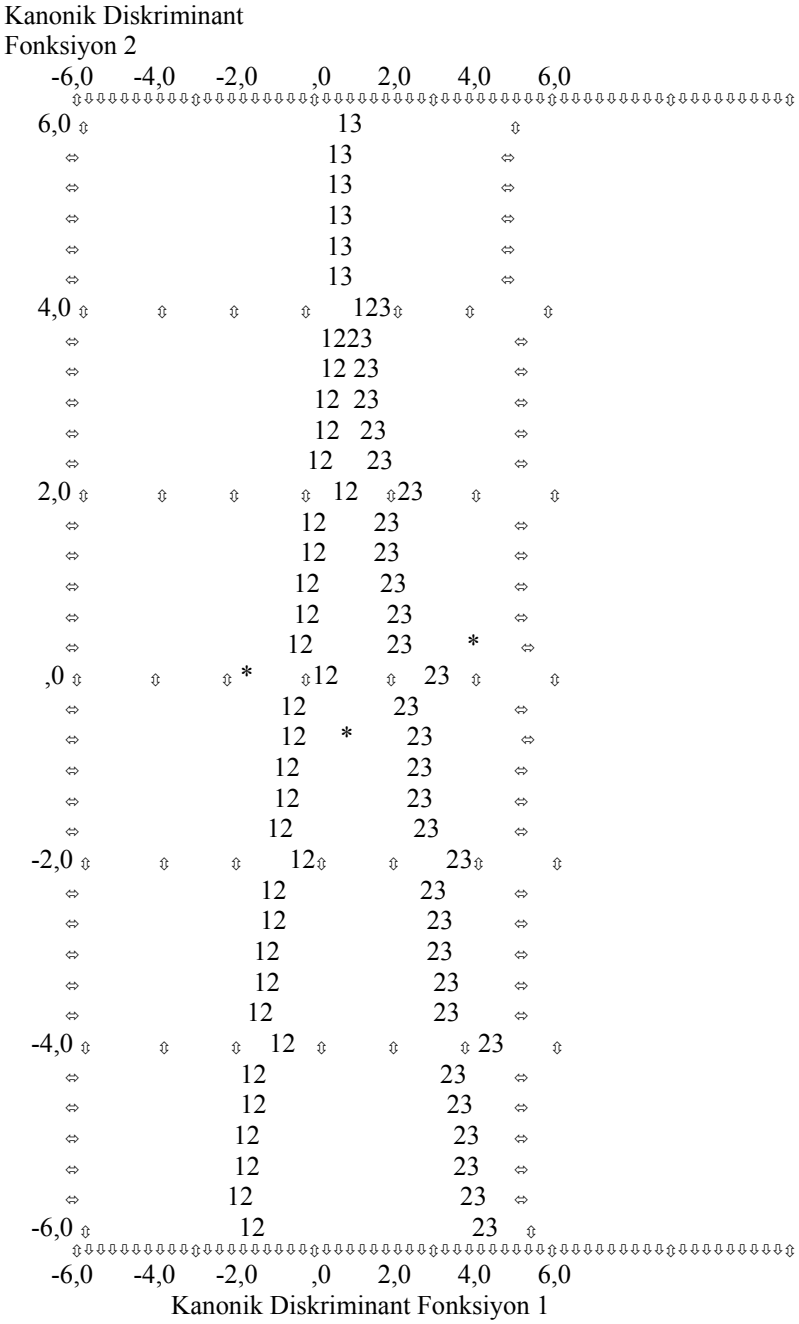
Skor (Isyeri_Buyuklugu 2) = - 0,526 * Genel Başarı Faktörü + 0,197 * Müşteriye Ulaşma Faktörü - 0,369 * Görsel Memnuniyetsizlik Faktörü + 0,008 * Görsel Etkililik Faktörü - 0,238 * Uzaktan İletişim Faktörü + 0,220 * Ürün Öncelik Faktörü + 0,733 * Stok Düzen Faktörü + 0,186 * Genel Ürün Faktörü + 0,056 Özel Ürün Faktörü (5.7)

Elde edilen denklemlerin grupları doğru tahmin etme başarımı Isyeri_Buyukluk "1" değeri için %100, "2" grubu için %75 ve "3" grubu için %92,3'dür. Genel olarak tüm veri setini doğru tahmin etme başarısı ise %94,7'dir. Diskriminant denklemleriyle oluşan gruplama Şekil 5.2 ve Şekil 5.3'de verilmektedir.

İşyeri büyüklükleri ve konumları bayilik sistemi için oldukça önemli niteliktedir. Burada diskriminant analizi sonucu elde edilen denklemler firmanın, bayi işyeri büyüklüklerini yüksek doğrulukta tahmin etmektedir. Firma bu bilgiyi yeni bayilikler verme noktasında kullanabilecektir. Ayrıca işyeri büyüklüklerinin satış miktarlarını da etkilediği düşünülürse, bayinin işyeri büyüklüğü bayi potansiyeli hakkında da fikir verebilecektir.



Şekil 5.2: Kanonik Diskriminat Fonksiyonu Grafiği



Şekil 5.3: Alan Haritası

5.2.2. Kümeleme Analizleri

Tez kapsamında incelenen veriler içerisinden uygun veri setleri oluşturacak veriler belirlenerek daha önce açıklanmış olan kümeleme teknikleri ile analiz edilmiştir. Söz konusu veri setleri ve kümeleme analizleri sonucunda elde edilen bilgiler aşağıda verilmiştir.

5.2.2.1. Kümeleme Analizi (Anket Değişkenleri)

Tablo 5.14'de verilen ve anket ile elde edilen öznel değerlendirmelere yönelik değişkenler kümeleme analizi yöntemleri ile analiz edilmiştir.

10'lu likert ölçeği ile ölçülmüş bayi tanımlamaya yönelik 23 değişkene farklı hiyerarşik kümeleme metotları uygulanmış ve elde edilen kümeler aşağıda verilmiştir.

Gruplar arası ortalama bağlantı metoduna göre;

1. kümeye 65 bayi üye olup, bayi tanımlayıcı numaraları: 1, 2, 3, 4, 5, 6, 7, 8, 11, 12, 13, 15, 16, 17, 19, 20, 21, 22, 23, 24, 25, 27, 28, 31, 32, 33, 34, 35, 37, 39, 40, 41, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 56, 57, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 74, 75, 76, 77, 78, 79'dur.

2. kümeye 11 bayi üye olup, bayi tanımlayıcı numaraları: 9, 10, 14, 18, 29, 30, 42, 54, 55, 58, 73'dür.

Gruplar içi ortalama bağlantı metoduna göre;

1. kümeye 24 bayi üye olup, bayi tanımlayıcı numaraları: 1, 6, 9, 10, 14, 15, 16, 17, 18, 20, 21, 24, 25, 29, 30, 31, 32, 35, 42, 48, 54, 55, 58, 72'dir.

2. kümeye 52 bayi üye olup, bayi tanımlayıcı numaraları: 2, 3, 4, 5, 7, 8, 11, 12, 13, 19, 22, 23, 27, 28, 33, 34, 37, 39, 40, 41, 43, 44, 45, 46, 47, 49, 50, 51, 52, 53, 56, 57, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 73, 74, 75, 76, 77, 78, 79'dur.

Tek bağlantı metoduna göre;

1. kümeye 75 bayi üye olup, bayi tanımlayıcı numaraları: 1, 2, 3, 4, 5, 6, 7, 8, 9, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 27, 28, 29, 30, 31, 32, 33, 34, 35, 37, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79'dur.

2. kümeye 1 bayi üye olup, bayi tanımlayıcı numarası: 10'dur.

Tam bağlantı metoduna göre;

1. kümeye 66 bayi üye olup, bayi tanımlayıcı numaraları: 1, 2, 3, 4, 5, 6, 7, 8, 11, 12, 13, 15, 16, 17, 19, 20, 21, 22, 23, 24, 25, 27, 28, 30, 31, 32, 33, 34, 35, 37, 39, 40, 41, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 56, 57, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79'dur.

2. kümeye 10 bayi üye olup, bayi tanımlayıcı numaraları: 9, 10, 14, 18, 29, 30, 42, 54, 55, 58'dir.

Merkezi bağlantı metoduna göre;

1. kümeye 75 bayi üye olup, bayi tanımlayıcı numaraları: 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 27, 28, 29, 30, 31, 32, 33, 34, 35, 37, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79'dur.

2. kümeye 1 bayi üye olup, bayi tanımlayıcı numarası: 13'dür.

Medyan bağlantı metoduna göre;

1. kümeye 75 bayi üye olup, bayi tanımlayıcı numaraları: 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 27, 28, 29, 30, 31, 32, 33, 34, 35, 37, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79'dur.

2. kümeye 1 bayi üye olup, bayi tanımlayıcı numarası: 13'dür.

Ward bağlantı metoduna göre;

1. kümeye 34 bayi üye olup, bayi tanımlayıcı numaraları: 1, 2, 6, 8, 9, 10, 14, 15, 16, 17, 18, 20, 21, 24, 25, 29, 30, 31, 32, 35, 41, 42, 48, 50, 51, 53, 54, 55, 58, 60, 63, 64, 65, 72'dir.

2. kümeye 42 bayi üye olup, bayi tanımlayıcı numaraları: 3, 4, 5, 7, 11, 12, 13, 19, 22, 23, 27, 28, 33, 34, 37, 39, 40, 43, 44, 45, 46, 47, 49, 52, 56, 57, 59, 61, 62, 66, 67, 68, 69, 70, 71, 73, 74, 75, 76, 77, 78, 79'dur.

Hiyerarşik kümeleme analizinde kullanılan yöntemlere göre, bayilerin küme üyelikleri incelendiğinde farklı sonuçların elde edildiği görülmektedir. Sonuçların karşılaştırılması için hiyerarşik olmayan kümeleme analiz yöntemi uygulanmıştır.

Bu yöntem, oluşturulması istenen küme sayısının belirlenmesini gerektirmektedir. Hiyerarşik kümeleme analizlerinin tamamında 2 küme oluşturulduğundan programa küme sayısı (k) 2 olarak girilmiştir.

Hiyerarşik kümeleme algoritması, küme merkezleri belirlendikten sonra bayilerin aldıkları değişken değerlerinin küme merkezine olan uzaklığına bağlı olarak bayilerin kümelere atanmasını gerçekleştirmiştir. Analiz yapıldığında oluşan iki kümenin, anket değişkenlerine bağlı olarak küme merkezleri Tablo 5.4'de verilmiştir.

Tablo 5.4: Son Küme Merkezleri

	Küme	
	1	2
Toplanti_Katilim	7,54	7,81
Iletisim_Duzeyi	5,79	5,58
Vitrin_Yeterlilik	6,29	6,04
Urun_Gorsel_Oncelik	5,39	5,77
Stok_Tut_Egilimi	2,36	1,60
Stok_Duzen	2,82	2,54
Indirim_Sip_Degisim	9,11	7,58
Kriz_Sip_Degisim	7,54	8,23
Potansiyeli	6,68	4,90
Kanal_Beceri	7,57	6,04
Sip_Yaklasimi	7,96	5,63
Dagitim_Agi	7,57	6,19
Reklam_Faaliyeti	3,89	2,15
Basarim_Duzey	5,86	4,15
Hedef_Tutturma	5,93	4,23
Musteri_Kazanma	7,64	5,46
Musteri_Yakinlik	6,21	4,54
Stand_Gorunum	5,29	2,73
Musteri_Memnuniyet	5,46	2,77
Geri_Bildirim	5,14	2,81
Sikayet_Cozum	4,64	2,67
Musteri_Odak	5,54	3,10
Bilgi_Düzey	7,32	5,58

Hiyerarşik olmayan kümeleme analizi sonucunda oluşan küme yapısı;

1. kümeye 28 bayi üye olup, bayi tanımlayıcı numaraları; 1, 6, 9, 10, 14, 15, 16, 17, 18, 20, 21, 24, 25, 29, 30, 31, 32, 35, 41, 42, 48, 50, 54, 55, 58, 60, 63, 72'dir.

2. kümeye 48 bayi üye olup, bayi tanımlayıcı numaraları; 2, 3, 4, 5, 7, 8, 11, 12, 13, 19, 22, 23, 27, 28, 33, 34, 37, 39, 40, 43, 44, 45, 46, 47, 49, 51, 52, 53, 56, 57, 59, 61, 62, 64, 65, 66, 67, 68, 69, 70, 71, 73, 74, 75, 76, 77, 78, 79'dur.

Yapılan kümeleme analizi sonucunda elde edilen küme yapısında, her değişken için anlamlı küme oluşup oluşmadığının kontrol edilmesi gerekmektedir. Bu nedenle grup içi varyanslarının kontrol edilmesi noktasında Tablo 5.5'de verilen anlamlılık (Sig.) değerine bakılmalıdır. Bu değerin 0,05'den küçük olması anket ile elde edilen değişkenlere göre küme ortalamaları arasında anlamlı fark olduğunu ifade eder ki bu zaten kümeleme analizinin ulaşmak istediği bir sonuçtur. Ancak tablo dikkatle incelendiğinde; Toplanti_Katilim, Iletisim_Duzeyi, Vitrin_Yeterlilik, Stok_Düzen, Urun_Gorsel_Oncelik, değişkenlerine ait anlamlılık değerinin 0,05'den büyük olması nedeniyle bu değişkenler için kümelerin birbirinden anlamlı derecede farklı olmadığı sonucuna varılmaktadır. Zira Tablo 5.4 incelendiğinde söz konusu değişkenlerin 1. ve 2. küme için aldığı değerlerin birbirine oldukça yakın olduğu gözlenmektedir.

Tablo 5.5: Kümeleme Analizi Varyans Tablosu

	Küme		Hata			
	Kareli Ortalama	df	Kareli Ortalama	df	F	Sig.
Toplanti_Katilim	1,355	1	1,247	74	1,086	,301
Iletisim_Duzeyi	,724	1	1,924	74	,376	,541
Vitrin_Yeterlilik	1,053	1	2,157	74	,488	,487
Urun_Gorsel_Onceli	2,526	1	1,016	74	2,488	,119
Stok_Tut_Egilimi	10,026	1	1,431	74	7,006	,010
Stok_Duzen	1,384	1	1,379	74	1,004	,320
Indirim_Sip_Degisim	41,063	1	1,356	74	30,282	,000
Kriz_Sip_Degisim	8,504	1	1,587	74	5,358	,023
Potansiyeli	56,203	1	1,792	74	31,369	,000
Kanal_Beceri	41,384	1	1,308	74	31,645	,000
Sip_Yaklasimi	96,773	1	2,273	74	42,572	,000
Dagitim_Agi	33,870	1	1,543	74	21,953	,000
Reklam_Faaliyeti	53,974	1	1,225	74	44,057	,000
Basarim_Duzey	51,790	1	1,857	74	27,891	,000
Hedef_Tutturma	51,072	1	1,842	74	27,720	,000
Musteri_Kazanma	84,392	1	1,221	74	69,123	,000
Musteri_Yakinlik	49,474	1	6,252	74	7,914	,006
Stand_Gorunum	115,583	1	1,989	74	58,108	,000
Musteri_Memnuniyet	128,293	1	2,074	74	61,871	,000
Geri_Bildirim	96,035	1	1,524	74	63,035	,000
Sikayet_Cozum	69,063	1	1,393	74	49,572	,000
Musteri_Odak	104,557	1	1,182	74	88,482	,000
Bilgi_Düzey	53,424	1	1,483	74	36,014	,000

Bu nedenle anlamlı ayırma niteliğine sahip olmayan; Toplanti_Katilim, Iletisim_Duzeyi, Vitrin_Yeterlilik, Stok_Düzen ve Urun_Gorsel_Oncelik değişkenleri çıkarılarak hiyerarşik yöntemlerden Ward metodu tekrar uygulandığında yeni oluşan küme yapısı aşağıda verilmiştir.

1. kümeye 37 bayi üye olup, bayi tanımlayıcı numaraları: 1, 2, 6, 8, 9, 10, 14, 15, 16, 17, 18, 20, 21, 24, 25, 29, 30, 31, 32, 35, 39, 41, 42, 48, 50, 51, 52, 53, 54, 55, 58, 59, 60, 63, 64, 65, 72'dir.

2. kümeye 39 bayi üye olup, bayi tanımlayıcı numaraları: 3, 4, 5, 7, 11, 12, 13, 19, 22, 23, 27, 28, 33, 34, 37, 40, 43, 44, 45, 46, 47, 49, 56, 57, 61, 62, 66, 67, 68, 69, 70, 71, 73, 74, 75, 76, 77, 78, 79'dur.

Aynı şekilde, 5 değişken çıkarılarak hiyerarşik olmayan kümeleme analizi uygulandığında oluşan küme yapısı aşağıdaki şekilde değişmektedir.

1. kümeye 28 bayi üye olup, bayi tanımlayıcı numaraları: 1, 6, 9, 10, 14, 15, 16, 17, 18, 20, 21, 24, 25, 29, 30, 31, 32, 35, 41, 42, 48, 50, 54, 55, 58, 60, 63, 72'dir.

2. kümeye 48 bayi üye olup, bayi tanımlayıcı numaraları: 2, 3, 4, 5, 7, 8, 11, 12, 13, 19, 22, 23, 27, 28, 33, 34, 37, 39, 40, 43, 44, 45, 46, 47, 49, 51, 52, 53, 56, 57, 59, 61, 62, 64, 65, 66, 67, 68, 69, 70, 71, 73, 74, 75, 76, 77, 78, 79'dur.

Hiyerarşik olmayan kümeleme analizinde oluşan kümelerin; Toplanti_Katilim, Iletisim_Duzeyi, Vitrin_Yeterlilik, Stok_Düzen ve Urun_Gorsel_Oncelik değişkenleri bakımından anlamlı derecede farklı olmamasından dolayı, İki adım analizinde söz konusu değişkenler çıkarılmış ve geriye kalan 18 değişken üzerinden kümeleme analizi yapılmıştır. Sonuç olarak, iki küme oluşmuş ve küme yapıları aşağıda verilmiştir.

1. kümeye 36 bayi üye olup, bayi tanımlayıcı numaraları: 1, 2, 6, 8, 9, 10, 14, 15, 16, 17, 18, 20, 21, 24, 25, 29, 30, 31, 32, 35, 41, 42, 48, 50, 51, 53, 54, 55, 58, 59, 60, 63, 64, 65, 72, 73'dür.

2. kümeye 40 bayi üye olup, bayi tanımlayıcı numaraları: 3, 4, 5, 7, 11, 12, 13, 19, 22, 23, 27, 28, 33, 34, 37, 39, 40, 43, 44, 45, 46, 47, 49, 52, 56, 57, 61, 62, 66, 67, 68, 69, 70, 71, 74, 75, 76, 77, 78, 79'dur.

Anlamlı ayırma niteliğine sahip olmayan değişkenlerin analizden çıkarılarak kümeleme analizlerine Ward yönteminin tekrarı, hiyerarşik olmayan kümeleme analizinin tekrarı ve iki adım kümeleme analizinin uygulanması sonucunda yukarıda verildiği gibi oluşan iki küme, farklı yöntemlerle farklı üyelik yapılarında elde edilmiştir. Hangi kümeleme çözümünün daha iyi kümeleme yaptığını test etmek için diskriminant analizinde verilen Wilks' Lambda değerleri, her üç yöntem için de karşılaştırıldığında her üçü için de anlamlı çıkmakla birlikte, Ward tekniği için bu değer (0,154), hiyerarşik olmayan kümeleme tekniği için (0,158) ve iki adım (twostep) tekniği için (0,143) olarak elde edilmiş olup, en iyi kümelemeyi hiyerarşik olmayan kümeleme analizinin sağladığı tespit edilmiştir.

Buna göre anket değişkenleri arasında anlamlı şekilde kümeleme yapma niteliğinde olan 18 değişken ile yapılan başarımı en yüksek küme yapısı aşağıda verilmiştir.

1. kümeye 28 bayi üye olup, bayi tanımlayıcı numaraları: 1, 6, 9, 10, 14, 15, 16, 17, 18, 20, 21, 24, 25, 29, 30, 31, 32, 35, 41, 42, 48, 50, 54, 55, 58, 60, 63, 72'dir.

2. kümeye 48 bayi üye olup, bayi tanımlayıcı numaraları: 2, 3, 4, 5, 7, 8, 11, 12, 13, 19, 22, 23, 27, 28, 33, 34, 37, 39, 40, 43, 44, 45, 46, 47, 49, 51, 52, 53, 56, 57, 59, 61, 62, 64, 65, 66, 67, 68, 69, 70, 71, 73, 74, 75, 76, 77, 78, 79'dur.

Tablo 5.6: Standartize Edilmiş Diskriminant Fonksiyon Katsayıları

	Fonksiyon
	1
Stok_Tut_Egilimi	-,241
Indirim_Sip_Degisim	,397
Kriz_Sip_Degisim	-,018
Potansiyeli	,262
Kanal_Beceri	-,054
Sip_Yaklasimi	,235
Dagitim_Agi	-,127
Reklam_Faaliyeti	,450
Basarim_Duzey	,527
Hedef_Tutturma	-,705
Musteri_Kazanma	,233
Musteri_Yakinlik	,333
Stand_Gorunum	-,101
Musteri_Memnuniyet	,462
Geri_Bildirim	,128
Sikayet_Cozum	,214
Musteri_Odak	,496
Bilgi_Düzey	,195

Skor (Fonksiyon) = -0,241*Stok_Tut_Egilimi + 0,397*Indirim_Sip_Degisim

-0,018*Kriz_Sip_Degisim +0,262*Potansiyeli

-0,054*Kanal_Beceri+0,235*Sip_Yaklasimi -0,127*Dagitim_Agi

+0,45*Reklam_Faaliyeti +0,527*Basarim_Duzey

–0,705*Hedef_Tutturma +0,233*Musteri_Kazanma

+0,333*Musteri_Yakinlik -0,101*Stand_Gorunum

+0,462*Musteri_Memnuniyet +0,128*Geri_Bildirim

+0,214*Sikayet_Cozum +0,496*Musteri_Odak

+0,195*Bilgi_Düzey (5.8)

Tablo 5.7: Merkez Grup Fonksiyonları

	Fonksiyon
Küme	1
1,00	2,980
2,00	-1,738

Tablo 5.6'da diskriminat analizi fonksiyonunun katsayıları verilmiştir. Bir bayi için tabloda yer alan değişkenlere göre, sahip olduğu değerler (5.8)'de verilen denklemde yerine konulduğunda elde edilen fonksiyon katsayısı, Tablo 5.7'de her iki küme için de verilen merkez değerlerinden en yakın olduğu kümeye atanmaktadır. Bu denklem sayesinde firma her bayi için yukarıdaki değişkenleri bilmesi durumunda, onun hangi küme üyesi olacağını tespit edebilecektir. Bu denklem firmanın yeni bayilikler verme noktasında da uygulayabileceği ek bir aşama olarak da kullanılabilecektir. Firma, bayi adayını ilk aşamada mümkün olan ve yukarıda verilen değişkenlere göre değerlendirebilecektir. Ancak, bazı değişkenler firma ile bayi arasında önceden var olması gereken ilişkileri gerektirmektedir. Dolayısıyla yeni bayilikler verme noktasında yeni bir model geliştirilmesi gerekmektedir.

İki adım kümeleme analizi sonucunda elde edilen ve Tablo 5.9'de verilen her iki kümenin orta noktaları incelendiğinde 1. küme üyesi bayilerinin tüm değişkenler bakımından 2. kümeden daha iyi konumda oldukları görülmektedir. 1. küme üyesi bayilerin indirim dönemlerinde sipariş artışları daha fazla iken, buna karşın kriz dönemlerindeki siparişlerindeki azalış daha azdır. Ayrıca stok tutma eğilimlerinin daha fazla olması 1. küme üyesi bayilerin finansal olarak daha iyi konumda

olduklarını göstermektedir. Kendi dağıtım ağının, bireysel reklam faaliyetlerinin, stand görünümünün, müşteri şikâyetlerini çözme özelliğinin ve ürün hakkındaki teknik bilgisinin fazla olması 1. küme üyelerinin firmaya olan katkısının fazla olduğunu, sadece firmanın reklam ve tanınırlığını kullanmak yerine, aynı zamanda bireysel olarak daha fazla çaba sarf ettiklerini göstermektedir. Firma ile belirlenen bayi hedeflerini yerine getirme konusunda da daha başarılıdırlar. Müşteri şikâyetleri konusunda firmayı bilgilendirmesi ve müşteri odaklı felsefeyi daha fazla benimsemeleri bu bayilerin firma imajına olumlu yönde katkı sağladıklarını göstermektedir.

Söz konusu bilgiler kullanılarak yeni bayilikler verilmesi noktasında, kullanılan değişkenler bakımından bayi adayının sahip olması gereken nitelikler belirlenerek daha seçici ve geleceğe dair daha planlı bir bayi yapılanmasına katkı sağlayabilecektir. Çünkü anketten elde edilen ve analizlerde kullanılan değişkenler artık müşteri odaklı çalışmak isteyen firmaların göz önünde bulundurması zorunlu olan unsurlardır. Firmanın bulunduğu sektör rekabeti yoğun olan ve marka değişikliklerinin kolaylıkla yapılabileceği bir pazardır. Müşteri ile yakın ilişki kurmayan ve müşteri problemlerine hızlı müdahale etmeyen bayilik yapısı firmaların rekabet gücünü zayıflatacağı gibi firma ömürlerini azaltacaktır.

Firma mevcut durumda firma imajına katkı sağlayacak yapılanma yoluna gittiğinde söz konusu unsurları mutlaka dikkate almalı, yeni bayiliklerde, işyeri büyüklüğünün, vitrin ve görsel argümanların gerekliliklerini sağlamaya zorlamalıdır. Bu stratejiyi mevcut durumda da uygulamaya koymalıdır.Ayrıca, yukarıda verilen değişkenlere firmanın verdiği önem derecesinde ağırlıklandırma yapılarak, daha önemli görülen değişkenlere göre ayrı bir model de geliştirmek mümkündür.

Değerlendirilen değişkenlere göre başarılı olan 1. küme üyesi bayilerin, bölgelere göre dağılımları Tablo 5.8'de verilmiştir.

Tablo 5.8: Küme Üyelerinin Bölgelere Göre Dağılımı

1. Bölge	15, 17, 20, 25, 31, 48, 72
2. Bölge	6, 10, 35
3. Bölge	16
5. Bölge	14, 24, 41, 50
6. Bölge	29, 32
8. Bölge	30, 42, 60, 63
9. Bölge	9, 18, 21, 54, 55, 58
12. Bölge	1

Birinci küme üyesi bayilerin bölgelere göre dağılımları incelendiğinde 1. bölgede 7 ve 9. bölgede 6 bayi yer almaktadır. Bu bilgi, her iki bölgede de %50 oranında iyi profile sahip bayi olduğunu göstermektedir. Burada dikkat edilmesi gereken bir nokta, 4., 7., 10. ve 11. bölgelerde iyi profile sahip bayilerin olmamasıdır. Firma bu bölgelere ağırlık vererek firma imajına katkı sağlayacak bayilikler tesis etmelidir.

Tablo 5.9: Küme Merkezleri Tablosu

		Indirim_Sip_Degisim		Kriz_Sip_Degisim		Potansiyeli		Kanal_Beceri		Sip_Yaklasimi	
		Ortalama	Std. Sapma	Ortalama	Std. Sapma	Ortalama	Std. Sapma	Ortalama	Std. Sapma	Ortalama	Std. Sapma
Küme	1	8,9167	1,42177	7,6667	1,35225	6,2500	1,53762	7,4722	,97060	7,8056	1,28329
	2	7,4500	,87560	8,2500	1,19293	4,9250	1,36603	5,8250	1,17424	5,3000	1,50555
	Birlesim	8,1447	1,37311	7,9737	1,29588	5,5526	1,58657	6,6053	1,35724	6,4868	1,87967

Dagitim_Agi		Reklam_Faaliyeti		Basarim_Duzey		Hedef_Tutturma		Musteri_Kazanma		Musteri_Yakinlik	
Ortalama	Std. Sapma	Ortalama	Std. Sapma	Ortalama	Std. Sapma	Ortalama	Std. Sapma	Ortalama	Std. Sapma	Ortalama	Std. Sapma
7,5278	1 10805	3,6389	1,43731	5,7778	1,22150	5,9167	1,20416	7,5000	,94112	5,8333	2,94230
5,9500	1 21845	2,0250	,76753	3,8750	1,32409	3,9000	1,23621	5,1500	1,00128	4,5500	2,13578
6,6974	1 40494	2,7895	1,38868	4,7763	1,58828	4,8553	1,58075	6,2632	1,52638	5,1579	2,61306

Stand_Gorunum		Musteri_Memnuniyet		Geri_Bildirim		Sikayet_Cozum		Musteri_Odak		Bilgi_Düzey		Stok_Tut_Egilimi	
Ortalama	Std. Sapma	Ortalama	Std. Sapma	Ortalama	Std. Sapma	Ortalama	Std. Sapma	Ortalama	Std. Sapma	Ortalama	Std. Sapma	Ortalama	Std. Sapma
4,6667	2,05635	4,8056	2,13568	4,8889	1,40972	4,4167	1,40153	5,1667	1,40408	7,3333	,75593	2,2778	1,40633
2,7750	1,09749	2,8250	1,10680	2,5750	,98417	2,4750	,90547	2,9500	,87560	5,2250	1,22971	1,5250	,96044
3,6711	1,87181	3,7632	1,93817	3,6711	1,66844	3,3947	1,51507	4,0000	1,60000	6,2237	1,47512	1,8816	1,24330

5.2.2.2. Kümeleme Analizi (Ürün Ailesi Miktar Verileri)

Analizler kapsamında üretmekte olduğu ve satışlarının büyük bir bölümünü oluşturduğu 5 ürün ailesi seçilmiştir. Bayiler tarafından söz konusu 5 ürün ailesine ilişkin sipariş verilen miktar verileri kümeleme analizi ile analiz edildiğinde elde edilen Tablo 5.10 varyans tablosunda değişkenlerin anlamlılık değerleri verilmiştir. Görüldüğü gibi 1. ürünün miktar (Miktar_1) verisine ait anlamlılık değeri 0,05'den büyük olması, kümeleme için uygun bir değişken olmadığını ortaya koymaktadır.

Tablo 5.10: Varyans Tablosu

	Küme		Hata			
	Ortalama Kare	df	Ortalama Kare	df	F	Sig.
Miktar_1	13608,001	1	57381,111	74	,237	,628
Miktar_2	5,147E+014	1	2,008E+012	74	256,275	,000
Miktar_3	5,549E+010	1	6623808430	74	8,377	,005
Miktar_4	2032991482	1	173898786,7	74	11,691	,001
Miktar_5	4,407E+012	1	3,793E+011	74	11,620	,001

Miktar_1 değişkeni çıkarılarak analiz farklı kümeleme teknikleriyle tekrarlandığında ve oluşan kümeleme analizlerinin başarımını ölçmek için diskriminant analizi uygulandığında, hiyerarşik olmayan kümeleme analizi ve hiyerarşik yöntemlerden (Ward, merkezi, gruplar arası) elde edilen kümeleme başarımlarını gösteren Wilks' Lambda değeri 0,22'dir. İki adım kümeleme yönteminin aldığı değer ise 0,241'dir. En yüksek değer olan 0,313 değeri ise gruplar içi kümeleme analizi ile elde edilmektedir. Gruplar içi yöntemiyle elde edilen küme yapısı;

1. kümeye 73 bayi üye olup, bayi tanımlayıcı numaraları: 1, 2, 3, 4, 6, 7, 8, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 27, 28, 29, 30, 31, 32, 33, 34, 35, 37, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 55, 56, 57, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79'dur.

2. kümeye 3 bayi üye olup, bayi tanımlayıcı numaraları: 5, 54, 58'dir.

Gruplar içi kümeleme yöntemiyle elde edilen kümelere diskriminant analizi uygulanması sonucu elde edilen fonksiyon (5.9)'da verilmiştir.

$$\text{Fonksiyon1} = 1{,}253*\text{Miktar_2} - 0{,}448*\text{Miktar_3} + 0{,}01*\text{Miktar_4} - 0{,}344*\text{Miktar_5} \quad (5.9)$$

Fonksiyonun kullanılması ile elde edilen yeni değerlendirmelerin hangi kümeye dâhil edileceğine dair küme merkezleri Tablo 5.11'da verilmiştir.

Tablo 5.11: Küme Merkezleri Tablosu

	Fonksiyon
Küme	1
1,00	-,299
2,00	7,167

Kümeleme analizi sonucunda 2. kümeye atanan 5, 54 ve 58 kodlu bayilerin veri setinden söz konusu ürün miktarlarına ilişkin değerlerine bakıldığında bu bayilerin mevcut bayi yapısında miktar olarak en çok sipariş veren bayiler olduğu tespit edilmiştir. Söz konusu ürün aileleri ayrı ayrı incelendiğinde, 1. küme üyesi olan bayilerden bazı ürün ailelerinde sipariş miktarları daha fazla olan bayiler vardır. Ancak bütün bir değerlendirme yapılacak olması durumunda yukarıdaki tespit geçerlidir.

Söz konusu 3 bayinin bölgeleri incelendiğinde 5 kodu verilmiş olan bayi 2. bölgede, 54 ve 58 kodlu bayiler ise 9. bölgede yer almaktadırlar. Bu bölgelerde yer alan bayi sayısına bakılacak olursa 2. bölgede toplam 10, 9. bölgede 12 bayi yer almaktadır. Bu durumda bu bölge pazarları ve bu bayiler oldukça önemlidir. Ayrıca bu bayilerin bulundukları bölge, firmanın faaliyet gösterdiği bölgeler içinde bayi sayısı en fazla bölgeler olması ve buna rağmen 1 veya 2 bayinin ön plana çıkmış olması bu bayilerin önemini arttırmaktadır. Özellikle bu bayilerin bulundukları bölgeler için potansiyelerinin de yüksek olduğunu söylemek yanlış olmayacaktır. Zira 54 ve 58 kodlu bayilerin ikisinin toplam sipariş miktarları kendi bölgelerinin yarısı kadardır. Dolayısıyla bölge pazarlarının kalan payı diğer bayiler arasında paylaşılmaktadır. Firma bu durumu, bölgedeki diğer bayilerin performanslarını arttırıcı yönde kullanmalıdır. Oluşan iki kümeden ilki "satış hacmi sınırlı", ikinci küme ise "satış hacmi yüksek" olarak adlandırılabilir.

Gerek bu bayilerin bulundukları bölge potansiyellerinin yüksek olması gerekse bireysel olarak yüksek miktarda ürünü satabilmeleri, firmanın satışlarında meydana gelecek dönemsel, mevsimsel vb. satış azalmalarında ilk aşamada değerlendirilecek uyarıcı gösterge olarak kullanılmalarını sağlayabilecektir.

5.2.2.3. Kümeleme Analizi (Finansal Veriler)

Top_Ciro (Toplam Ciro), Bayi_Kredi (Bayi Kredibilitesi), Teminat_Mik (Teminat Miktarı) ve Odenmemis (Ödenmemiş Teminat Miktarı) değişkenlerine hiyerarşik kümeleme metotlarından Ward analizi uygulanması sonucunda; 1. küme 51, 2. küme 25, Gruplar arası, Gruplar içi, Tam bağlantı metotları uygulandığında; 1. küme 68, 2. küme 8, hiyerarşik olmayan kümeleme analizi uygulandığında 1. küme 63, 2. küme 13 son olarak iki adım yöntemi uygulandığında 1. küme 63, 2. küme 13 bayiden oluşan kümeleme yapıları oluşmuştur. Kullanılan değişkenler kümeleri oluşturmada anlamlıdır. Ayrıca her kümelemeye diskriminat analizi uygulandığında en yüksek grup içi varyansını oluşturan kümelemenin hiyerarşik metotlarla (Gruplar içi, Gruplar arası ve Tam bağlantı) elde edildiği tespit edilmiştir ve Wilks' Lambda değerleri 0,324'dür. Buna karşın iki adım ve hiyerarşik olmayan kümeleme metotlarının söz konusu değeri 0,249 ve Ward yönteminin 0,296'dır. Buna göre, gruplar içi kümeleme yönteminin sonucu oluşan küme yapıları aşağıdaki gibidir.

1. kümeye 68 bayi üye olup, bayi tanımlayıcı numaraları: 1, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 27, 28, 30, 31, 32, 33, 34, 35, 37, 39, 40, 41, 43, 44, 45, 46, 47, 49, 51, 52, 53, 55, 56, 57, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 73, 74, 75, 76, 77, 78, 79'dur.

2. kümeye 8 bayi üye olup, bayi tanımlayıcı numaraları: 2, 29, 42, 48, 50, 54, 58, 72'dir.

Gruplar içi kümeleme yöntemi sonucu oluşan kümelere diskriminant analizi uygulandığında, elde edilen fonksiyon aşağıda verilmiştir.

Fonksiyon 1 = 1,166 * Top_Ciro – 0,54 * Teminat_Mik + 0,597 * Odenmemis (5.10)

Burada dikkat edileceği gibi Bayi_Kredi (Bayinin kredibilitesi) değişkeni, fonksiyonda yer almamaktadır. Bu değişken belirli tolerans dâhilinde tüm bayileri etkilediğinden program bu değişkeni fonksiyona dâhil etmemiştir. Bu denklemdeki değişkenlere ilişkin bayi değerleri denklemde yerine konulduğunda, bayiye ilişkin fonksiyon değerini verecektir. Bu değer ise Tablo 5.12'de yer alan kümeler için fonksiyon değerleriyle karşılaştırıldığında, hangi fonksiyon değerine yakın ise söz konusu bayinin o kümeye atanacağını ifade etmektedir.

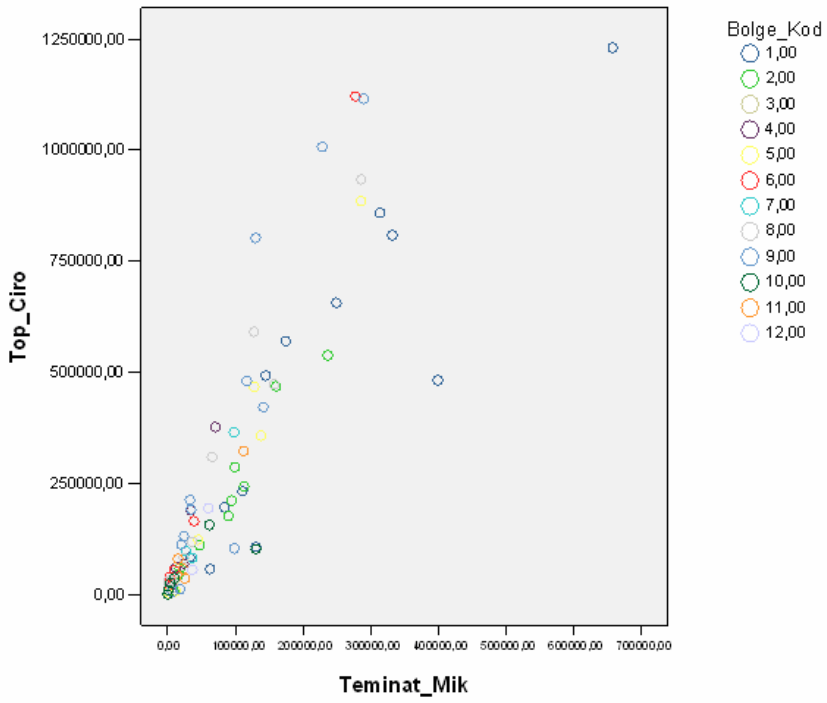
Tablo 5.12: Grup Merkezleri Fonksiyonu

Küme	Fonksiyon 1
1,00	-,489
2,00	4,155

Yukarıda verilen küme yapısında "finansal olarak güçlü" olarak adlandırılması mümkün 2. küme üyesi olan bayiler incelendiğinde, 2, 48 ve 72 kodlu üç bayi 1. bölge, 50 kodlu bayi 5. bölge, 29 kodlu bayi 6. bölge, 42 kodlu bayi 8. bölge ve 54 ile 58 kodlu bayiler 9. bölge üyesi bayilerdir. Firmanın söz konusu bayilerinin satış kotaları belirlenirken daha önce gerçekleştirdikleri satışlar yanında finansal yapıları da göz önünde tutulmalıdır. Bayiler rakip ürünleri de satabildikleri için bu oldukça önemlidir. Daha fazla satış yapan bayilere daha yüksek kâr payı uygulanabilir.

5.2.3. Regresyon Analizi (Top_Ciro ve Teminat_Mik)

Bu aşamada sonuçlara katkı sağlamak ve firma bayilik sistematiği için bayiden elde edilen ciro (Top_Ciro) ile bayiden istenen teminat miktarı (Teminat_Mik) arasında nasıl bir ilişkinin oluğunu analiz etmek için regresyon analizi kullanılmıştır. Şekil 5.4'de verilen değişkenlere ait saçılım diyagramında, arada yüksek bir korelasyonun olduğu gözlenmektedir.



Şekil 5.4: Saçılım Diyagramı

İki değişken arasındaki ilişkiyi bulmak için regresyon analizi yapıldığında Top_Ciro (Toplam Ciro) ve Teminat_Mik (Teminat Miktarı) değişkenleri arasında Tablo 5.13'de görüldüğü gibi 0,882 gibi yüksek bir korelasyon derecesi tespit edilmiştir. Buna ek olarak bağımlı değişkendeki (Teminat_Mik) %77,7'lik değişim modele eklenen bağımsız değişken (Top_Ciro) tarafından açıklanmaktadır.

Tablo5.13: Korelasyon Tablosu

Model	R	R Kare	Ayarlanmis R kare	Tahmin Std. hatas
1	,882[a]	,777	,774	54573,47551

a. Tahmin Eden: (Sabit), Top_Ciro

Regresyon denklemi için bulunan Tablo 5.14'deki sabit terim 0,49 anlamlılık değerindedir ve denklemin katsayısı 5798,67'dir. Buna göre söz konusu iki değişken arasındaki ilişkiyi veren denklem (5.11)'de verilmiştir.

Top_Ciro = 5798,67 + 0,324 * Teminat_Mik (5.11)

Tablo 5.14: Model Katsayı

Model		Standatize Edilmemis Katsayi		Standartize Edilmis Katsayi	t	Sig.
		B	Std. Hata	Beta		
1	(Sabit)	5798,670	8338,871		,695	,489
	Top_Ciro	,324	,020	,882	16,064	,000

Bu denklem, bayilerin verdiği 1000 YTL'lik Teminat_Mik artışının söz konusu bayinin Top_Ciro değerini 324 YTL arttıracağını ifade etmektedir. Bu bilgi firmanın yeni bayilikler vermesi ve mevcut bayilerin teminat miktarlarının düzenlenmesi için kullanabilecek bir bilgidir. Buna göre yeni bayiliklerde beklenen cirosuna göre bayilerin teminat miktarlarının belirlenmesinde kullanılabilir.

6. SONUÇ

Müşteri ilişkileri yönetimi kapsamında değerlendirilen ve tedarik zinciri içerisinde de önemli bir yere sahip olan bayi kavramının yönetilmesi için çeşitli değişkenler kullanılarak kümeleme analizi teknikleri uygulanmıştır.

Yapılan kümeleme analizleri sonucunda kullanılan değişkenlere bağlı olarak en yüksek anlamlılıkta kümelemeyi gerçekleştiren yöntem, değişken ve elde edilen küme yapılarının özeti Tablo 6.1'de verilmiştir.

Tablo 6.1: Kümeleme Analizleri Özet Tablosu

YÖNTEM	DEĞİŞKENLER	KÜME YAPISI	
		1. KÜME	2. KÜME
Hiyerarşik Olmayan Kümeleme Analizi	Stok_Tut_Egilimi, Indirim_Sip_Degisim Kriz_Sip_Degisim, Potansiyeli Kanal_Beceri, Sip_Yaklasimi Dagitim_Agi, Reklam_Faaliyeti Basarim_Duzey, Hedef_Tutturma Musteri_Kazanma, Musteri_Yakinlik Stand_Gorunum, Musteri_Memnuniyet Geri_Bildirim, Sikayet_Cozum Musteri_Odak, Bilgi_Düzey	1, 6, 9, 10, 14, 15, 16, 17, 18, 20, 21, 24, 25, 29, 30, 31, 32, 35, 41, 42, 48, 50, 54, 55, 58, 60, 63, 72	2, 3, 4, 5, 7, 8, 11, 12, 13, 19, 22, 23, 27, 28, 33, 34, 37, 39, 40, 43, 44, 45, 46, 47, 49, 51, 52, 53, 56, 57, 59, 61, 62, 64, 65, 66, 67, 68, 69, 70, 71, 73, 74, 75, 76, 77, 78, 79
Hiyerarşik Kümeleme Analizi (Gruplariçi)	Miktar_2 Miktar_3 Miktar_4 Miktar_5	1, 2, 3, 4, 6, 7, 8, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 27, 28, 29, 30, 31, 32, 33, 34, 35, 37, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 55, 56, 57, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79	5, 54, 58
Hiyerarşik Kümeleme Analizi (Gruplariçi, Gruplararası)	Top_Ciro Bayi_Kredi Teminat_Mik Odenmemis	1, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 27, 28, 30, 31, 32, 33, 34, 35, 37, 39, 40, 41, 43, 44, 45, 46, 47, 49, 51, 52, 53, 55, 56, 57, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 73, 74, 75, 76, 77, 78, 79	2, 29, 42, 48, 50, 54, 58, 72

Kümeleme analizleri öncesinde bayilerin sipariş verdikleri ürün miktarlarından hareketle faktör analizi yapılmış ve 5 ürün miktarını temsil eden 2 faktör denklemi elde edilmiştir. Ürün ailelerinin satış miktarları arasında ilişkiyi ortaya koyan ve analizle elde edilen denklemlerle gerek mevcut bayilerin sipariş miktarlarına göre diğer ürünlerin ağırlıklandırılması yapılabilir, gerekse bir bayinin zaman içinde sipariş değişimi izlenebilir. Ayrıca diğer bayilerle karşılaştırma yapılabilir veya dönemsel olarak değerlendirilerek kendi içinde veya bölgelerde değişimleri incelenebilir ve üretim planlamaya katkı sağlanabilir.

Mevcut veri setinde toplam ciro (Top_Ciro) ve teminat miktarı (Teminat_Mik) arasında yüksek bir korelasyon olduğu tespit edilmiştir. Buna ilişkin denklem, Teminat_Mik değişimine bağlı olarak Top_Ciro değişkeninin ne oranda etkileneceğini ortaya koymaktadır. Bu bilgi firmanın yeni bayilikler vermesi sırasında kullanabileceği bir bilgidir.

Anket sorularından elde edilen 7 ve 5 ürün ailesi miktar verilerinden elde edilen 2 faktörün diskriminant analizinden elde edilen fonksiyonlar sayesinde, Sozlesme değişkeni %59,2, Odeme_Prob değişkeni %80,3, Per_Yeterlilik değişkeni %63,2 ve Isyeri_Buyukluk değişkenini %94,7 oranında doğru tahmin eden fonksiyonlar elde edilmiştir. Bu denklemler aynı zamanda yeni bayilikler verme aşamalarında da kullanılabilecektir.

Kümeleme analizlerinde, anket sorularının analizi sonunda elde edilen 1. küme, ürün aileleri miktar değişkenleri kullanılarak elde edilen 2. küme ve finansal değerlere ilişkin değişkenlerin kullanılması ile elde edilen 2. küme, değerlendirme kriterlerine göre en iyi kümeler olmuştur.

Farklı değişkenlerin kullanılması ile elde edilen iyi profile sahip kümelerin ortak bayiler olup olmadığı incelendiğinde, anket değişkenleri ve finansal değişkenlerin ortak bayileri 29, 42, 48, 50, 54, 58, 72 numaralı bayilerdir. Dolayısıyla bu bayiler hem öznel değerlendirme kriterlerine göre hem de finansal verilere göre diğer bayilere göre daha iyi konumda bulunmaktadırlar. Anket ile değerlendirilen bayilerin stok tutma seviyeleri, indirim ve kriz dönemlerindeki sipariş değişimleri finansal verilerle yakından ilişkili olmasına rağmen diğer değerlendirmeler bakımından da bu küme üyesi bayiler, diğer küme üyesi bayilere göre daha iyi konumdadırlar.

Kullanılan üç değişken grubu için de ortak bir değerlendirme yapılacak olursa, en iyi profile sahip kümelerin ortak bayileri 54 ve 58 kodlu bayilerdir. Bu bayiler hem finansal değişkenler bakımından hem öznel değerlendirmeler bakımından hem de

ürün sipariş miktarları bakımından firma bayileri içerisinde en iyi olarak adlandırılabilecek iki bayidir. Bu bayilerin yakından izlenmesi ile firma genel yapı hakkında fikir sahibi olabilecektir. Özellikle satış miktarlarındaki değişimlerin (mevsimsel, dönemsel vb.) izlenmesinde kullanılabileceklerdir.

Yapılan analizler sonucunda firmanın bayi profili belirlenmiş, elde edilen denklemler sayesinde bayilerin bireysel olarak da hangi profile sahip küme üyesi olabileceği tahmin edilebilecektir. Bu denklemler aynı zamanda yeni bayilikler verilmesi noktasında da ya doğrudan ya da bazı değişikliklerle beraber kullanılabilecek, bu yönde yapılacak çalışmalara katkı sağlayabilecektir.

Eğer firmanın genel olarak cirosunda bir düşme söz konusu ise, öncelikli olarak finansal olarak güçlü olan bayilerden elde edilen cirolar incelenerek söz konusu değişikliğin dönemsel, mevsimsel, bölgesel vb. farklılıklardan kaynaklanıp kaynaklanmadığı sonucuna kısa vadede ulaşılabilecektir. Dolayısıyla bu bayilerin durumundaki değişiklikler diğer bayilerin göstergesi olarak kullanılabilecektir.

Yapılan analizler sonucunda firmanın bayi profili belirlenmiş, kullanılan değişkenlere göre en iyi bayiler tespit edilmiştir. Bu analizler sonucunda elde edilen bilgilerle firma, kampanya, reklam, personel eğitimi, hedef belirleme, satış kotaları belirleme vb. değişik stratejileri daha çok hedefe yönelik yapabilecektir. Hatta stratejileri gerçekleştirme noktasında her bayi ile özelleşmiş ilişkiler kurulabilecektir.

Elde edilen sonuçlar firmanın mevcut veri setinden elde edilmiş olup mevcut durumunu ortaya koymaktadır. Veri madenciliği sürekli bir süreç olduğundan yeni bayilikler verilmesi, bayilikten ayrılma ya da mevcut durumda kullanılan değişkenler bakımından fark oluşturacak durumlarda analizlerin tekrarlanması ve modelin güncellenmesi gerekmektedir.

KAYNAKLAR

[1] **Akpınar H.**, 2004. Veri tabanlarında bilgi keşfi ve veri madenciliği, *İ.Ü. İşletme Fakültesi Dergisi*, **29,** 1-22.

[2] **Andrews, R.L. ve Currim, I.S.**, 2003. Recovering and profilling the true segmentation structure in markets: an emprical investigation, *International Journal of Researcsh in Marketing*, 20, 177-192.

[3] **Bandyopadhyay, S. ve Maulik U.**, 2005. Advanced methods for knowledge discovery and data mining from complex data, Springer, Texas.

[4] **Berry, M.J.A., Linoff, G.S.**, 2000. Mastering data mining: the art an science of customer relationship management, John Wiley & Sons, New York.

[5] **Berson A., Smith S. ve Thearling K**., 2004. Data mining applications for CRM, www.thearling.com.

[6] **Best, R.J.**, 2005. Market-based management strategies for growing customer value and profitability, Pearson Hall, New Jersey.

[7] **Bounsaythip C. ve Runsala E. R.**, 2001. Research Report Overview of data mining for customer behaviour modeling, VTT Information Technology.

[8] **Churchill G. A.**, 1992. Basic marketing research, The Dryden Press, Orlando.

[9] **Erdoğan, Ş. Z. ve Timor, M.**, 2005. A data mining application in a student, database, *Journal of Abronautics and Space Technologies*, **2,** 53-57.

[10] **Erdoğan, Ş. Z.**, 2004. Veri madenciliği ve veri madenciliğinde kullanılan K-means algoritmasının öğrenci veritabanında kullanılması, *Yüksek Lisans Tezi*, İ.Ü Sosyal Bilimler Enstitüsü, İstanbul.

[11] **Ertay, T. ve Çekyay, B.**, 2005. Integrated clustering modeling with backpropagation neural network for efficient customer relationship management, Intelligent Data Mining Techniques and Applications, Springer, Berlin.

[12] **Goller, S., Hogg, A. ve Kalafatis S.P.**, 2002. A new research agenda for business segmentation, *European Journal of Marketing*, **36,** 252-271.

[13] **Hair J. F., Black, W. C., Babin B. J., Anderson R. E., Tatham, R. L.**, 2006. Multivariate data analysis, Prentice Hall, New Jersey.

[14] **Han, J. ve Kamber, M.**, 2001. Data mining: concepts and techniques, Morgan Kaufmann, San Francisco.

[15] **Hill, T. ve Lewicki, P.**, 2006. Statistics methods and applications - a comprehensive reference for science, Industry and Data Mining, Statsoft Inc., Tulsa.

[16] **Hwang, H., Jung, T. ve Suh, E.**, 2004. An LTV model and customer segmentation based on customer value: A case study on the wireless telecommunication industry, *Expert Systems with Applications*, **26, 1**81-188.

[17] **Jonker, J. J., Piersma N. ve Pool, D. V.**, 2004. Joint optimization of customer segmentation and marketing policy to maximize long-term profitability, *Expert Systems with Applications*, **27,** 159-168.

[18] **Kırım, A.**, 2001. Strateji, biri bir pazarlama ve CRM, Sistem Yayıncılık, İstanbul.

[19] **Kim, Y. S. ve Street, W. N.**, 2004. An intelligent system for customer targeting: A data mining approach, *Decision Support Systems*, **37**, 215-228.

[20] **Knox, S.**, Loyalty-based segmentation and the customer development process, European Management Journal, 16, 729-737.

[21] **Kotler, P., Amstrong, G.**, 2004. Principle of marketing, Pearson Prentice Hall, New Jersey.

[22] **Lancaster, G. ve Massingham, L.**, 2001. Marketing Management, McGrawhill, London.

[23] **Liu, D. R. ve Shih, Y. Y.**, 2005. Integrating AHP and data mining for product recommendation based on customer lifetime value, *Information & Management,* **42,** 387-400.

[24] **Meltzer, M.**, 2007. Customer relationship management – segment your customers based on profitability, Active Management Techniques, www.amt.com.

[25] **Oğuzlar, A.,** 2003. Veri ön işleme, *Erciyes Üniversitesi İktisadi ve İdari Bilimler Fakültesi Dergisi***, 21,** 67-76.

[26] **Özekes, S.,** 2003. Veri madenciliği modelleri ve uygulama alanları, *İstanbul Ticaret Üniversitesi Dergisi,* **3,** 65-82.

[27] **Özkay, M.B.**, 2000. Kümeleme analizi ve bankalar üzerine bir uygulama, *Yüksek Lisans Tezi*, İ.Ü. Sosyal Bilimler Enstitüsü, İstanbul.

[28] **Ruan, D., Chen, G. ve Kerre E.E.**, 2005. Intelligent data mining techniques and applications, Springer, Berlin.

[29] **Rygielski, C., Wang, J. C. ve Yen, D. C.**, 2002. Data mining techniques for customer relationship management, *Technology in Society,* **24**, 483-502.

[30] **Sağlam, B., Salman, S. F., Sayın, S. ve Türkay, M.**, 2006. A mixed-integer programming approach to te clustering problem with an application in customer segmentation, *European Journal of Operational Research*, **173,** 866-879.

[31] **Schiffman, L.G. ve Kanuk, L.L.**, 1987. Consumer behaviour, Prentice Hall, New Jersey.

[32] **Sezen, B.**, 2004. Veri zarflama analizi ile tedarik zinciri ortaklarının performans değerlendirmesi, 24. Yöneylem Araştırması / Endüstri Mühendisliği Kongresi, Gaziantep-Adana, 15-18 Haziran.

[33] **Shaw, M. J., Subramaniam, C., Tan, G. W. ve Welge, M. E.**, 2001. Knowledge management and data mining for marketing, *Decision Support Systems* **31**, 127-137.

[34] **Swift, R.S.**, 2001. Accelerating customer relationship / using CRM and relationship technologies, Prentice Hall, New Jersey.

[35] **Tonta, Y.**, 2004. Bilgi yönetiminin kavramsal tanımı ve uygulama alanları, Kütüphaneciliğin Destanı Sempozyumu, 21-24 Ekim, Ankara.

[36] **Turban, E., Aronson J. E.**, 1990. Decision Support Systems and Intelligent Systems, Prentice Hall, New Jersey.

[37] **Vaura, T.G.**, 1999. Müşteri tatmini ölçümlerinizi geliştirmenin yolları, Kalder, İstanbul.

[38] **Wedel, M. ve Kamakura, W.**, 2000. Market segmentation conceptual and methodological foundations, Kluwer Academic Publishers, Boston.

EK A - ANKET SORULARI

S1 Bayinin düzenlenen toplantı ve eğitimlere katılımını nasıl değerlendirirsiniz.
Düşük 1 2 3 4 5 6 7 8 9 10Yüksek

S2 Bayinin firma ile olan iletişim düzeyinin nasıl değerlendirirsiniz
Düşük 1 2 3 4 5 6 7 8 9 10Yüksek

S3 Bayiyi tabela, vitrin vb. görsel argümanların yeterliliği konusunda nasıl değerlendirirsiniz [37].
Düşük 1 2 3 4 5 6 7 8 9 10Yüksek

S4 Bayide firma ürünlerinin, rakip ürünlerden görsel olarak ön planda olma derecesini nasıl değerlendirirsiniz? [37]
Düşük 1 2 3 4 5 6 7 8 9 10Yüksek

S5 Bayi firma ürünlerini ödeme kolaylığı, kampanya vb. destekleyici çabalarla ön plana çıkarmaktadır.
1) Evet 2) Hayır

S6 Bayinin stok tutma eğilimini nasıl değerlendirirsiniz.
Düşük 1 2 3 4 5 6 7 8 9 10Yüksek

S7 Bayi sipariş vermek için kendi güvenli stok seviyesini göz önünde bulundurur (ürün bitmeden sipariş verir)
1) Evet 2) Hayır

S8 Bayinin ürün stok ve depo düzenini nasıl değerlendirirsiniz.
Düşük 1 2 3 4 5 6 7 8 9 10Yüksek

S9 Kampanya, indirim vb. durumlarda bayinin sipariş değişimi (artış yönünde) nasıl değişmektedir.
Düşük 1 2 3 4 5 6 7 8 9 10Yüksek

S10 Zam dönemlerinde bayinin sipariş değişimi (azalış yönünde) nasıl değişmektedir.
Düşük 1 2 3 4 5 6 7 8 9 10Yüksek

S11 Kriz dönemlerinde bayi sipariş değişimi (azalış yönünde) nasıl değişmektedir.
Düşük 1 2 3 4 5 6 7 8 9 10Yüksek

S12 Bayinin performans ve potansiyel bakımından gelişim fırsatını nasıl değerlendirirsiniz.

Düşük 1 2 3 4 5 6 7 8 9 10Yüksek

S13 Bayinin satış stratejisi, sadece markanın tanınırlığı ve reklamlarıdır. (bireysel girişimlerde bulunmaz)

1) Evet 2) Hayır

S14 Bayinin farklı satış kanalları kullanım becerisini nasıl değerlendirirsiniz.

Düşük 1 2 3 4 5 6 7 8 9 10Yüksek

S15 Bayiyi sipariş miktarı ve sıklığı göz önüne alındığında nasıl değerlendirirsiniz.

Temkinli 1 2 3 4 5 6 7 8 9 10 Saldırgan

S16 Bayi servis ağı ve dağıtım kanalları yeterliliğini nasıl değerlendirirsiniz.

Düşük 1 2 3 4 5 6 7 8 9 10Yüksek

S17 Bayinin bölgesel reklam çalışmalarına katılım derecesini nasıl değerlendirirsiniz.

Düşük 1 2 3 4 5 6 7 8 9 10Yüksek

S18 Bayinin hedeflerini başarma derecesini nasıl değerlendirirsiniz.

Düşük 1 2 3 4 5 6 7 8 9 10Yüksek

S19 Bayinin hedeflerini düzenli tutturma derecesi nedir?

Düşük 1 2 3 4 5 6 7 8 9 10Yüksek

S20 Bayi iş yeri mülkiyet durumu nedir [8] ?

1) Kendi mülkiyeti 2) Kira

S21 Bayi iş yeri büyüklüğü ne kadardır?

1) Küçük 2) Orta 3) Büyük

S22 Bayinin kendi imkânları ile müşteri kazanma ve tutma başarısını nasıl değerlendirirsiniz.

Düşük 1 2 3 4 5 6 7 8 9 10 Yüksek

S23 Bayi iş yerinin müşterilerine yakınlık durumunu nasıl değerlendirirsiniz.

Uzak 1 2 3 4 5 6 7 8 9 10 Yakın

S24 Bayinin stand görünümünü firma imajına olumlu katkısı bakımından nasıl değerlendirirsiniz.

Düşük 1 2 3 4 5 6 7 8 9 10 Yüksek

S25 Bayiyi müşteri hizmet ve müşteri memnuniyeti bakımından seviyesini nasıl değerlendirirsiniz [32].

Düşük 1 2 3 4 5 6 7 8 9 10 Yüksek

S26 Bayinin müşteri şikâyet ve isteklerini firmaya geri bildirim derecesini nasıl değerlendirirsiniz [8], [37].

Düşük 1 2 3 4 5 6 7 8 9 10 Yüksek

S27 Bayinin müşteri şikâyetlerine yaklaşımını ve çözüme katkısını nasıl değerlendirirsiniz [8].

Düşük 1 2 3 4 5 6 7 8 9 10 Yüksek

S28 Bayinin yeni veya geliştirilmiş ürünü pazara sürme konusunda istek seviyesi nedir [38] ?

Düşük 1 2 3 4 5 6 7 8 9 10 Yüksek

S29 Bayilik yapısı nasıldır [8] ?

1) Bireysel 2) Ortaklık

S30 Bayinin, ürünlerin teknik özellikleri ve kullanım alanları konusunda bilgi düzeyleri ne seviyededir [8] ?

Düşük 1 2 3 4 5 6 7 8 9 10 Yüksek

S31 Bayi personel sayısı yeterlidir [32].

1) Evet 2) Hayır

EK B – FREKANS TABLOLARI

Firmanın Türkiye genelinde sahip olduğu ve tez kapsamında değerlendirilen 76 bayisinin bölgelere göre dağılımı Tablo B.1'de bölge kodlarına bağlı olarak verilmiştir. Buna göre, firma, 1 bölgesinde 12 (%15,8'i), 2 bölgesinde 10 (%13,2)'si, 3 bölgesinde 2 (%2,6'sı), 4 bölgesinde 5 (%6,6'sı), 5 bölgesinde 6 (%7,9'u), 6 bölgesinde 5 (%6,6'sı), 7 bölgesinde 3 (%3,9'u), 8 bölgesinde 9 (%11,8'i), 9 bölgesinde 12 (%15,8'i), 10 bölgesinde 6 (%7,9'u), 11 bölgesinde 4 (%5,3'ü) ve 12 bölgesinde 2 (%2,6'sı) bayiye sahiptir.

Tablo B.1: Bolge_Kodu Değişkeni Frekans Tablosu

		Frekans	Yüzde	Mevcut Yüzde	Kümülatif Yüzde
Mevcut	1	12	15,8	15,8	15,8
	2	10	13,2	13,2	28,9
	3	2	2,6	2,6	31,6
	4	5	6,6	6,6	38,2
	5	6	7,9	7,9	46,1
	6	5	6,6	6,6	52,6
	7	3	3,9	3,9	56,6
	8	9	11,8	11,8	68,4
	9	12	15,8	15,8	84,2
	10	6	7,9	7,9	92,1
	11	4	5,3	5,3	97,4
	12	2	2,6	2,6	100,0
	Toplam	76	100,0	100,0	

Sozlesme (Sözleşme) değişkenine göre bayilerin frekansları Tablo B.2'de verilmiştir. Firmanın, çeşitli nedenlerle sözleşme yapılmamış bayi (0) sayısı 21, sözleşme yapılmış bayi (1) sayısı 40 ve sözleşme imzalama konusunda çekinceleri olan bayi (3) sayısı 15'dir.

Tablo B.2: Sozlesme Değişkeni Frekans Tablosu

		Frekans	Yüzde	Mevcut Yüzde	Kümülatif Yüzde
Mevcut	,00	21	27,6	27,6	27,6
	1,00	40	52,6	52,6	80,3
	2,00	15	19,7	19,7	100,0
	Toplam	76	100,0	100,0	

Odeme_Prob (Ödeme problemi) değişkeninin frekansları Tablo B.3'de verilmiştir. Odeme_Prob değişkeni, ödeme problemi olmayan bayi (0) sayısı 64 (%84,2'si), ve ödeme problemi olan bayi (1) sayısı 12 (%15,8)'dir.

Tablo B.3: Odeme_Prob Değişkeni Frekans Tablosu

		Frekans	Yüzde	Mevcut Yüzde	Kümülatif Yüzde
Mevcut	0	64	84,2	84,2	84,2
	1	12	15,8	15,8	100,0
	Toplam	76	100,0	100,0	

Bayi_Vade (Bayinin ödeme vadesi) değişkenine ilişkin frekanslar Tablo B.4'de verilmiştir. Bayi_Vade değişkenine göre, ödeme vadesi 45 gün ve daha az olan bayi (1) sayısı 2 (%2,6), ödeme vadesi 45 ve 90 gün arasında olan bayi (2) sayısı 14 (%18,4), ödeme vadesi 90 gün olan bayi (3) sayısı 58 (%76,3) ve ödeme vadesi 90 günden fazla olan bayi (4) sayısı 2 (%2,6) şeklinde dağılmaktadır.

Tablo B.4: Bayi_Vade Değişkeni Frekans Tablosu

		Frekans	Yüzde	Mevcut Yüzde	Kümülatif Yüzde
Mevcut	1,00	2	2,6	2,6	2,6
	2,00	14	18,4	18,4	21,1
	3,00	58	76,3	76,3	97,4
	4,00	2	2,6	2,6	100,0
	Toplam	76	100,0	100,0	

Per_Yeterlilik (Personel sayısının yeterliliği) değişkenine ilişkin frekanslar Tablo B.5'de verilmiştir. Personel sayısı yeterli bayi (1) sayısı 26 (%34,2) ve personel sayısı yetersiz bayi (2) sayısı 50 (%65**,**8'i)'dir.

Tablo B.5: Per_Yeterlilik Değişkeni Frekans Tablosu

		Frekans	Yüzde	Mevcut Yüzde	Kümülatif Yüzde
Mevcut	1,00	26	34,2	34,2	34,2
	2,00	50	65,8	65,8	100,0
	Toplam	76	100,0	100,0	

Isyeri_Buyukluk (İşyeri büyüklüğü) değişkenine ilişkin frekanslar Tablo B.6'da verilmiştir. Isyeri_Buyukluk değişkeni; küçük bayi (1) sayısı 51 (%67,1), orta bayi (2) sayısı 12 (%15,8) ve büyük bayi (3) sayısı 13 (%17,1)'dir.

Tablo B.6: Isyeri_Buyukluk Değişkeni Frekans Tablosu

		Frekans	Yüzde	Mevcut Yüzde	Kümülatif Yüzde
Mevcut	1,00	51	67,1	67,1	67,1
	2,00	12	15,8	15,8	82,9
	3,00	13	17,1	17,1	100,0
	Toplam	76	100,0	100,0	

Mulkiyet (İşyerinin bayi mülkiyeti olup olmaması) değişkenine ilişkin frekanslar Tablo B.7'de verilmiştir. Mulkiyet değişkeni; bayinin mülkiyeti olan (1) ve kira olan (2) olmak üzere 2 değer almaktadır. Buna göre, işyeri bayinin mülkiyetinde olan (1) bayi sayısı 46 (%60,5'i) ve kira olan (2) bayi sayısı 30 (%39,5'i)'dur.

Tablo B.7: Mulkiyet Değişkeni Frekans Tablosu

		Frekans	Yüzde	Mevcut Yüzde	Kümülatif Yüzde
Mevcut	1,00	46	60,5	60,5	60,5
	2,00	30	39,5	39,5	100,0
	Toplam	76	100,0	100,0	

Analizlerde kullanılan ve firmanın Türkiye genelindeki 76 bayisinin genel yapısını ortaya koyan sürekli değişkenlerin tanımlayıcı istatistikleri Tablo B.8'de verilmiştir. Bunlar; Top_Ciro (Toplam ciro), Top_Sip_Sayi (Toplam sipariş sayısı), Teminat_Mik (Teminat miktarı), Odenmemis (Henüz ödenmemiş teminat miktarı) ve Bayi_Kredi (Bayi kredibilitesi) değişkenleridir.

Tablo B.8: Sürekli Değişkenlerin Tanımlayıcı İstatistikleri

	N	Minimum	Maksimum	Ortalama	Std. Sapma
Top_Sip_Sayi	76	1,00	525,00	139,5263	136,62186
Top_Ciro	76	134,50	1230277,71	272795,6	312043,83485
Teminat_Mik	76	95,63	657610,30	94296,08	114830,50612
Odenmemis	76	-287811	627610,30	49054,23	121588,94568
Bayi_Kredi	76	74,72	683487,62	151553,1	173357,68557
Mevcut N	76				

EK C – FAKTÖR ANALİZİ (ANKET) TABLOLARI

Tablo C.1: Ortak Varyans Tablosu

	Baslangic	Cikarim
Toplanti_Katilim	1,000	,665
Iletisim_Duzeyi	1,000	,400
Vitrin_Yeterlilik	1,000	,745
Urun_Gorsel_Oncelik	1,000	,825
Stok_Tut_Egilimi	1,000	,766
Stok_Duzen	1,000	,726
Indirim_Sip_Degisim	1,000	,692
Potansiyeli	1,000	,633
Kanal_Beceri	1,000	,585
Sip_Yaklasimi	1,000	,680
Dagitim_Agi	1,000	,706
Reklam_Faaliyeti	1,000	,635
Basarim_Duzey	1,000	,905
Hedef_Tutturma	1,000	,912
Musteri_Kazanma	1,000	,790
Musteri_Yakinlik	1,000	,739
Stand_Gorunum	1,000	,910
Musteri_Memnuniyet	1,000	,925
Geri_Bildirim	1,000	,712
Sikayet_Cozum	1,000	,713
Musteri_Odak	1,000	,673
Bilgi_Düzey	1,000	,764
Kriz_Sip_Degisim	1,000	,660

Tablo C.2: Faktör Sayısı ve Açıklanan Varyans

Faktör	Baslangic Özdeger		Cikarimlarin Kareli Toplami	
	Toplam	Varyansin %'si	Toplam	Varyansin %'si
1	8,111	35,266	8,111	35,266
2	2,023	8,797	2,023	8,797
3	1,705	7,412	1,705	7,412
4	1,445	6,284	1,445	6,284
5	1,403	6,099	1,403	6,099
6	1,047	4,550	1,047	4,550
7	1,026	4,459	1,026	4,459
8	,974	4,233		
9	,830	3,610		
10	,774	3,365		
11	,596	2,590		
12	,590	2,566		
13	,521	2,265		
14	,408	1,774		
15	,351	1,526		
16	,319	1,388		
17	,229	,997		
18	,200	,868		
19	,166	,722		
20	,140	,608		
21	,106	,460		
22	,022	,094		
23	,015	,065		

Tablo C.3: Faktör Matrisi

	Faktör						
	1	2	3	4	5	6	7
Toplanti_Katilim							
Iletisim_Duzeyi							
Vitrin_Yeterlilik				,731			
Urun_Gorsel_On						,613	
Stok_Tut_Egilimi	,548	-,512					
Stok_Duzen				-,510			,558
Indirim_Sip_Degi	,522						
Potansiyeli	,598						
Kanal_Beceri	,696						
Sip_Yaklasimi	,638						
Dagitim_Agi	,684						
Reklam_Faaliyeti	,731						
Basarim_Duzey	,694						
Hedef_Tutturma	,730						
Musteri_Kazanm	,864						
Musteri_Yakinlik		,803					
Stand_Gorunum	,693		-,540				
Musteri_Memnun	,688		-,560				
Geri_Bildirim	,797						
Sikayet_Cozum	,793						
Musteri_Odak	,702						
Bilgi_Düzey	,719						
Kriz_Sip_Degisim		,500					

Tablo C.4: Faktör Skor Katsayı Matrisi

	Faktör						
	1	2	3	4	5	6	7
Toplanti_Katilim	-,120	-,110	,317	-,027	,485	,485	-,260
Iletisim_Duzeyi	,176	,034	-,248	,356	-,283	-,112	,293
Vitrin_Yeterlilik	,219	-,136	,346	,731	,065	-,024	,138
Urun_Gorsel_Oncelik	-,037	-,376	,193	,222	-,468	,613	,042
Stok_Tut_Egilimi	,548	-,512	,319	-,174	-,100	,052	,242
Stok_Duzen	,144	,155	,283	-,510	-,158	-,065	,558
Indirim_Sip_Degisim	,522	,174	,014	,486	,274	-,082	,266
Kriz_Sip_Degisim	-,419	,500	,341	,143	,086	,166	,249
Potansiyeli	,598	-,014	-,453	,181	-,145	,039	-,121
Kanal_Beceri	,696	,163	,211	,006	,147	,042	-,078
Sip_Yaklasimi	,638	,378	,279	-,101	,110	,113	,132
Dagitim_Agi	,684	-,280	,154	,207	,044	-,298	-,061
Reklam_Faaliyeti	,731	,029	,023	-,175	-,145	,090	,198
Basarim_Duzey	,694	,396	,037	,016	-,437	,152	-,225
Hedef_Tutturma	,730	,358	,076	-,003	-,440	,093	-,209
Musteri_Kazanma	,864	,014	,068	,041	-,092	-,078	-,150
Musteri_Yakinlik	,097	,803	-,182	,094	,183	,094	,032
Stand_Gorunum	,693	-,120	-,540	-,065	,191	,218	,185
Musteri_Memnuniyet	,688	-,133	-,560	-,107	,188	,220	,158
Geri_Bildirim	,797	-,129	,080	-,002	,209	-,033	,099
Sikayet_Cozum	,793	-,233	,137	-,041	,051	-,072	-,041
Musteri_Odak	,702	-,003	,079	-,120	,352	,146	-,120
Bilgi_Düzey	,719	,052	,232	-,166	,011	-,313	-,254

EK D - FAKTÖR ANALİZİ (ÜRÜN AİLELERİ) TABLOLARI

Tablo D.1: Ortak Varyans Tablosu

	Baslangic	Çikarim
Miktar_1	1,000	,937
Miktar_2	1,000	,664
Miktar_3	1,000	,583
Miktar_4	1,000	,473
Miktar_5	1,000	,490

Tablo D.2: Faktör Sayısı ve Açıklanan Varyans

	Baslangic Özdeger			Çikarimlarin Kareli Toplami		
Faktör	Toplam	Varyansin %'si	Kümülatif %	Toplam	Varyansin %'si	Kümülatif %
1	2,130	42,606	42,606	2,130	42,606	42,606
2	1,016	20,330	62,936	1,016	20,330	62,936
3	,806	16,112	79,048			
4	,569	11,385	90,433			
5	,478	9,567	100,000			

Tablo D.3: Faktör Matrisi

	Faktör	
	1	2
Miktar_2	,793	
Miktar_3	,757	
Miktar_5	,690	
Miktar_4	,662	
Miktar_1		,960

Tablo D.4: Faktör Katsayıları Tablosu

	Faktör	
	1	2
Miktar_1	,039	,945
Miktar_2	,389	,148
Miktar_3	,344	-,130
Miktar_4	,328	,151
Miktar_5	,310	-,151

EK E - DİSKRİMİNANT ANALİZİ (ODEME_PROB)

Tablo E.1: Box's M Test Sonuçları

Box's M		94,840
F	Yaklasik	1,417
	df1	45
	df2	1276,344
	Sig.	,037

Tablo E.2: Özdeğer Tablosu

Fonksiyon	Özdeger	Varyansin %'si	Kümülatif %	Kanonik korelasyon
1	,273	100,0	100,0	,463

Tablo E.3: Wilk's Lamda Değer Tablosu

Test Fonksiyonu	Wilks' Lambda	Chi-kare	df	Sig.
1	,786	16,771	9	,052

Tablo E.4: Kanonik Diskriminant Fonksiyon Katsayıları

	Fonksiyon
	1
Genel Basari Faktör Skoru	,519
Müsteriye Ulasma Faktör Skoru	,228
Görsel Memnuniyetsizlik Faktör Skoru	,385
Görsel Etkililik Faktör Skoru	,719
Uzaktan Iletisim Faktör Skoru	,064
Ürün Öncelik Faktör Skoru	,403
Stok Düzen Faktör Skoru	,096
Genel Ürün Faktör Skoru	,023
Özel Ürün Faktör Skoru	-,018

Tablo E.5: Grup Merkez Fonksiyonları

	Fonksiyon
Odeme_Prob	1
0	-,223
1	1,190

Tablo E.6: Sınıflandırma Başarım Sonuçları

			Tahmin Edilen Grup Üyeligi		
		Odeme_Prob	0	1	Toplam
Orjinal		0	52	12	64
		1	3	9	12
	%	0	81,3	18,8	100,0
		1	25,0	75,0	100,0
Çapraz-Geçerlilik		0	50	14	64
		1	4	8	12
	%	0	78,1	21,9	100,0
		1	33,3	66,7	100,0

ÖZGEÇMİŞ

Ahmet Cüneyt BAYSAL, 10 Mayıs 1979'da Tavşanlı'da doğdu. Tavşanlı Atatürk Lisesi'nden 1997 yılında mezun oldu ve Yıldız Teknik Üniversitesi Endüstri Mühendisliği Bölümüne girdi ve 2002 yılında mezun oldu. Bir firmanın üretim planlama bölümünde Üretim Mühendisi olarak çalışmaktadır.