

**T.C.**  
**AFYON KOCATEPE ÜNİVERSİTESİ**  
**SOSYAL BİLİMLER ENSTİTÜSÜ**  
**İŞLETME ANABİLİM DALI**  
**DOKTORA TEZİ**

**ÜRÜN KATEGORİLERİ ARASINDAKİ SATIŞ  
İLİŞKİSİNİN BİRLİKTELİK KURALLARI VE  
KÜMELEME ANALİZİ İLE BELİRLENMESİ VE  
PERAKENDE SEKTÖRÜNDE BİR UYGULAMA**

**Hazırlayan**  
**Ertuğrul ERGÜN**

**Danışman**  
**Yrd. Doç Dr. Ali ELEREN**

**AFYONKARAHİSAR 2008**

Doktora tezi olarak sunduđum “rn Kategorileri Arasındaki Satıř İliřkisinin Birliktelik Kuralları ve Kmeleme Analizi İle Belirlenmesi ve Perakende Sektrnde Bir Uygulama” adlı alıřmanın tarafımdan bilimsel ahlak ve geleneklere aykırı dřecek bir yardıma bařvurmaksızın yazıldıđını ve yararlandıđım eserlerin Kaynaka’da gsterilen eserlerden oluřtuđunu, bunlara atıf yapılarak yararlanmıř olduđumu belirtir ve bunu onurumla dođrularım.

14/11/2008

Ertuđrul ERGN

**TEZ JÜRİSİ KARARI VE ENSTİTÜ ONAY SAYFASI**

JÜRİ ÜYELERİ

İMZA

**Tez Danışmanı:** Yrd. Doç. Dr. Ali ELEREN

.....

**Jüri Üyeleri:** Prof. Dr. Mahmut TEKİN

.....

Prof. Dr. Kemalettin CONKAR

.....

Yrd. Doç .Dr. Süleyman DÜNDAR

.....

Yrd. Doç Dr. Harun ÖZTÜRKLER

.....

İşletme anabilim dalı doktora öğrencisi Ertuğrul ERGÜN'ün “Ürün Kategorileri Arasındaki Satış İlişkisinin Birliktelik Kuralları ve Kümeleme Analizi İle Belirlenmesi ve Perakende Sektöründe Bir Uygulama” başlıklı tezi 14/11/2008 tarihinde, saat 14:00’de Lisansüstü Eğitim Öğretim ve Sınav Yönetmeliğinin ilgili maddeleri uyarınca, yukarıda isim ve imzaları bulunan jüri üyeleri tarafından değerlendirilerek kabul edilmiştir.

**Doç. Dr. Mehmet KARAKAŞ**  
**MÜDÜR**

**DOKTORA TEZ ÖZETİ****ÜRÜN KATEGORİLERİ ARASINDAKİ SATIŞ İLİŞKİSİNİN BİRLİKTELİK KURALLARI VE KÜMELEME ANALİZİ İLE BELİRLENMESİ VE PERAKENDE SEKTÖRÜNDE BİR UYGULAMA****Ertuğrul ERGÜN****AFYON KOCATEPE ÜNİVERSİTESİ  
SOSYAL BİLİMLER ENSTİTÜSÜ  
İŞLETME ANABİLİM DALI  
Kasım 2008****Tez Danışmanı: Yrd. Doç Dr. Ali ELEREN**

Günümüzde birçok işletme, müşterileri ile ilgili çok miktarda ve çok çeşitli veriye sahiptir. Bu verilerin çeşitli analiz teknikleriyle işlenerek anlamlı ve işletmenin işine yarayacak bilgilerin elde edilmesi, işletmenin daha etkin kararlar vermesini sağlayacak ve küreselleşen, rekabetin arttığı dünyada işletmenin rekabet etme gücünü arttıracaktır. Veri madenciliği, matematiksel ve istatistiksel teknikler yardımı ile veri seti içindeki gizli bağlantıları tespit ederek, işletmelere karar verme sürecinde yardımcı olmaktadır.

Bu çalışmada ürün kategorileri ve ürün sınıfları arasındaki satış ilişkisinin tespit edilmesi amaçlanmıştır. Bu amaçla perakendeci bir işletmenin bir yıllık süre boyunca topladığı alışveriş fişi verileri üzerinde birliktelik kuralları analizi ve hiyerarşik kümeleme gibi veri madenciliği yöntemleri uygulanmıştır. Araştırmada 1.096.125 adet fiş analiz edilmiştir. Birliktelik kuralları analizi sonucunda en büyük kural desteğinin %18,69 ile içecekler⇒tatlı ürünler bağıntısında, en yüksek güven değerinin %69,46 ile market markası⇒tatlı ürünler bağıntısında olduğu belirlenmiştir. Kümeleme analizi sonucunda en kuvvetli ilişkiler, makarna-bakliyat (0,403), zeytin-peynir (0,354), çay-şeker (0,337), pasta malzemeleri-hazır tatlı (0,308) ürün çiftleri arasında tespit edilmiştir. Kümeleme aşaması 82 adımda tamamlanmıştır.

Bu araştırmanın sonuçları perakendeci işletmeler tarafından çeşitli promosyon kampanyalarında değerlendirilebilir. İşletmenin kategori yönetiminde kullandığı ürün kategorileri ve sınıfları, kümeleme analizi sonucunda elde edilen kümelere göre tekrar oluşturularak daha verimli bir şekilde kullanılabilir. Aynı zamanda bulunan örüntüler, kararsız müşteriler için ikna edici, yön gösterici çeşitli faaliyetler tasarlanması için de kullanılabilir.

**Anahtar kelimeler: Perakende sektörü, alışveriş davranışları, veri madenciliği, birliktelik kuralları analizi, kümeleme analizi**

**ABSTRACT****DETERMINING PRODUCT CATEGORIES SALES RELATIONSHIP WITH  
ASSOCIATION RULES AND CLUSTER ANALYSIS AND AN  
APPLICATION IN RETAILING SECTOR****Ertuğrul ERGÜN****AFYON KOCATEPE UNIVERSITY  
THE INSTITUTE OF SOCIAL SCIENCES  
DEPARTMENT OF BUSINESS ADMINISTRATION****November 2008****Advisor: Assoc. Prof. Dr. Ali ELEREN**

Many enterprises have lots of data about their customers. Processing these data with various analysis techniques to make meaningful information for enterprises will make enterprises to take beneficial decisions in the competitive world. Data mining determine hidden informations and patterns in a data set with the help of mathematical and statistical techniques to help enterprises to make choices.

In this research it is aimed to determine the sales relationship between product categories and product classes. To achieve this goal, data mining methods, like association rule analysis and hierarchical cluster analysis, applied to one year sales database constructed by a retailer. In this research 1.096.125 sales slips were analyzed. In association rules analysis, the biggest rule support (%18.69) was found between beverages and sweet categories, the highest confidence value (%69.46) was found between private label and sweet categories. In cluster analysis the strongest relationship was found between macaroni-legumes (0,403), olives-cheese (0,354), tea-sugar (0,337), pastry ingredients-ready-made sweet (0,308) classes. Clustering stage was completed in 82 steps.

Results of this research can be used in promotion campaigns by retailers. Product categories and classes, which retailers use in category management, can be formed efficiently by examining cluster analysis results. Patterns, which were founded by this research, can be used to design activities for indecisive customers to persuade them.

**Keywords: Retail sector, shopping behaviours, data mining, association rule analysis, cluster analysis**

## ÖNSÖZ

Bu tez çalışmasında göstermiş olduğu yakın ilgi, destek ve yönlendirmelerinden dolayı Sayın Danışman Hocam Yrd. Doç. Dr. Ali Eleren'e; tez izleme komitesi toplantılarında yardımlarını, desteklerini esirgemeyen, yol gösteren, bu tezin en iyi şekilde hazırlanmasında emekleri olan Sayın Yrd.Doç.Dr. Süleyman Dünder'a ve Sayın Yrd.Doç. Dr. Harun Öztürkler'e; tez savunmasında değerli eleştiriyile tezi son haline getirmemi sağlayan Sayın Prof. Dr. Mahmut Tekin'e ve Sayın Prof. Dr. Kemalettin Conkar'a saygılarımı ve teşekkürlerimi sunarım.

Ayrıca tez çalışması boyunca ihmal ettiğim, bana karşı sabırlı ve anlayışlı davranışları nedeniyle eşim Selcen Süheyla'ya, kızlarım Sena ve Sude'ye; benim bu günlere gelmemde büyük emekleri olan babam Prof. Dr Mustafa Ergün ve annem Duda Ergün'e; tez çalışması sırasında eleştirileriyle yardımcı olan Uçman Ergün'e ve bilgisayarımı uzun bir süre elinden aldığım Bircan Ergün'e sonsuz teşekkürlerimi sunarım.

Ertuğrul Ergün

## İÇİNDEKİLER

Sayfa

YEMİN METNİ.....	ii
TEZ JÜRİSİ VE ENSTİTÜ MÜDÜRÜ ONAYI .....	iii
ÖZET .....	iv
ABSTRACT .....	v
ÖNSÖZ.....	vi
İÇİNDEKİLER.....	vii
ŞEKİLLER LİSTESİ .....	x
TABLolar LİSTESİ .....	Xi
GİRİŞ .....	1

### BİRİNCİ BÖLÜM

#### PERAKENDECİLİK SEKTÖRÜ

1. GENEL BİLGİLER, TANIM VE KAVRAMLAR.....	4
2. PERAKENDECİ İŞLETMELERİN SINIFLANDIRILMASI.....	6
3. SÜPERMARKETLER .....	7
4. TÜRKİYE'DE PERAKENDE SEKTÖRÜ.....	10
5. PERAKENDE SEKTÖRÜ VE BİLGİ TEKNOLOJİLERİ .....	17
6. KATEGORİ YÖNETİMİ VE TANZİM-TEŞHİR KAVRAMLARI.....	18

### İKİNCİ BÖLÜM

#### VERİ, VERİTABANI VE VERİ MADENCİLİĞİ

1. VERİ, BİLGİ, VERİTABANI KAVRAMLARI .....	28
1.1. VERİ VE BİLGİ .....	28
1.2. VERİTABANI VE VERİ AMBARI.....	29
1.2.1. Veri Ambarı Yapısı .....	32
1.2.2. Veri Ambarı Tipleri .....	36
1.2.3. Veritabanı ve Veri Ambarı Arasındaki Farklar .....	37
2. VERİ MADENCİLİĞİ .....	38
2.1. VERİ MADENCİLİĞİNİN İŞLEVLERİ.....	42
2.1.1. Keşif .....	42
2.1.2. Tahmini modelleme .....	43
2.1.3. Adli analiz.....	43

2.2. VERİ MADENCİLİĞİ SÜRECİ.....	44
2.2.1. Çalışmanın, Araştırmanın Planlanması (İş anlama).....	45
2.2.2. Verilerin Anlama.....	46
2.2.3. Verilerin Hazırlanması.....	46
2.2.4. Modelleme aşaması.....	47
2.2.5. Değerlendirme aşaması.....	48
2.2.6. Uygulama aşaması.....	48
2.3. VERİ MADENCİLİĞİ MODELLERİ.....	49
2.3.1. Sınıflama ve Regresyon.....	50
2.3.2. Kümeleme.....	52
2.3.3. Birliktelik Kuralları ve Ardışık Zamanlı Örüntüler.....	55
2.3.4. Tahmine Dayalı Karar Verme.....	57
2.3.5. Ayırma.....	57
2.3.6. İstisnalar (Fark Saptaması).....	58
2.3.7. Metin Madenciliği.....	58
2.3.8. Web Madenciliği.....	58
2.4. VERİ MADENCİLİĞİNDE KULLANILAN YÖNTEMLER.....	59
2.4.1. Karar Ağaçları.....	59
2.4.2. Yapay Sinir Ağları.....	61
2.4.3. Genetik Algoritmalar.....	63
2.4.4. k - En Yakın Komşu.....	64
2.4.5. Bellek Tabanlı Yöntemler.....	64
2.4.6. Geleneksel Sorgu Araçları.....	65
2.4.7. Görsel Yöntemler.....	66
2.4.8. Sapan Değer Çözümlemesi.....	66
2.4.9. Evrimsel çözümleme.....	66
2.4.10. Naive-Bayes.....	66
2.5. VERİ MADENCİLİĞİNİN UYGULAMA ALANLARI.....	67
2.6. LİTERATÜR TARAMASI.....	69

## ÜÇÜNCÜ BÖLÜM

### BİRLİKTELİK KURALLARI ANALİZİ VE KÜMELEME ANALİZİ

1. BİRLİKTELİK KURALLARI ANALİZİ.....	74
1.1. BİRLİKTELİK KURALLARINDA KULLANILAN ALGORİTMALAR.....	80
1.1.1. AIS Algoritması.....	80
1.1.2. SETM Algoritması.....	81
1.1.3. Apriori Algoritması.....	82
1.1.4. AprioriTid Algoritması.....	88
1.1.5. Diğer Algoritmalar.....	89
1.2. BİRLİKTELİK KURALLARI HESAPLAMA YÖNTEMLERİ.....	90
1.2.1. Piatetsky-Shapiro.....	91



1.2.2. Brin-Motwani-Silverstein.....	91
1.2.3. Ki-Kare Analizi İle Birliktelik Kurallarının Değerlendirilmesi .....	93
1.3 . PAZAR SEPET ANALİZİ.....	96
1.4. ÖRNEK BİR PAZAR SEPETİ ANALİZİ.....	99
<b>2. KÜMELEME ANALİZİ .....</b>	<b>101</b>
2.1. UZAKLIK ÖLÇÜLERİ .....	103
2.2. KÜMELEME YÖNTEMLERİ .....	104
2.2.1. Hiyerarşik Kümeleme.....	104
2.2.2. Hiyerarşik Olmayan Kümeleme .....	107

## DÖRDÜNCÜ BÖLÜM

### BİR SÜPERMARKETİN SATIŞ FİŞLERİ ÜZERİNDE ÜRÜNLERİN SATIŞINDA BİRLİKTELİK KURALLARI VE KÜMELEME ANALİZLERİ UYGULAMASI

<b>1. ARAŞTIRMANIN AMACI, ÖNEMİ VE KAPSAMI.....</b>	<b>109</b>
1.1. AMAÇ.....	109
1.2. ÖNEM .....	109
1.3. KAPSAM.....	110
<b>2. MODELİN OLUŞTURMASI VE UYGULANMASI.....</b>	<b>110</b>
2.1. ARAŞTIRMANIN OLUŞTURULMASI.....	110
2.2. VERİLERİN OLUŞTURULMASI.....	111
2.3. VERİLERİN HAZIRLANMASI.....	111
2.4. MODELLEME.....	116
2.5. DEĞERLENDİRME.....	117
2.5.1. Kategoriler Arasındaki Birlikte Satış İlişkilerinin Grafikselsel Olarak Gösterilmesi.....	117
2.5.2. Kategoriler Arasındaki İkili Birlikte Satış İlişkilerinin İncelenmesi	120
2.5.3. Kategoriler Arasındaki Çoklu Birlikte Satış İlişkilerinin İncelenmesi .....	125
2.5.4. Kategorilerin Birlikte Satış İlişkilerinin Tek Tek İncelenmesi.....	129
2.5.5. Ürün Sınıfları Arasındaki Kümelenemenin Belirlenmesi.....	149
2.5.6. Ürün Sınıfları Arasındaki Birliktelik Kurallarının Belirlenmesi .....	157
<b>SONUÇ.....</b>	<b>165</b>
<b>KAYNAKÇA .....</b>	<b>170</b>
<b>EKLER.....</b>	<b>182</b>
<b>ÖZGEÇMİŞ.....</b>	<b>1961</b>

## ŞEKİLLER LİSTESİ

Şekil 1. Türkiye Perakende Sektörünün Toplam Ciro Grafiği .....	11
Şekil 2. Planogram Örneği (Mağaza Genel Yerleşimi) .....	25
Şekil 3. Planogram Örneği (Ürün Yerleşimi-şema) .....	25
Şekil 4. Planogram Örneği (Ürün Yerleşimi-fotoğraf) .....	26
Şekil 5. Veri Ambarlarının Oluşturulması .....	30
Şekil 6. Veri Madenciliği Aktiviteleri .....	42
Şekil 7. CRISP-DM Metodoloji Döngüsü .....	45
Şekil 8. Veri Madenciliği Aşamaları Arasındaki İlişki .....	49
Şekil 9. Veri Madenciliği Modelleri .....	50
Şekil 10. Kümeleme Analizi Örneği (İnal,2008:26) .....	54
Şekil 11. Karar Ağacı Örneği .....	60
Şekil 12. Apriori Algoritmasının Adımları.....	87
Şekil 13. Örnek bir dendogram ve küme grafiği .....	106
Şekil 14. Alışveriş Merkezinde Kullanılan Ürün Kategorileri .....	113
Şekil 15. Clementine Programında Kullanılan Model Görüntüsü .....	116
Şekil 16. Kategoriler Arası Birlikte Satış Bağlantılarını Gösteren Web Grafiği .....	119
Şekil 17. Kişisel Bakım Kategorisi İlişki Grafiği .....	130
Şekil 18. Kültür-Hobi Kategorisi İlişki Grafiği.....	131
Şekil 19. Tamamlayıcı ürünler Kategorisi İlişki Grafiği .....	133
Şekil 20. Züccaciye Kategorisi İlişki Grafiği .....	134
Şekil 21. İçecekler Kategorisi İlişki Grafiği .....	135
Şekil 22. Kuru Gıda Kategorisi İlişki Grafiği .....	137
Şekil 23. Market Markaları Kategorisi İlişki Grafiği.....	138
Şekil 24. Evcil Hayvanlar Kategorisi İlişki Grafiği .....	139
Şekil 25. Tatlı Ürünler Kategorisi İlişki Grafiği .....	141
Şekil 26. Temizlik Kategorisi İlişki Grafiği.....	142
Şekil 27. Açık Unlu Mamul Kategorisi İlişki Grafiği .....	143
Şekil 28. Açık Şarküteri Kategorisi İlişki Grafiği.....	145
Şekil 29. Etler Kategorisi İlişki Grafiği .....	146
Şekil 30. Sebze-meyve Kategorisi İlişki Grafiği .....	147
Şekil 31. Ürün Sınıfları İlişkilerini Gösteren Dendogram .....	154
Şekil 32. Analiz sonucunda ürün sınıflarının oluşturduğu küme grafiği .....	156

## TABLOLAR LİSTESİ

<i>Tablo 1. Türkiye’de Perakende Sektöründe Faaliyet Gösteren İşletmelerin Sayıların Yıllara Göre Dağılım</i> .....	13
<i>Tablo 2. Türkiye’de Gıda Perakendecilerinin Pazar Payları (%)</i> .....	13
<i>Tablo 3. Teknolojik Gelişmeler Işığında Veri Analiz Tekniklerinin Gelişimi</i> .....	41
<i>Tablo 4. A, B ve C Nesneleri Arasındaki Birliktelik Kuralları</i> .....	76
<i>Tablo 5. Nesne Sayısına Göre Ortaya Çıkabilecek Birliktelik Kuralları Sayısı</i> .....	76
<i>Tablo 6. Örnek Alışveriş Sepetleri</i> .....	77
<i>Tablo 7. Alışveriş Numaraları ve Alınan Ürünlerin Tablosu</i> .....	84
<i>Tablo 8. Çay-Kahve Satışları Tablosu</i> .....	93
<i>Tablo 9. Ki-kare Analizi İçin Örnek Veri Dağılım Tablosu</i> .....	95
<i>Tablo 10. I(r) Değerleri Tablosu</i> .....	96
<i>Tablo 11. Bir E-Ticaret Sitesinin Alışveriş Kayıt Örneği</i> .....	100
<i>Tablo 12. Analiz Sonucunda Elde Edilen Birliktelik Kuralları</i> .....	100
<i>Tablo 13. Aylara Göre Fiş ve Hareket Miktarının Dağılımı</i> .....	112
<i>Tablo 14. Kategorilerin Hareketlerdeki Görülme Sayılarının Aylara Göre Dağılımı</i> .....	114
<i>Tablo 15. Kategorilerin Fişlerde Görülme Sayılarının Aylara Göre Dağılımı</i> .....	115
<i>Tablo 16. Kategorilere Uygulanmış Apriori Modelinin Sonuçları (Kural Desteğine Göre Sıralı)</i> .....	120
<i>Tablo 17. Kategorilere Uygulanmış Apriori Modelinin Sonuçları (Güven Değerine Göre Sıralı)</i> .....	121
<i>Tablo 18. Kategorilerdeki Birliktelik Kurallarının Ki-Kare Tablosu</i> .....	122
<i>Tablo 19. Tatlı ürünler &amp; İçecekler İlgili Değerleri Tablosu</i> .....	123
<i>Tablo 20. Kuru gıda-tatlı ürünler İlgili Değerleri Tablosu</i> .....	123
<i>Tablo 21. İçecekler- kuru gıda İlgili Değerleri Tablosu</i> .....	124
<i>Tablo 22. Açık Şarküteri-tatlı ürünler İlgili Değerleri Tablosu</i> .....	124
<i>Tablo 23. Açık Şarküteri-içecekler İlgili Değerleri Tablosu</i> .....	124
<i>Tablo 24. Açık Şarküteri-kuru gıda İlgili Değerleri Tablosu</i> .....	125
<i>Tablo 25. Kategorilere Uygulanmış Çoklu GRI Modelinin Sonuçları (Kural Desteğine Göre Sıralı)</i> .....	125
<i>Tablo 26. Kategorilere Uygulanmış Çoklu GRI Modelinin Sonuçları (Güven Değerine Göre Sıralı)</i> .....	126
<i>Tablo 27. Kategorilerdeki Çoklu Birliktelik Kurallarının Ki-Kare Tablosu</i> .....	126
<i>Tablo 28. İçecekler&amp;Kuru Gıda &amp; Tatlı Ürünler Bağıntısı İçin İlgili Tablosu</i> .....	127
<i>Tablo 29. Kuru Gıda &amp; Tatlı Ürünler &amp; İçecekler Bağıntısı İçin İlgili Tablosu</i> .....	127
<i>Tablo 30. İçecekler&amp;Tatlı Ürünler &amp; Kuru Gıda Bağıntısı İçin İlgili Tablosu</i> .....	127
<i>Tablo 31. Tatlı Ürünler&amp;Açık Şarküteri &amp; Kuru Gıda Bağıntısı İçin İlgili Tablosu</i> .....	128
<i>Tablo 32. Kuru Gıda&amp;Açık Şarküteri &amp; Tatlı Ürünler Bağıntısı İçin İlgili Tablosu</i> .....	128
<i>Tablo 33. Kuru gıda&amp;Tatlı Ürünler &amp; Açık Şarküteri Bağıntısı İçin İlgili Tablosu</i> .....	128
<i>Tablo 34. İçecekler&amp;Açık Şarküteri &amp; Tatlı Ürünler Bağıntısı İçin İlgili Tablosu</i> .....	129
<i>Tablo 35. Açık Şarküteri&amp;Tatlı Ürünler &amp; İçecekler Bağıntısı İçin İlgili Tablosu</i> .....	129
<i>Tablo 36. Tatlı Ürünler&amp;İçecekler &amp; Açık Şarküteri Bağıntısı İçin İlgili Tablosu</i> .....	129

<i>Tablo 37. “Kişisel Bakım” Kategorisi Birliktelik Kuralları Tablosu.....</i>	<i>131</i>
<i>Tablo 38. “Kültür-Hobi” Kategorisi Birliktelik Kuralları Tablosu.....</i>	<i>132</i>
<i>Tablo 39. “Tamamlayıcı Ürünler” Kategorisi Birliktelik Kuralları Tablosu .....</i>	<i>133</i>
<i>Tablo 40. “Züccaciye” Kategorisi Birliktelik Kuralları Tablosu .....</i>	<i>135</i>
<i>Tablo 41. “İçecekler” Kategorisi Birliktelik Kuralları Tablosu.....</i>	<i>136</i>
<i>Tablo 42. “Kuru Gıda” Kategorisi Birliktelik Kuralları Tablosu .....</i>	<i>137</i>
<i>Tablo 43. Market Markaları Kategorisi Birliktelik Kuralları Tablosu .....</i>	<i>139</i>
<i>Tablo 44. Evcil Hayvanlar Kategorisi Birliktelik Kuralları Tablosu.....</i>	<i>140</i>
<i>Tablo 45. Tatlı ürünler Kategorisi Birliktelik Kuralları Tablosu .....</i>	<i>141</i>
<i>Tablo 46. Temizlik Kategorisi Birliktelik Kuralları Tablosu .....</i>	<i>143</i>
<i>Tablo 47. Açık Unlu Mamul Kategorisi Birliktelik Kuralları Tablosu.....</i>	<i>144</i>
<i>Tablo 48. Açık Şarküteri Kategorisi Birliktelik Kuralları Tablosu .....</i>	<i>145</i>
<i>Tablo 49. Etlер Kategorisi Birliktelik Kuralları Tablosu .....</i>	<i>147</i>
<i>Tablo 50. Sebze-Meyve Kategorisi Birliktelik Kuralları Tablosu .....</i>	<i>148</i>
<i>Tablo 51. Kategoriler Arasındaki İlgi Değerleri .....</i>	<i>149</i>
<i>Tablo 52. Ürün Sınıflarının Fişlerde Görülme Frekansları Ve Yüzdeleri .....</i>	<i>150</i>
<i>Tablo 53. Ürün Sınıfları Arasındaki Kümeleme Aşamalarını Gösteren Toplu Tablo .....</i>	<i>152</i>
<i>Tablo 54. Ürün Sınıfları Arasındaki Birliktelik Kuralları (Kural Desteğine Göre Sıralı) .....</i>	<i>157</i>
<i>Tablo 55. Ürün Sınıfları Arasındaki Birliktelik Kuralları (Güven Değerine Göre Sıralı) .....</i>	<i>157</i>
<i>Tablo 56. Ürün Sınıfları Arasındaki Birliktelik Kuralları (Kaldırma Oranı Değerine Göre Sıralı).....</i>	<i>158</i>
<i>Tablo 57. Ürün Sınıfları Arasındaki Birliktelik Kuralları Tablosu.....</i>	<i>159</i>

## GİRİŞ

Günümüz ekonomilerinde işletmelerin teknoloji ve bilgisayar kullanımlarının artmasıyla birlikte müşteri ve satış bilgileri ve verileri elektronik ortamda depolanabilmektedir.

Müşteri ve satış verilerinin elektronik ortamda tutulabilmesi, bu verilerin kolayca belli amaçlara yönelik kullanılmasını gündeme getirmektedir. Özellikle müşteri bölümlerinin geniş olduğu sektörlerde oluşturulan müşteri veri tabanlarının, işletme amaçları ve politikaları çerçevesinde kullanılmasının oldukça önemli hale geldiği görülmektedir.

Yeni teknolojilerden etkilenen ve dijital ekonomi olarak adlandırılan bu yenedünyada bilgi ve zaman boyutlarının önemi artmaktadır. Buna bağlı olarak işletmelerin doğru ve anlamlı bilgiye dayalı, hızlı karar alma gereği her zamankinden daha fazla ön plana çıkmaktadır.

Günümüzde birçok işletme, müşterileri ile ilgili çok miktarda ve çok çeşitli veriye sahiptir. Bu verilerin çeşitli analiz teknikleriyle işlenerek anlamlı ve işletmenin işine yarayacak bilgilerin elde edilmesi, işletmenin daha etkin kararlar vermesini sağlayacak ve küreselleşen, rekabetin arttığı dünyada işletmenin rekabet etme gücünü arttıracaktır. Bilgisayar sistemleri ile üretilen veriler tek başlarına değersizdir, çünkü çıplak gözle bakıldığında bir anlam ifade etmezler. Bu veriler belli bir amaç doğrultusunda işlendiği zaman bir anlam ifade eder. Bu yüzden büyük miktardaki verileri işleyebilen teknikleri kullanabilmek büyük önem kazanmaktadır. Bu ham veriyi bilgiye dönüştürme işlemleri veri madenciliği ile yapılabilmektedir. Veri madenciliği veri setlerinde saklı durumda bulunan örüntü ve eğilimleri keşfetme işlemidir.

Eskiden süpermarketlerde basit toplama makinesinden ibaret olan kasalar, müşterinin o anda satın aldığı ürünlerin toplamını hesaplamak için kullanılmaktaydı. Günümüzde ise kasa yerine kullanılan satış noktası terminalleri sayesinde satın alma işlemlerinin bütün detayları saklanabilmektedir. Kaydedilen büyük miktardaki veri,

veri madenciliği yöntemleriyle analiz edilebilir. Böylece veri seti içindeki gizli bağlantılar tespit edilerek, işletmelere karar verme sürecinde yol gösterilebilir. Perakende işletmelerinde müşterilerin demografik özellikleri ve alışverişleri ile ilgili veriler (işlem saati, alınan ürün, miktarı, tutarı, vb.) müşteri bazında veya ürün bazında kaydedilmekte, veriler üzerinde analizler yapılarak müşterinin satın alma tutum ve davranışları hakkında bilgiler üretilebilmektedir. Örneğin işletmeler her müşteriye bir kod numarası verip, müşterinin tüm alışveriş hareketlerini izleyebilmekte ve kayıt edebilmekte, bu verileri veri madenciliği teknikleriyle analiz ederek, müşteri kârlılığı ve sadakati konusunda yeni bilgilere ve örüntülere ulaşabilmekte ve buna göre pazarlama stratejileri geliştirilebilmektedir.

Bu çalışmada, perakende sektöründe ürünlerin birlikte satış ilişkisinin, veri madenciliği yöntemleri kullanılarak belirlenmesi amaçlanmıştır. Veri madenciliği sürecinde CRISP-DM süreç modeli kullanılmıştır. Bu model verilerin toplanması, verilerin analizi, sonuçların yorumlanması, düzenlemelerin yapılması ve sonuçların izlenmesi aşamalarında veri madenciliği projelerde kullanılabilecek bazı genel çerçeveleri içermektedir. Veriler SPSS Clementine yazılımında Apriori ve GRI (Generalized Rule Indiction) ve SPSS 13.0 yazılımında hiyerarşik kümeleme analizleriyle incelenmiş ve ağ (web) grafikleri, birliktelik kuralları ilişkileri ve kümeleme tablo-grafikleri elde edilmiştir.

Birinci bölümde perakende sektörü hakkında bilgiler verilmiş ve perakendeci işletmelerin sınıflandırılması genel olarak açıklanmıştır. Araştırmanın da temelini oluşturan süpermarket türü perakendeciler de bu bölümde incelenmiştir. Daha sonra perakendecilikte bilgi teknolojilerinin kullanımı, pazar sepet analizi, kategori yönetimi ve tanzim-teşhir konularında bilgilere yer verilmiştir.

İkinci bölümde veri, bilgi, bilgi sistemleri, veri ambarı ve veri tabanı kavramları açıklanmıştır. Veri madenciliği kavramı açıklanıp, veri madenciliği süreci, veri madenciliği modelleri ve veri madenciliği teknikleri incelenmiştir. Aynı zamanda veri madenciliği yöntemleri kullanılarak yapılmış çalışmalar incelenmiştir.

Üçüncü bölümde birliktelik kuralları (association rules) analizi hakkında temel bilgiler yer verilmiş, birliktelik kuralları analizi için kullanılan algoritmalar ve birliktelik kurallarının hesaplama yöntemleri incelenmiştir. Daha sonra kümeleme

analizi açıklanmış ve kümeleme analizi için kullanılan yöntemler ve uzaklık ölçüleri incelenmiştir.

Dördüncü bölümde de Afyonkarahisar il merkezinde faaliyet gösteren bir perakendeci işletmenin 2007 yılı satış fişleri veritabanı kullanılarak, ürünlerin birlikte satış ilişkileri belirlenmiştir. Elde edilen bu bulgular doğrultusunda, müşterilerin alışverişi sırasında hangi ürün sınıflarının, hangi ürün sınıflarının satışını olumlu yönde etkileyebileceği konusunda değerlendirmelerde bulunulmuştur. Sonuç bölümünde ise araştırma bulguları değerlendirilmiş, araştırma sonuçlarının işletmecilik biliminde ve perakendecilik sektöründe nasıl kullanılabilirliği yönünde önerilerde bulunulmuştur. Son olarak ise gelecek araştırmaların planlanmasına ilişkin önerilere yer verilmiştir.

## **BİRİNCİ BÖLÜM**

### **PERAKENDECİLİK SEKTÖRÜ**

#### **1. GENEL BİLGİLER, TANIM VE KAVRAMLAR**

Tüketicilere gereksinim ve istekleri nedeniyle satın almak istedikleri malların ulaştırılmasından, pazarlama kanalındaki aracılar sorumludur. Aracılar, üreticiler ile tüketiciler arasında köprü işlevi görmekte ve bu arada da modern işletmecilik anlayışı uyarınca kendi amaçlarına ulaşmaya çalışmaktadırlar. Aracılardan, gereksinimleri nedeniyle alışveriş eden tüketiciye mal ya da hizmetlerin satışından sorumlu olan kanal perakendecilerdir. Perakendecilik, mal ve hizmetlerin, ticari bir amaç için kullanılmaksızın, tüketicilerin gereksinimlerinin karşılanması amacıyla, doğrudan doğruya son tüketiciye pazarlanmasıyla ilgili faaliyetler bütünüdür (Tek, 1984:28). Mucuk (1994: 256) ise perakendeciliği, mal ya da hizmetlerin doğrudan doğruya son kullanıcılara ya da tüketicilere pazarlanması ile ilgili etkinlikler bütünü olarak tanımlamaktadır.

Faaliyet alanları açısından farklılık gösteren perakendeci işletmeler değer oluşturmayı hedeflerler. Oluşturulan bu değerler fayda olarak isimlendirilir. Perakendeciler üreticilerin ürün ve hizmetlerini tüketiciye sunarak zaman, yer, sahiplik ve şekil faydası meydana getirirler. Perakendeciler tüketicilerin ihtiyaçları olduğu zaman sunulmak üzere ürünleri stoklayarak zaman faydasını oluştururlar. Perakendeci üreticilerden ya da toptancılardan satın aldığı ürünleri taşıyarak ve depolayarak yer faydasının yanı sıra zaman faydası da yaratmaktadır. Perakendeciler tarafından ürünlerin tüketicilere satılması ile yaratılan fayda ise sahiplik faydası



olarak tanımlanmaktadır. Ürünlerin sahipliği perakendecilerden satın alma yolu ile tüketicilere geçmekte ve böylece perakendeci sahiplik faydası yaratmaktadır. Son olarak perakendeciler ürünleri geliştirerek ve ürünlere yeni değerler ekleyerek şekil faydası yaratmaktadır (Timur, Öztürk ve Oyman, 1996: 78).

Perakendecilik, tarihi seyri içerisinde seyyar satıcılıktan, organize ve büyük perakendeciliğe uzanan bir gelişme göstermiştir. Ara aşamalardan biri olan geleneksel perakendeciliğin genel özellikleri, buldukları mahalli pazar bölümüne hitap etmeleri, mal bileşimlerinin sınırlı olması ve bu malların düşük hizmet düzeyi ve yüksek kâr marjları ile satılmasıdır. Geleneksel perakendecilikte genellikle bakkaliye malları ve bazen bu tür malların yanında bir kısım dayanıksız tüketim malları yer almaktadır. Son aşamalardan biri olan organize perakendecilikte ise gıda perakendeciliği başta olmak üzere birçok alanda büyük ölçekli, modern perakendeci teknikleri ile desteklenmiş perakende mağazaların ortaya çıktığı görülmektedir (Güllü, 2005: 5-6).

Aksulu'ya (2002: 1) göre perakendeci işletmelerin sahip olduğu genel özellikler şunlardır:

1. Perakendeci işletmeler faaliyetlerini genellikle sabit mekânlarda gerçekleştirirler. Mağazasız perakendecilikle ilgili uygulamalar, teknolojinin ucuzlaması, kullanım kolaylığı ve dolayısıyla teknolojinin yaygınlaşması sayesinde gün geçtikçe gelişmekte ve mağazasız perakendeciliğin perakendecilikteki payı artmaktadır.

2. Perakendeciler halkı satın almaya yöneltici ve mağazaya çekmeye yöneltici etkinlikler ile uğraşırlar. Perakendeci, genellikle son satın alıcı ve tüketici ile ilişkide bulunan dağıtım kanalı halkası olarak müşteri elde etme, mevcut müşterileri elde tutma ve mağaza bağlılığı yaratmaya yönelik çeşitli promosyonel eylemlerde bulunmak zorundadır. Artan rekabet, perakendecilerin bu tür faaliyetlerinin geçmişe göre çok daha fazla yoğunlaşmasına neden olmaktadır.

3. Perakendeciler ticari malları çoğunlukla üretici/tedarikçilerden alırlar ve satarlar. Ancak organize perakendeciliğin gelişmesiyle perakendecilerin kendi ürünlerini üretebilir/ürettirebilir duruma geldikleri görülmektedir. Birçok büyük perakendeci bazı kategorilerdeki ürünlerini üretebilmek için kendi üretim

merkezlerini oluřturmakta ya da fason olarak küçük ya da ulusal üreticiler ile çalışarak üretme yoluna gitmektedir. Bu kuruluşlar bir kısım malı kendileri üretseler de ticaret içinde adları perakendeci olarak geçmektedir.

4. Perakendeciler, müşterilere tek tek ya da küçük miktarlarda satış yapan işletmelerdir. Ancak günümüzde perakendeci işletme hacimlerinin büyümesi, bu işletmelerin başka ticari kuruluşlara da büyük ölçüde satış yapmalarına olanak sağlamıştır. Bu anlamda bazı büyük perakendecilerin bir anlamda toptancı gibi çalıştığı görülmektedir.

## **2. PERAKENDECİ İŐLETMELERİN SINIFLANDIRILMASI**

Perakendeci işletmeler çeşitli özelliklerine göre değişik şekillerde sınıflandırılmaktadır. Sınıflandırmada daha çok genel ölçek büyüklüğü, örgütsel yapı ve uygulanan yöntemler kullanılmaktadır.

Varinli (2005: 8-11) perakendeci işletmeleri büyüklük esasına göre, satılan mal türüne göre, mülkiyet durumuna göre, örgütsel yapı ve uyguladıkları yöntemlere göre ve satış yöntemlerine göre olmak üzere 5 farklı kritere göre sınıflandırmaktadır.

Perakendeci işletmeler aynı zamanda faaliyet yönetimine göre, iki açıdan incelenebilmektedir. Mağazalı perakendecilik ve mağazasız perakendecilik. Mağazalı ve mağazasız perakendecilik arasındaki temel farklılık, mağazalı perakendecinin hedef noktadaki tüketiciye fiilen sahip olduđu bir satış noktasından hitap etmesi, mağazasız perakendecinin ise satış faaliyetini yürüttüğü sabit bir mekânın olmamasıdır.

Mağazalı perakendecilikteki işletme türlerine indirimli mağazalar, süperetler, süpermarketler, hipermarketler ve alışveriş merkezleri örnek olarak gösterilebilir. Katalog perakendeciliđi, telefonla satış, elektronik ticaret, doğrudan satış ve elektronik makineler de mağazasız perakendecilik uygulamalarındandır.

### 3. SÜPERMARKETLER

Bu arařtırmada verilerin alındığı iřletme mađazalı perakendeciliđin bir alt sınıfı olan süpermarket sınıfındadır. Etimolojik yönden “üstün pazar” anlamına gelen süpermarket, mađaza satıř alanı, satıř hacmi ve yönetim biçimleri bakımından, aynı tür ticareti yapan, alışılmıř geleneksel gıda maddesi perakendecilerinden daha çok büyük ve farklı bir tasarımda olan perakendeci türüdür (Tek, 1973’den aktaran Demir, 2007: 19).

Süpermarketler self servis anlayıřla çalıřan, en az 1000 m<sup>2</sup> satıř alanına sahip, gıda ürünlerinin ađırlıklı olduđu perakendecilerdir. Süpermarketlerde satılan ürünlerin çođunu gıda vb. temel ihtiyaç maddeleri oluřturmaktadır.

Bennett (1995: 279) süpermarketleri, kuru bakkaliye, taze et, kolay bozulabilen ürünler ve süt ürünleri gibi ürünlerinin tamamının yanında uygun bir çeřitlilikle bulunan gıda dıřı ürünleri oldukça geniş biçimde sunan, genişçe bölümlendirilen ve self-servis esasına göre iřletilen perakende kurumlar olarak tanımlamaktadır. Süpermarketler 4 ana sınıfa ayrılabilir: klasik süpermarketler (conventional), süper süpermarketler (superstore), birleřik mađazalar (combination) ve gıda ticaret merkezleri (warehouse) (Güllü, 2005: 14). Bu süpermarketler arasındaki temel farklılıklar, mađaza büyüklükleri ve bulduklarını ürün karmasıdır.

Klasik süpermarketler, 1.500-3.000 m<sup>2</sup> büyüklükte olup, gıdaya dayalı ürünlerde uzmanlırlar (satıřlarının %5’den azı genel ürün satıřlarından gelmektedir).

Süper süpermarketler, 3.000-6.500 m<sup>2</sup> büyüklükte dirler ve geniş bir genel ürün seçeneđi sunarlar. Satıřlarının %20 ile %25 arası genel ürünlerden elde edilir. Klasik süpermarketlere kıyasla süper süpermarketler, genellikle daha geniş bir ticari alana sahiptirler ve bir defada çok miktarda alışveriř yapma kolaylığı sunarlar.

Birleřik mađazalar, bir süpermarket ve bir eczane ya da, bir süpermarket ve bir genel ürün perakendecisi biçiminde tek bir mađazada birleřtirilirler. Süper süpermarketlerden farklı olarak, birleřik mađazada alanın %25’i gıda dıřı ürünlere tahsis edilir. Birleřik mađaza birleřen birimler arasında bir sinerjiye sahiptir. Genel ürün perakendecisi veya eczane, süpermarket müşterilerinin sıklığından faydalanır.

Süpermarket ise, ticari alandaki artış ve daha yüksek ortalama satış dolayısıyla genel ürün perakendecisi ve eczanenin daha yüksek kâr marjlarından fayda sağlar.

Gıda ticaret merkezlerinin büyüklüğü değişmektedir. Büyüklük ölçülerine bağlı olarak, çeşitli biçimlerde sınıflandırılabilirler. Gıda ticaret merkezleri düşük fiyatla satış yaparlar. Gıda ticaret merkezleri, düşük yerleşim maliyetlerinden, düşük sabit maliyetlerden ve uygun satın almadan yararlanarak faaliyet maliyetlerini azaltırlar.

Süpermarketler geniş bir tüketici kitlesine hizmet sunmaktadır. Başta gıda olmak üzere, küçük mutfak eşyalarından kırtasiyeye, oyuncaktan giysiye değin uzanan geniş bir ürün yelpazesi vardır. Bu özelliğinden dolayı tüketicilere tek duraklı alışveriş imkânı sunmaktadır. Satış fiyatları düşük olması nedeniyle tüketiciye para, zaman ve enerjiden tasarruf sağlayan mekânlar oldukları söylenebilir.

Güllü'ye (2005: 15) göre süpermarket tanımları yapılırken esas alınan ölçülerin farklılıklar arz ediyor olması ortak bir tanımın geliştirilmesini zorlaştırmaktadır. Ekonomik gelişmişlik düzeyinin, bir ülkede bulunan süpermarketin mal bileşimini, çeşidini ve sayısını etkiliyor olması, mal bileşimini ortak bir kriter olarak almayı zorlaştırmaktadır. Bunun yanında satış cirosunun da ülkelerin gelişmişlik düzeyine göre farklılık gösterdiği, hatta aynı ülkede bölgeden bölgeye değişebileceği gibi işletmeden işletmeye de değişebilmesi, bunun da ortak bir kriter olarak kabul edilmesine imkân tanımamaktadır.

Süpermarketlerin temel özelliklerini şu şekilde sıralamak mümkündür (Timur vd., 1996:91).

- Temel olarak bakkaliye, kuru gıda, yaş sebze ve meyve, et ve et ürünleri, süt ve süt ürünlerine ek olarak kozmetik, temizlik, tuhafiyeye malzemelerinden oyuncağa değin gıda dışı ürünlerin de satılması,
- Her grup ürünün ayrı reyonlarda, çok çeşit ve sayıda satışa sunulması,
- Self servis yöntemi ile ürünlerin raflarda, açık vitrinlerde satılması,
- Genellikle tek katlı ve geniş bir satış alanına sahip olması,
- Otopark olanağı,

- Ürünlerin mümkün olduğunca üreticiden aracı olmaksızın doğrudan satın alınarak satışa sunulması,
- Büyük ölçekli süpermarket zincirlerinin kendi markalarını taşıyan ürünler satması.

Süpermarketler için yapılan tanımlamalarda sıkça yer alan ve bu perakendeci türünün tarifinde karakteristik bir niteliğe sahip olan self-servis, süpermarketlerin bir perakendeci türü olarak ortaya çıkmasını ve atılım yapmasını sağlamıştır. Self-servis, müşteri yardımcı aramadıkça, satış elemanının yardımı olmadan müşterinin kontrol edebilmesi, dokunabilmesi için, ürünün sergilendiği bir faaliyet şeklidir. Müşteriye, sadece hesaplamada/tartmada veya bazı paketleme işlemlerinin yapıldığı yerlerde refakat edilir. Self-servis yöntemi müşterilerin malları mağaza içerisinde raflardan ya da sergilenen bölümlerden bizzat kendilerinin muayene ederek, seçerek alabilmelerine, yazar kasa çıkışlarına kadar götürüp bedellerini ödemek suretiyle, tüketim alanına kadar taşıma işlemini de kendilerinin yapmalarına izin veren bir faaliyet biçimidir. Bu yöntemin işlemesi ile perakendeciler tarafından icra edilen fonksiyonların bir kısmı nihaî tüketiciye yüklenmiş olmaktadır. Bu değişiklik vasıtasıyla perakendeciler, bazı avantajlar elde ederken uygulamanın bazı dezavantajları ile de karşı karşıya kalmaktadırlar. Varinli'ye (2005: 41) göre self-servis yönteminin perakendeciye getirdiği bazı avantajlar ve dezavantajlar şöyle sıralanabilir:

#### Self-servis yönteminin avantajları

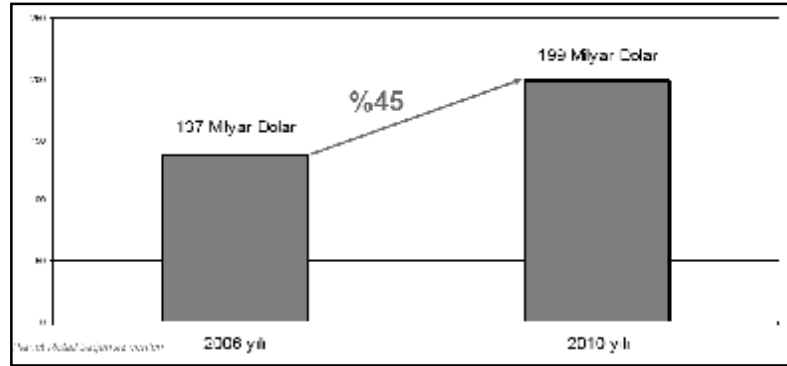
- Müşteriye malı kolayca inceleyebilme ve tezgâhtardan bağımsız olarak karar verebilme imkânı sağlar.
- Daha az satış elemanı gerektirdiği için, satış masrafları ve personel ile ilgili problemler azalır.
- Çeşitli yollarla elde edilen tasarruflar sonucu, diğer mağazalardan daha düşük fiyatla mal satabilme imkânını elde ederler.
- Müşteriler daha rahat bir ortamda alışveriş yapabildikleri için miktar ve çeşit olarak müşteri başına düşen satış daha fazladır.

### Self-servis yönteminin dezavantajları

- Self-serviste, tezgâhta satış yönteminde olduğundan daha fazla satış alanına ihtiyaç duyulur.
- Bazı müşteriler, belirli bir malı arayıp bulmaktan ve bedelini ödemek üzere kasaya götürmekten hoşlanmazlar ve bu yüzden tezgâhtarlar tarafından hizmet sunulmasını beklerler.
- Bazı mağazalarda, self-servis yöntemine dayalı olarak satış hacminin artması, özellikle alışverişin yoğun olduğu zamanlarda, kasa önlerinde kargaşaya neden olur. Bu durum ise, kasiyerlerin daha fazla hata yapması ve bu nedenle de müşterinin güvensizlik duyması sonucunu doğurur.
- Daha ziyade orta ve düşük gelirli tüketici gruplarına ambalajlanmış ve iyi bilinen malları sunmada geçerli olan self-servis yöntemi, yüksek gelirli gruba hitap eden ve ambalajlanmamış mallar satan perakendeciler için pek uygun değildir.
- Self-servis yönteminde hırsızlık olayı daha yaygındır.
- Dayanıklı tüketim malları gibi yüksek fiyatlı malların satışında, satış elemanına ihtiyaç duyulduğu bilinmekle beraber yüksek fiyatlı olmadıkları halde, (ayakkabı, elbise gibi) bazı malların da satış elemanı gerektirdiğini ortaya çıkarmıştır. Satış elemanının bulunmadığı bir satış ortamında, müşteri bağımlılığının kazanılabilesinde güçlükler de söz konusudur.

## **4. TÜRKİYE'DE PERAKENDE SEKTÖRÜ**

Perakende sektörü Türkiye'nin en büyük sektörlerden birisidir. Türkiye'de perakende sektörü enerji, eğitim ve sağlıktan sonra dördüncü büyük sektördür. Şekil 1'de de görüldüğü gibi 2006 yılı toplam cirosu 137 milyar dolar, 2007 yılı toplam cirosu 150 milyar dolardır. 2010 yılına kadar cironun 199 milyar dolar olacağı tahmin edilmektedir (ampd.org, 10.04.2008).



Şekil 1. Türkiye Perakende Sektörünün Toplam Ciro Grafiği

Aydın'a (2005:21) göre perakende sektörünün en önemli alanlarından birisi olan hızlı tüketim ve ihtiyaç malzemelerinin pazar büyüklüğü yaklaşık 60 milyar dolardır. Bu pazar payının yaklaşık 38 milyar dolarlık kısmı gıda ve temizlik pazarına, geri kalan 22 milyar dolarlık kısmı ise gıda ve temizlik dışı pazara aittir. Toplam perakende tüketiminin %65'ini gıda, %35'ini gıda dışı tüketim oluşturmaktadır. Bu verilerde organize perakendeciliğin payı ise 20 milyar dolar düzeyindedir. Türkiye'de organize perakendenin genel perakende içerisindeki oranı %30-33 arasında değişirken, geri kalan %67-70'lik kısmını ise geleneksel perakendecilik oluşturmaktadır. Gelişmiş ülkelerde ise geleneksel perakendenin toplam perakende cirosuna oranı %15-20'dir.

Sektörde toplam 2,5 milyon çalışan bulunmaktadır. Organize perakende sektöründe 3 yıl içinde 250 bin yeni istihdam beklenmektedir. Organize perakendede yaklaşık 16 milyon metrekare satış alanı bulunmaktadır. Mevcut 186 alışveriş merkezine, 3 yıl içinde 170 yeni alışveriş merkezi eklenmesi beklenmektedir (ampd.org, 10.04.2008).

Ülkemizde modern gıda perakendeciliğinin gelişim sürecinde süpermarketler 1955 yılında İsveç Migros Koop. Federasyonu tarafından İstanbul'da kurulan Migros Türk A.Ş. ile başlamıştır. Migros-Türk'ten sonra, Türkiye'de modern gıda perakendeciliğinde doğrudan bir kamu girişimi olarak 1956 yılında Gıda TAŞ (Gıda ve İhtiyaç Maddeleri)'nin kuruluşu gerçekleşmiştir. Ancak, Türkiye'de süpermarketlerin asıl gelişmesi, ekonomik gelişmeye ve özellikle gıda sanayindeki ilerlemeye bağlı olarak, 1970 sonrası başlamış ve 1980 yılından sonra da artış kaydetmiştir (Demir, 2007: 20).

Güllü'ye (2005: 22) göre, ithalatın serbestleştirildiği 1980 sonrası, yabancı tüketim mallarının piyasaya girebilmesi, tüketicinin bu mallara olan bağımlılığının artması, devlet eliyle işletilen mağazaların tüketicinin ihtiyaçlarını karşılamada yetersiz kalması, bu dönemde, özel girişimciliğin gelişmesi ve özel perakende zincirlerinin sayıca artması sonucunu doğurmuştur.

1990'lı yıllarda yabancı perakendeci kuruluşların Türkiye perakende sektörüne girmesi sonucu sektörde temel değişimler meydana gelmiştir. 1988'de Metro, 1991'de Carrefour ve 1992'de Continent Türkiye perakende pazarında yerlerini almışlardır. Avrupa'da kişi başına düşen tüketim oranlarının sabitleştirilmiş olması, bu oranların istikrarlı bir yapıya kavuşması, gittikçe artan mevzuatlar bu yabancı kuruluşların yeni pazar arayışına girmelerini zorunlu hale getirmiştir. Türkiye'de bu tür kuruluşlara, özellikle hipermarketler ve açılış yerlerine ilişkin mevcut bir mevzuatın olmaması, dolayısıyla bir sınırlamanın bulunmaması, buna karşılık çok dinamik, çok genç bir nüfusun var olması, bu tüketim potansiyelini değerlendirmek isteyen yabancı perakendeci kuruluşlara perakende sektörünü çekici hale getirdiği söylenebilir. Bu yabancı kuruluşlar, gelişmiş tüketici hizmetleri, büyük sermayeleri ve know-howları ile perakende pazarında, rekabet artırıcı ve eğitici görev icra etmişlerdir. Büyük miktarlarda alımlardan elde ettikleri avantajlar ve sahip oldukları sermayeye bağlı olarak yüksek kalitede hizmet sunabilmeleri neticesinde rekabet gücü artan büyük mağazalar karşısında, küçük mağazaların bazıları ya kapanmış ya da aralarında örgütlenerek, alım birlikleri kurarak (Anmar, İsmar, Karmar), rekabet güçlerini artıracak çabalar içine girmişlerdir (Güllü, 2005: 22).

Yabancı sermayenin perakende piyasasına girmesi, yerli perakendecileri süpermarketleşme ve hipermarketleşme biçiminde yeni stratejiler belirlemeye zorlamıştır. Zincir mağaza biçiminde varolan marketlerin Türkiye pazarının bütününe hedef alarak yayılma girişiminde oldukları, uygun durumda olan bazı perakende satış mağazalarını hipermarkete dönüştürdükleri, mevcut girişimlere ilâve olarak da yeni girişimcilerin Anadolu'nun büyükşehirlerinde market ve hipermarket açtıkları görülmektedir.

Türkiye'de perakende ticaretin %8,1'ini küçük süpermarketler, %2,5'ini büyük süpermarketler oluşturmaktadır. Avrupa'da 1 milyon kişiye ortalama 150



süpermarket, 15 hipermarket (karma mağaza) düşerken, ülkemizde 1 milyon kişiye düşen hipermarket sayısı 2, süpermarket sayısı ise 16'dır (ampd.org, 10.04.2008).

Nielsen (2007) tarafından açıklanan Hızlı Tüketim Ürünleri Pazar Değerlendirme raporu verilerine göre, 2007 yılı sonunda toplam hiper ve süper market sayısı geçen yıla göre yüzde 14 artarak 8.252'e ulaşmıştır. Bakkal sayısı yaklaşık %1,6 gerilerken, orta marketler ise yüzde 3 büyümeye ile 15.273 âdete ulaşmıştır (Tablo 1).

*Tablo 1. Türkiye'de Perakende Sektöründe Faaliyet Gösteren İşletmelerin Sayılarının Yıllara Göre Dağılımı*

	2006	2007	2008*
Hipermarket	164	178	183
Büyük Süpermarket	504	568	623
Süpermarket	1.567	1.712	1.902
Küçük Süpermarket	4.239	4.763	5.544
Orta Market	14.775	14.876	15.273
Bakkal	116.857	115.220	113.295
Toplam Özel Market	52.207	52.946	53.214
<b>TOPLAM</b>	<b>190.313</b>	<b>190.263</b>	<b>190.034</b>

\*: Tahmini

Akan'a (2007: 5) göre Türkiye'de 2003 yılından itibaren süpermarketlerin ve indirim mağazalarının (discount) sayısında önemli bir artış olmuştur. Süpermarketlerin ve indirim mağazalarının yükselişi Türkiye ekonomisi açısından yeni bir oluşumdur. Söz konusu bu marketlere ulaşımın kolay olmasından, örneğin konut alanlarına yakın olması ve alışveriş için daha az zaman harcanması sebebiyle, Türkiye'de giderek artan sayıda insan bu tür yerlerden alışverişini tercih etmektedir. Türkiye'deki çeşitli tipteki gıda perakendecilerin pazar paylarının yıllara göre değişimi Tablo 2'de verilmiştir (PricewaterhouseCoopers, 2007: 7).

*Tablo 2. Türkiye'de Gıda Perakendecilerinin Pazar Payları (%)*

	2000	2001	2002	2003	2004	2005
Hipermarketler	2,8	2,9	2,9	3,0	3,1	3,2
Süpermarketler	19,3	19,8	20,8	21,8	23,0	24,2
İndirim Mağazaları	3,2	3,7	4,0	4,4	5,0	5,6
Marketler	1,1	1,1	1,1	1,0	1,0	1,0
Bakkallar	47,4	46,4	45,3	44,3	43,2	42,0
Şarküteriler	16,8	16,8	16,7	16,5	16,0	15,6
Diğer	9,4	9,3	9,2	9,0	8,7	8,4

Nielsen (2007) araştırma şirketinin 2007 yılında yayınladığı raporda ise yukarıda bahsedilen araştırmadan daha farklı sonuçlar elde edildiği görülmektedir. Bu araştırmaya göre toplam pazar paylarına bakıldığında, bakkallar pazarın yüzde 29,5 ile en yüksek payı oluştururken, orta marketler yüzde 6,6'sını, hipermarketler yüzde 10'unu, büyük süpermarketler yüzde 11,9'unu, süpermarketler yüzde 15,3'ünü ve küçük süpermarketler de yüzde 16,7'sini oluşturmaktadır. Geri kalan yüzde 10'luk kısmı ise özel market kategorisinde yer alan kuruyemiş, eczane, büfe, parfümeri vb. satış alanlarının oluşturduğu görülmektedir. Bu sonuç bakkalların pazar payının düşme ve süpermarket-hipermarket gibi perakendeci işletmelerin pazar payının yükselme trendinde olduğu yorumunu doğrulamaktadır. Temel gıda ürünleri pazarına bakıldığında ise, bakkallar yüzde 34,3 ile diğer satış alanları arasında en yüksek paya sahip olarak görülmektedir.

İnternet ve yeni teknolojilerin perakende sektöründe kullanılmasıyla birlikte pek çok yenilik ve değişiklik meydana gelmiştir. Maliyetlerin düşmesi, rekabetin artması, fiyat mekanizmasının daha sağlıklı işlemesi, piyasaya giriş engellerinin eskiye oranla azalması, fiyatların şeffaflaşarak alıcı ve satıcı arasındaki enformasyon akışının düzenlenmesi ve paylaşımın sınırsız olması, piyasanın daha verimli işlemesi, ekonomik kaynak kullanımı ve dağıtımının daha doğru bir şekilde yapılması bu değişiklikler kapsamında incelenebilmektedir (Demir, 2007: 60).

Perakende sektörü özellikle 1980'lerden başlayarak sürekli bir devinim ve gelişim içindedir. Yerel oluşumların gösterdikleri gelişme ve büyüme çabaları, yabancı perakendecilerin de pazara girmesiyle birlikte artan rekabetin ardından daha da yoğun hale gelmiştir. Barutçugil'e (2006: 1) göre perakende sektöründe rekabet temelde, hemen her sektörde olduğu gibi dört ana boyutta yaşanmaktadır. Bunlar; ürün ve hizmetlerde, fiyatlarda, yer ve dağıtım anlamında kuruluş yerlerinde ve tanıtım-tutundurma faaliyetlerinde sürdürülen rekabettir.

- Ürün ve hizmetlerde rekabet; ürün türü, ürün çeşitliliği ve müşteriye sunulan hizmetlerin düzeyi ile bağlantılıdır. Daha fazla ve farklı ürünler bulundurmak, müşteriye ürün seçenekleri ve bunları sunma biçimleri ve ürünlerin içerdiği ek hizmetlerdeki farklar rekabetçi üstünlük sağlayabilmektedir.

- Fiyat rekabeti, aslında sonu, anlamı ve uzun dönemde de faydası olmayan bir çabadır. Kısa dönemde belki bir fark yaratabilir; ancak dikkatli yönetilmediğinde sonuçta herkesin kaybettiği fiyat savaşlarına dönüşebilmektedir.
- Kuruluş yerleri açısından rekabet yavaş yavaş eski önemini kaybetmektedir. İletişim ve ulaşım olanaklarındaki gelişmeler, müşterinin evine yakın olmanın sağlayacağı avantajları önemli ölçüde azaltmıştır.
- Tanıtım ve tutundurma amaçlı etkinliklerin, yani reklam, promosyon, halkla ilişkiler gibi çabaların, özellikle ilgi ve dikkat çekme anlamında, başlangıçta bir rolü bulunmaktadır.

Daha fazla sayıda personel, daha fazla sayıda kasa ve daha fazla sayıda ürün seçenekleri sunan mağazalar, tüketicilerin daha önce yaşadığı tecrübelerden farklı tecrübeleri yaşatmıştır. Rekabet düzeyinin yüksekliği ve tüketicilerin eskisine oranla daha fazla seçenek istemesi, farklı markaları bir arada sunma anlayışını beraberinde getirmiştir. Böylece tüketiciler mağazadan çıkmadan marka kıyaslaması yapma imkânı bulmaktadırlar (Özdemir, 2006: 10).

Gürman'a (2006: 259) göre, yabancı yatırımcıların perakende pazarına girmesiyle birlikte artan rekabet, yerli perakendecilerin verimlilik kavramına bakış açılarını değiştirmiş, operasyonel kârlılığa, insan kaynağının geliştirilmesine, teknolojinin kullanımına, müşteriyle ilişkilerin yönetimine, bilimsel mağazacılığa karşı yaklaşımını şekillendirmiştir. Perakendeciliğin odak noktasında olan müşteri çok çeşitli perspektiflerden incelenmeye başlanmış, onu satın almaya yönelten psikolojik, sosyal, maddi, vb. nedenlere açıklık getirebilmek için tüketiciler bilimsel bir yaklaşım ile mercek altına alınmıştır. Klasik indirim ve promosyon yöntemleriyle perakendeciliğin geride kaldığı bir dönemde artık teknolojik gelişmelerden de faydalanarak akıllı (bilimsel) mağazacılığı uygulamak ön plana çıkmıştır.

Dal (2006: 2) klasik indirim ve promosyon kampanyalarıyla gerçekleştirilen mağazacılık döneminin kapandığını, perakende sektöründe artık başarının anahtarının "Bilimsel Mağazacılık" olduğunu öne sürmektedir. Bilimsel mağazacılığın amacı, satış, stok, müşteri hareketleri gibi gerçekleşen operasyonlar

sonucunda oluşan veriye dayalı, yönetim kültürü ve bakış açısı sağlamaktır. Bilimsel mağazacılık yaklaşımı ile bir perakende tedarik zincirindeki hem talep-tüketici odaklı süreçlerin hem de üretici-tedarik odaklı süreçlerin bütünleştirilmesi ve optimizasyonu, hazırlanan matematiksel modelleme ve optimizasyon yöntemleriyle sağlanmaktadır. Sistem sayesinde daha etkili tahminler yapılabilmekte (tahmin tutarlılığında %14-40 arasında artış), stok seviyeleri indirgenebilmektedir. Bu sayede satışlarda ortalama %50, kâr marjında da %5-20 arasında, çalışan verimliliğinde %20-40 arasında artış sağlanabilmektedir.

Rakiplerin uluslararası işletmeler oluşu; onlarla rekabet edebilmek için büyümeyi gerekli kılmıştır. Ölçek büyüklüğünün sağladığı maliyet avantajları ve sinerjik etki sayesinde müşteri tatmini artırılırken, mağaza da kârlılığını artırma imkânına kavuşmaktadır.

Günümüz piyasalarında rekabet edebilmek için artan veri çeşitliliğine rağmen, analizlerin daha hızlı yapılması ve sonuçta anlamlı ve eyleme yönelik bilgiler ortaya çıkarılması gerekmektedir. Rekabetin yoğunlaşması; uyum sürecinin gerektirdiği hızı yakalayabilmenin, müşteri odaklı olmanın ve verimliliğin önemini her zamankinden daha çok arttırmıştır.

Dünyanın ve Türkiye'nin en dinamik sektörü olan perakende sektörü, geleneksel sınırları aşarak, yenilikçi ve müşterileriyle açık iletişim kurabileceği yollar aramaktadır. Buna paralel olarak, artan bir hızla büyümektedir. Sektörün büyümesiyle birlikte rekabet de artmaktadır. Farklılaşmanın ve öne çıkmanın yolu, son on yılda sektöre damgasını vuran ve perakende kuruluşların önemle üzerinde durduğu teknoloji kavramıdır. Perakende sektörü, teknolojiye en çok yatırım yapan sektörlerin başında gelmektedir. Tedarik zinciri yönetimi, mal çeşitliliğinin optimizasyonu ve müşteri yönetim sistemleri şirketlerin karlılığını arttırmaktadır (Demir, 2007: 2).

Perakendecilik alanındaki kuruluşlar, eldeki verilerden faydalanarak ve daha çok müşteri ilişkileri yönetimi (Customer Relationship Management) uygulamaları yaparak (kampanya yönetimi, promosyonlar vb.) hangi müşteri profilinin hangi ürünü daha çok talep ettiği bilgisi, ne tip ürünlerin daha fazla gelir getirdiği, hangi ürün satın alınırken buna bağlı diğer bir ürünün de satın alınma olasılığı gibi

çalışmalara odaklanarak var olan potansiyel müşteriyle satışı elde tutabilme ve müşteri sayısını artırma amacına yönelik çalışmalar yaparak sektördeki varlıklarını sürdürmektedirler (Yıldırım, 2005: 85). Bu kuruluşlar ellerindeki potansiyel verileri veri madenciliği teknolojisi ile işleyerek bu bilgileri elde edebilirler.

Günümüzde işletmeler açısından yeni bir müşteri kazanmak, mevcut müşteriyi sadık hale getirmekten çok daha fazla maliyetlidir. Firmalar müşteri sadakatini en yüksek seviyeye çıkarmayı ana amaç haline getirmişlerdir. Bunun için de müşterileri tanıma, kişiselleştirilmiş hizmetler sunma, şikâyetleri dinleme, giderebilme, ödüller verme ve memnuniyeti arttırabilmeyi sağlama amaçlarına yönelik uygulamalar yapmaya yönelmişlerdir.

## **5. PERAKENDE SEKTÖRÜ VE BİLGİ TEKNOLOJİLERİ**

Bilgi teknolojileri son yıllarda çok hızlı bir gelişme göstermektedir. Bilgi teknolojilerinin perakende sektöründe etkin bir şekilde kullanılması perakendeciye sürdürülebilir bir rekabet avantajı kazandırabilir. Perakende sektöründe bilgi teknolojileri, özellikle etkili bir veri yönetimi sağlamak amacıyla kullanılmaktadır. Güllü'nün (2005: 189) Dawson'dan (1994) aktardığına göre, perakendecilikte kullanıldığı şekliyle, bilgi teknolojisi satış noktalarında satış faaliyetlerine ilişkin verilerin düzenli bir şekilde toplanmasını, daha sonra bu verilerin mağaza faaliyetleri, lojistik ve satınalma gibi çeşitli konularda karar vermede kullanılmasını içerir. Bilgi ve iletişim teknolojilerindeki gelişmeler işletmelere, perakendeci maliyetlerinde azalma, daha geniş tüketici tercihinine hitap etme ve tüketicilere yerel ve kişisel sunumlarda bulunabilme gibi önemli avantajlar sağlamıştır.

Elektronik satış noktası uygulamalarının (POS-Point of Sale), elektronik bilgi değişiminin, veri yönetiminin, satış alanı optimizasyonu modellerinin gelişmesi, perakendecilik sektörünün yapısını geliştirmiştir. Yeni teknolojilere dayalı bilgi yönetimine sahip olmak ve bunları etkin bir şekilde kullanmak, perakendecilerin rakipleriyle rekabet edebilmelerini ve pazarda varlıklarını sürdürebilmelerini sağlar. Bilgi ve iletişim teknolojilerindeki ilerlemeler, mağazasız perakendeciliğin özellikle elektronik ticaretin kullanımının ve pazar payının artmasını sağlamıştır.

Güllü'ye (2005:190) göre bilginin perakendeciler ile mal tedarikçileri arasındaki dağıtım kanallarındaki akışı kadar, perakendeciler ile tüketiciler arasındaki akıştaki önemi de gittikçe artmaktadır. Bilgi akışı sayesinde, mağaza operasyonları ve tüketici işlemlerinde verim artırılabilir, karar verme mekanizması derinleşecek, tanzim-teşhir ve stok yönetimi geliştirilecek ve çalışan üretkenliğinde ilerleme kaydedilebilecektir. Dünyanın en büyük perakendecisi olan Wal-Mart'ın başarısında teknolojinin önemi büyüktür. Mağazalarının birinde bir kot pantolon satıldığında, hem üretici hem de Wal-Mart, ürünün satıldığını ve yenisinin mağazaya gönderilmek üzere işlemin başladığını bilmektedir. Aynı zamanda Wal-Mart her satışı elektronik ortamda depolayarak çok büyük bir veri tabanı meydana getirmekte ve ürün kategorisi bazında tanımlanan satışlar üzerinde daha sonra gerçekleştirilen analizlerle tüketici ihtiyaçlarını daha iyi tahmin etmeye çalışmaktadır.

Perakendeciler bilgi teknolojilerini kullanarak, her bir tüketicinin neyi, ne zaman, ne kadar aldığını, hangi markanın, hangi boyunu ve rengini tercih ettiğini izleyebilmektedir. Tüketicileri daha iyi tanıma, her bir tüketicinin ihtiyaç ve beklentilerini çok daha etkin bir biçimde tatmin etme imkânı sağlamaktadır. Öte yandan, tüketicilerin de ürünler ve perakendeciler hakkında özellikle Internet ortamında bilgi edinebilmesi de süreci etkilemektedir.

Ürünlerin ambalajlarında yer alan barkodlar da gelişen bilgi teknolojisi sayesinde işlev kazanmaktadır. Barkod kullanımının hem tüketicilere, hem perakendecilere pek çok faydası bulunmaktadır; ürünün satışının hatasız ve hızlı yapılması, envanterin doğru olarak tutulabilmesi, satış yöntemini ve kasa işlemlerini kolaylaştırması, müşteri ve satıcılarla iletişimi sağlaması gibi perakendecilere yararlar sağlamaktadır. Bunların sonucunda da işletmede verimliliği artırmaktadır.

## **6. KATEGORİ YÖNETİMİ VE TANZİM-TEŞHİR KAVRAMLARI**

Kategori, ortak bir özelliği paylaşan nesnelerin bir araya toplanmasıdır. Kategorileştirme ise nesnelerin aynı tür sınıf ya da kategorilere dağıtılması işlemidir (Aydın ve Candan, 2007: 8). Perakende sektöründe ürün çeşitlerinin kategorilere

ayrılması genellikle ürünlerin yapısına, nerede-nasıl tüketildiklerine ya da hangilerinin satın alındığına göre yapılır. Geleneksel kategorileştirmeden farklı olarak, günümüzde ürünlerin kategorilere ayrılması işleminin tüketicilerden gelen bilgiler doğrultusunda yapılmasına daha fazla önem verilmektedir.

Perakendecilik açısından stok ve depolarda bulunan ürünlerin ayrıntılı bir şekilde sınıflandırılması gerekir. Bu bakımdan kategori yönetimi, bir ürün kategorisi içindeki en temel düzeyde tanımlanan ürünlerin yönetimi sürecidir. Bu süreç fiyat, raf yeri, tanzim, teşhir stratejisi, tutundurma çabaları ile pazarlama karmasının bütün diğer unsurları ile birlikte kategori temeline göre işletme hedefleri, değişen çevre ve tüketici davranışları üzerine sürekli yönetimini kapsar (Taşkın, 2002: 104).

ACNielsen (1992:9) kategori yönetimini; müşteri tatminini sağlayabilmek amacıyla, ürün gruplarının bağımsız birer işkolu gibi ele alınarak, mağaza veya mağaza grupları bazında gerekli düzenlemelerin yapılmasını işlemi olarak tanımlamaktadır. Bu süreçte bireysel ürün ya da marka yerine ürün kategorileri esas alınmaktadır. Böylelikle etkinlik ve verimlilik sağlanarak satışın ve kârlılığın artırılması hedeflenmektedir.

Aydın ve Candan'ın (2007: 21) Brookman'dan aktardığına göre kategori yönetimi tedarikçi, dağıtıcı, toptancı ve perakendecinin; bir kategorinin toplam kârlılığında etkisi olan her ürüne ayrı bir biçimde odaklanmasını esas alan esnek/değişken bir organizasyonel yaklaşımdır.

Aydın'a (2007) göre kategori yönetimi; ürün gruplarını bağımsız birer işkolu gibi ele alarak, mağaza veya mağaza kümeleri bazında, müşterileri tatmin etmek için gerekli düzenlemeleri yapan dairesel bir süreçtir. Bir başka deyişle üretici perakendeci işbirliği ile birlikte müşteriye fiyat, çeşitlilik, tutundurma ve stok konularında maksimum değeri yaratmak üzere kategorilerin düzenlenmesi sürecidir.

Kategori yönetimi, müşteri tepkilerinin etkili olarak ölçülmesi, en uygun mal seçimi, etkili tutundurma ve etkili mağaza politikası konularıyla yakından ilgilidir. Etkili mağaza yönetimi için mağazaya en uygun ürünlerin sağlanması, ürün yönetimi, fiyat yönetimi, tutundurma faaliyetlerinin yönetimi ile raf yerlerinin yönetimi çalışmaları kategori yönetiminin konuları içinde yer alır (Taşkın, 2002: 104).

Diğer taraftan kategori yönetimi, müşteri odaklı pazarlama stratejilerini desteklemektedir. Kategori yönetimi ile sadık müşterilerin alışveriş alışkanlıkları ve ihtiyaçları daha iyi incelenerek, daha iyi tedarik zincirleri kurularak etkinlik yükseltilmektedir (Taşkın, 2002: 105). Kategori yönetimi, sadece belirli bir markanın değil, o kategorideki tüm markaların satışlarının yükseltilmesini hedefler.

Berman ve Evans'a (2001: 475) göre kategori yönetiminin başarı koşullarından biri kategorilerin oluşturulmasında raflarda tüketicilerin ürün taleplerine uygun düzenlemeler yapılmasıdır. Aynı zamanda kategori yönetimi, kategori içerisinden bir çeşidin satın alınması yerine çok fazla ürün satın alınmasına yönelik motive edici olmalıdır. Herkese neyi satacağınız kadar neyi satmayacağınızla ilgili kararları da içerir. Kategori yönetimi için veri tabanı; ticari alanın ihtiyaçlarına yönelik ön müşteri arayüz analizlerinden sağlanmalıdır.

Kategori yönetiminde satın alma noktası ya da satın alıcıya odaklanması önemli bir faktördür. Öneri geliştirmek ve satın almaya motive etmek için onların mağazadaki tutum ve davranışlarını bilmek önemlidir. Bu yüzden bir kategori ile ilgili ticari geri dönüşleri maksimize etmek için müşteriye anlamak önemlidir (Aydın ve Candan, 2007: 38).

Kategori yönetimi, tanzim teşhir, satın alma, satış tahmini ile birlikte; satış geliştirme, fiyatlandırma gibi pazarlama karmasının tutundurma unsurları ile daha fazla bilişim teknolojilerinden yararlanılmasına ihtiyaç göstermektedir. Bu nedenle mevcut büyük satış noktaları bilişim teknolojilerinden daha çok yararlanmalıdır (Taşkın, 2002: 112).

Kategori yönetiminin verimli ve doğru yapılabilmesi için veri akışı tam olarak sağlanmalıdır. Veriler alışveriş detayları, demografik, psiko-grafik ve diğer birçok veri tiplerinden oluşur. Kategori yönetimi temelde veri odaklı ve yoğun analizlerin gerçekleştirildiği bir işletme sürecidir.

Kategori yönetiminin işletme düzeyinde çalışması, yönetim bilişim sistemlerinin kurulması ve bilgi işleme birlikte uygulanmasına bağlıdır. Böyle bir sistem perakendeciye giren müşterilerin alışverişlerini inceleyerek ürün karmasının belirlenmesini ve siparişlerin düzenli olarak verilmesini sağlar. Böylece stoksuz kalma durumları önlenerek aynı müşteriye daha fazla ürün satılarak başarı yükseltilir.



Burada ilk olarak önemli olan kasadan çıkan bir ürünün okutulan barkod etiketi ile envanter ve tedarik arasındaki veri bağlantısını kurmaktır (Taşkın, 2002: 114).

Aydın ve Candan'a (2007: 28) göre kategori yönetiminin çoğu kez yıldırıcı gözükten bir yönü de veri analizidir. Tüketici ve satış verilerinin kullanımı; kategori dinamiklerinin anlaşılmasına yeni bir bakış açısı sağlar. Bununla beraber mevcut verilerin dikkatsiz yönetimi etkin karar verme yerine yanlış kararlara yol açabilir. Kategori yönetiminde verilerin kullanımı müşteri davranışı, satınalma kararları, sınıflandırma/gruplandırma, alan ayırma/taahsis, kategori trendleri, finansal fırsatlar, kategori etkinliği, stok seviyeleri ve bulunulurluk konularda değerli bilgiler sağlar.

Türkiye'deki çeşitli marketlerdeki kategori yönetimi uygulamalarını inceleyen Beksırgakızı (2001), şu sonuçlara ulaşmıştır: Migros'un 35, Gima'nın 35, Yimpaş'ın 30, Beğendik'in 30 ve Tansaş'ın 25 ürün kategorisi bulunmaktadır. Seçilen süpermarketlerde kategoriler öncelikle iki temel kategoriye (Gıda ve Gıda Dışı) ayrılarak, daha sonra da alt kategoriler oluşturulmaktadır. Kategoride yer alacak markalar; tüketici tercihleri de dikkate alınarak mağazadan mağazaya değişebilmektedir. Kategori yönetimi uygulanmasının öncesi ve sonrasındaki dönemlerde; rafların paylaşımı ve çeşitlendirilmesinde önemli farklılıklar bulunmaktadır. Kategori yönetiminin uygulanması ile hem müşteriler hem de kendileri için çeşitlilik artmakta, fiyatlar düşmekte, stoklar azalmakta, yenilikler ortaya çıkmakta, alışveriş yapılması kolaylaşmakta, müşteri sayıları artmakta ve buna paralel olarak pazar payı da yükselmektedir.

Günümüzde perakendeci ve üreticiler, her bir mağazanın çevresinde oturan müşteri nüfusunu tanımlayan demografik verilerinden, hane halkı için uzun dönemli satınalma kalıpları/modelleri ile ilgili ayrıntılı bir görünüm sağlayan panel verilerinden, perakendecinin satış noktası verileri ile üreticilerin herhangi bir seviyede analiz için daha ayrıntılı bir bakış sağlayacak kategori ve tüketici verilerinden faydalanmaktadırlar. Bu araçlara sağlanacak verilere dayalı olarak, alınacak daha akılcı kararlar, müşterilerin alışveriş deneyimlerine uygun satış ve karlılığı sağlamaktadır (Aydın ve Candan, 2007: 85).

Kategori yönetimi sürecinde karar verme esnasında, tedarikçiden sağlanan bilgilerin yanında, perakendecilerin kullanabileceği birçok analitik araç vardır. Bunlar (ACNielsen, Karolefski ve Heller, 2006:79);

- Pazar sepet analizi: Alışveriş sırasında insanlar hangi ürünleri beraber alırlar? Kategoriler arasında daha güçlü bağlantılar kurularak, nasıl daha çok sayıda ve çeşitte ürün satılması sağlanır?
- Sık alışveriş eden müşteri verileri: En iyi müşterileriniz ne tercih ederler? Onlar için hangi kategoriler çok önemlidir?
- Tüketici panel verileri: Bu veriler; yaygınlığı (penetration), satın alma sıklığını, değişimi ve daha birçok veriyi içermektedir. Mendil alan müşterilerin yüzde kaçını sizden alışveriş yapıyor? Bir perakendecinin müşterilerinin yüzde kaçını bu ürünleri başka yerden alır ve diğer perakendecilerde ne kadar harcama yapar? Mendil alan müşterileri nasıl kendi işletmenize çekebilirsiniz ve onların kategorilerden daha çok alışveriş yapması nasıl sağlanır?
- Satış Noktası (POS) verileri: Bu veriler kategori büyüklüğünü ve satış trendlerini içerir. Her mağazadaki satış verileri analiz edilerek şu sorulara cevap verilebilir: Ulusal satışlar, yerel pazar satışları karşılaştırıldığında nasıl bir sonuç elde edilir? Kategori büyüyor mu, küçülüyor mu? Kategori büyümesini fiyat artışları mı motive ediyor ya da belirli bir kategoride gerçekte tüketiciler daha çok mu, daha az mı satın alıyorlar?
- Mevsimsel Analiz: Bir kategorinin tatil günlerine ve diğer etkinliklere katkısını belirlemek için yapılır.
- Doğru Pay ve Talep Açığı: Şu anda neyin gerçekten satıldığı ve neyin satılabileceği arasındaki farktır. Ticari alanda müşteriler hangi ürünü ne kadar sever, müşteriler ürününüzden ne kadar tüketirler, daha ne kadar satabilirsiniz?
- Finansal Analiz: Satış, stok devir hızları, kâr miktarı (brüt kâr veya satış fiyatı ile satılan malların maliyeti arasındaki fark), müşteri

değişimi (herhangi başka bir perakendeciden, herhangi bir kategorideki ürünleri satın alan insanlar, sizin mağazanızdan ne kadarını alırlar, siz bunu ne kadar arttırabilirsiniz) ve faaliyet esaslı maliyetleme (satılan ürüne ait tüm masraflar; yükleme, depolama, elleçleme, pazarlama vb.).

Genellikle süpermarketler ve hipermarketler ürün gruplarına göre bölümlere ayrılmakla birlikte, kategori yönetimini uygulamaya başlayan marketlerde tüketicinin satın alma alışkanlıklarına göre yerleşim tercih edilmeye başlanmıştır (Varinli , 2005: 197).

Türkçe literatürde genellikle tanzim-teşhir olarak kullanılan merchandising terimi, pazarlama ile ilgili bütün faaliyetleri içine alacak kadar geniş bir terimdir. Tanzim-teşhir hangi malların satılacağı, ne miktar stok yapılacağı, fiyatların nasıl saptanacağı, ürünlerin nerelerde ve nasıl tüketiciye sunulacağı gibi işlemleri kapsayan genel bir terimdir. Hatipoğlu (1986: 67) merchandising kelimesini ürünlerin en iyi yerlerde, en iyi zamanda, en iyi miktarda ve en iyi fiyatla satışa sunulması olarak tanımlamaktadır.

Perakendecilikte üründen daha çok “merchandise” terimi kullanılmaktadır. Bu bağlamda tanzim-teşhir kavramı çeşit, fiyat/kalite, marka ve stok devir hızı gibi kavramları da içermesi nedeniyle daha geniş bir anlam içermektedir. Diğer taraftan merchandise perakendecilikte iyi tanımlanmamış terim (ill-defined term) olarak da bilinir. Bazılarına göre ürünlerin teşhiri, sunumu, bazılarına göre ise çok daha geniş bir anlam ifade etmektedir. (Aydın ve Candan, 2007: 62). Ürkmez (2008: 13) de tanzim-teşhiri satış noktalarında son tüketiciye yönelik yapılan ve satış artışı sağlamayı amaçlayan her türlü iletişim ve dikkat çekici aktiviteler bütünü olarak tanımlamaktadır.

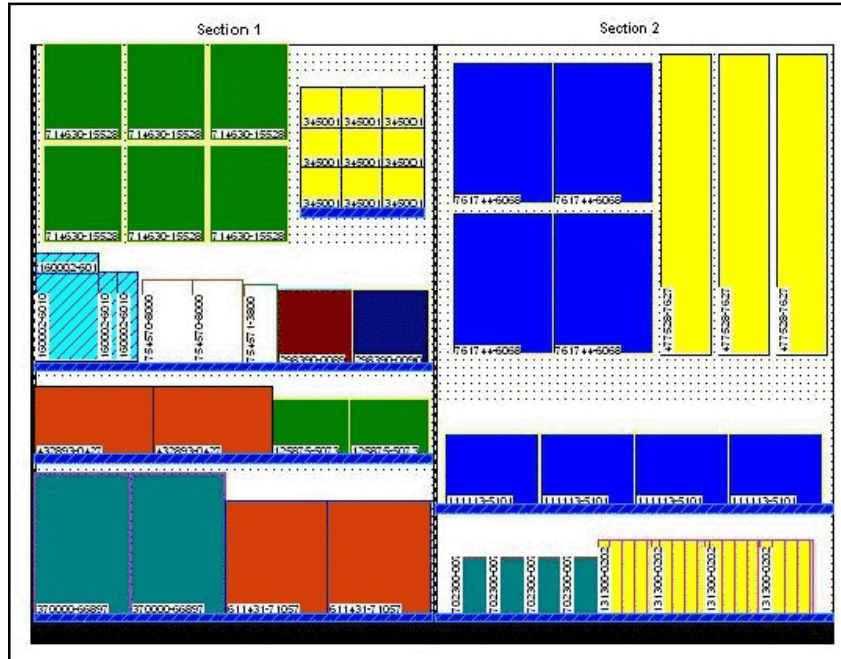
Tanzim-teşhir planlaması, şirket içerisinde kategorinin stratejik rolü esas alınarak; kategori için satış ve kâr hedeflerini başarmak amacıyla stratejik pazarlama ve finansal plan yaratılmasını içermektedir. Teknolojik yenilikler bu tür planların hazırlanmasını ve geliştirilmesini kolaylaştırmaktadır (Aydın ve Candan, 2007: 92).

Tanzim-teşhir planlaması tüketicinin satın alma kararını etkilemek, rakibe göre avantaj sağlamak, tutarlı marka imajı yaratabilmek, kârı-satışı arttırabilmek ve

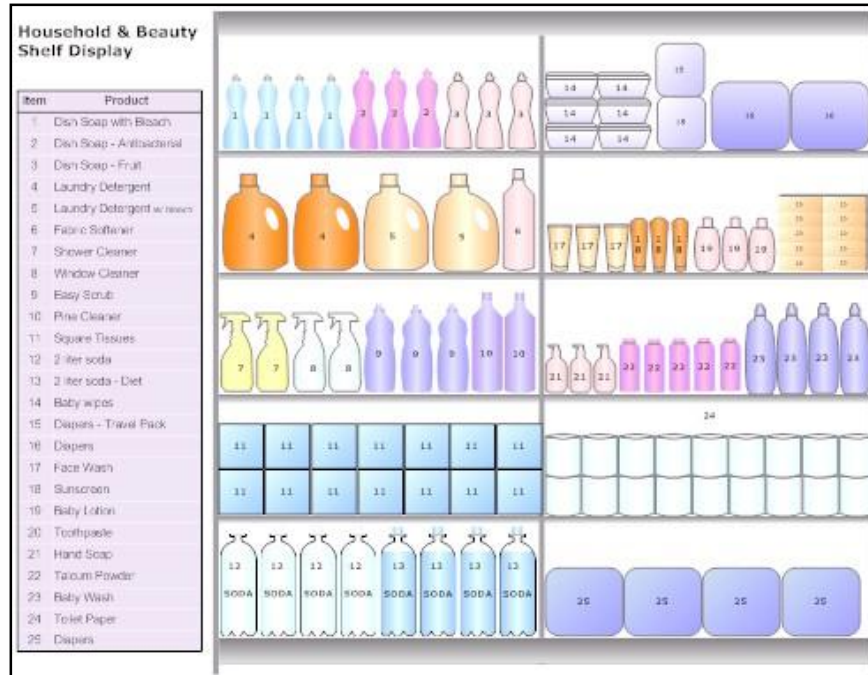
satış noktasını yönetmek gibi amaçlar doğrultusunda gerçekleştirilir (Ürkmez, 2008: 15).

Perakendecilerin kategoriye nasıl sunacakları konusundaki temel belirleyicilerden biri satış alanıdır. Çok başarılı perakendeciler yer-satış rasyosunu dikkate alarak kategori, alt kategori ve segmentleri çok iyi bir biçimde yönetmektedirler. Perakende satış noktaları için; doğru büyüklük, iyi bir konum ve doğru stok çok önemlidir. Her perakende mağaza sabit bir alana sahiptir ve bunu arttıramaz. Bir kategoriye, alt kategoriye ya da segmenti genişletmek için aynı şekilde diğerlerinden kısmak gerekecektir. Bu basit gerçeği görmemek ve satış alanı konusunda hatalı hesaplamalar yapmak kategorilerde karışıklıklara yol açmaktadır. Örneğin; belirlenmiş sabit bir yer için daha büyük miktarda ürün gelirse, bunun nereye yerleştirileceği ve ilave yer bulunulup bulunamayacağı önemli bir sorundur. Mağazalar fazla gelen bu ürünleri ilk seferde raf dışında teşhir edebilirler, ancak yeni ürün geldiğinde fazla gelen bu ürünler hiçbir zaman rafta yer bulamazlar. ACNielsen verilerine göre ABD'de stoksuzluk nedeniyle %8 ile % 12 arasında kayıplar yaşanmaktadır. Bunun nedeni ise çoğunlukla satış alanı yetersizliğidir (Aydın ve Candan, 2007: 170).

Raf sunumu, çok kritik bir değer taşımaktadır. Çünkü ya insanların alışveriş yapmasını sağlayacaktır ya da onları uzaklaştıracaktır. Kontrolü kaybetmemek için, perakendeciler satış alanının kullanımında her ayrıntıyı dikkate almak zorundadırlar. Planogram, bir ürün ya da ürün grubunun perakendeci mağazasının raflarında nasıl sergileneceğini gösteren çizimdir. Arslan ve Bayçu (2006: 254) planogramı, mağaza içindeki alanın satış, ürünler, personel, müşteriler ve ürün grupları için görsel gösterimi şeklinde tanımlamaktadırlar. Şekil 2-4'de çeşitli planogram örnekleri görülmektedir.



Şekil 2. Planogram Örneği (Mağaza Genel Yerleşimi)



Şekil 3. Planogram Örneği (Ürün Yerleşimi-şema)



Şekil 4. Planogram Örneği (Ürün Yerleşimi-fotoğraf)

Planogram tasarlamasında etkili olan faktörler brüt kâr marjı, ürün kârlılığı ve her bir kalem ürünün satılabilirliğidir. Planogramlar ile aynı ürünün farklı ebatlarının hangi raflara yerleştirilmesi gerektiği de planlanır. Deneyimsiz stok görevlileri dikkatsizlikle ve istemeyerek de olsa planogram planlarını mahvedebilirler. Bazen de yerleştirme kurallarına uyulmadığında çok kâr yapan deterjanlar göz hizası yerine rafın en alt kısmına konularak kayıpların ortaya çıkmasına neden olunabilir (Aydın ve Candan, 2007: 170).

Kategori alanının yanlış hesaplanması, mağaza deposundan daha sık bir biçimde rafa ikmal yapılmasına yol açmaktadır. Bu durum bir taraftan işgücü maliyetlerinin artmasına yol açarken, diğer taraftan ürünün taşıma (handling) nedeniyle daha fazla zarar görmesine neden olmaktadır. Bu şekilde zarar gören ürünler sunumda da sorun olmaktadır.

Bugün, her bir mağaza için 100'den daha fazla kategori planogramı bulunması ve bunların karmaşıklığına rağmen, mağaza büyüklüğüne (400 m<sup>2</sup>'den

10.000 m<sup>2</sup>'ye kadar) bağılı olarak; planogramlar yaklaşık %60 dolayında birbirlerine uygunluk göstermektedir. Sonuç olarak; perakendeciler çoğunlukla "rehber planogram"lardan yararlanmaktadırlar. Mağaza yöneticileri bu modellerden yararlanarak kendileri için en uygun olanını hazırlamaktadırlar. (Aydın ve Candan, 2007: 171).

Günümüzde dinamik veritabanının bir parçası olarak planogramlar; satış/alan ilişkisi raporlarına bakılarak yeniden düzenlenebilir ve satış raporları ile yeniden ayarlanabilir. Yapılan analizler ve değerlendirmeler sonucunda ürünün dördüncü rafın ortasında olması yerine, ikinci sıranın köşesinde olması halinde daha fazla satış yapılacağı gibi bir bilginin bulunması ve bunun planogramlarda uygulanması bu şekilde bir yeniden düzenlemeye örnek olarak gösterilebilir.

## İKİNCİ BÖLÜM

### VERİ, VERİTABANI VE VERİ MADENCİLİĞİ

#### 1. VERİ, BİLGİ, VERİTABANI KAVRAMLARI

##### 1.1. VERİ VE BİLGİ

Sanayi toplumunda bilgi toplumuna dönüşümü yaşadığımız günümüzde, bilişim teknolojilerinin giderek yaygınlaşması ve bilginin beşinci üretim faktörü olarak ekonomik sisteme dâhil edilmesi bilgiye verilen önemi arttırmıştır. Günümüzde bilginin elde edilmesi, işlenmesi, depolanması ve dağıtılması alanlarında çok büyük gelişmeler kaydedilmiştir. Çok boyutlu bir kavram olan bilgi, sıklıkla veri kavramı ile karıştırılmaktadır.

Alpaydın'a (2000:1) göre veri sadece sayılar veya harfler değil; sayılar, harfler ve onların anlamıdır. Bilgi ise verinin belli bir amaca yönelik olarak işlenmesiyle elde edilir. Bir diğer ifade ile bilgi, bir soruya yanıt vermek için veriden çıkardığımız sonuç olarak tanımlanabilir. Özdemir'e (2004:5) göre veri, bir bilgi sistemine girilen veya diğer cihazlardan alınan, yapısal olmayan, işlenmemiş girdilerin tümüdür. Han ve Kamber (2001: 4) bilgiyi, potansiyel olarak ilginç ve faydalı olan veriler arasındaki ilişki olarak tanımlamaktadır. Veriyi bilgiye çevirme işlemine veri çözümlemesi denmektedir.

Tekin, Güleş ve Öğüt (2003: 2) veriyi, işlendiğinde bilgi haline getirilebilecek ve sonuçlar çıkarılabilecek; kişilerle, olaylar, süreçlerle ve fikirlerle ilgili ham olgular olarak tanımlamaktadır. Bu bağlamda veri şu anda gerçekleşen ve geçmişte gerçekleşmiş sürekli olayla ilgili ham gerçekleri kapsamaktadır. Bilgi ise verinin analiz edilerek, karar vermek için anlamlı ve kullanılabilir bir hale getirilmesidir.

Vahaplar ve İnceoğlu'na (2006: 1-2) göre dünyadaki veri miktarı, gelişen teknolojinin yardımıyla her an çok büyük miktarlarda artmasına rağmen, verilerden



elde edilen anlamlı bilgi miktarı aynı şekilde artmamaktadır. Verilerin dijital ortamda saklanmaya başlanması ile birlikte, yeryüzündeki veri miktarının her 20 ayda bir iki katına çıktığı günümüzde, veri tabanlarının sayısı da benzer bir oranda artmaktadır. San Francisco’da faaliyet gösteren Golden Gate Software şirketinin bir araştırmasına göre, günümüzde veri ambarlarının neredeyse yarısı içerdiği veri miktarı açısından yıllık yüzde 10-50 arası büyüme göstermektedir (Kumar ve Wahidabanu, 2008: 43). Hatta perakendecilik gibi sektörlerde bu artış rakamları daha büyük olabilmektedir. Yüksek kapasiteli işlem yapabilme maliyetinin ucuzlamasının bir sonucu olarak, veri saklama hem daha kolay olmuş, hem de verinin kendisi de ucuzlamıştır (Vahaplar & İnceoğlu, 2006: 2).

Tekin ve diğer.’ne (2003:3) göre, bilgi ancak karar vericinin amaçlarını karşıladığı sürece bir anlam taşıyacağından, kişisel ve örgütsel karar vermede kullanılacak bilginin belirli niteliklere sahip olması gerekir. Bu nitelikler doğruluk, uygunluk, zamanlılık, eksizliklik, denetlenebilirlik, kısalık, güncellik ve ekonomiklik olarak sıralanabilir.

## 1.2. VERİTABANI VE VERİ AMBARI

Bir karar verici için vereceği kararın doğruluğu, onun yeteneklerine, deneyimine ve bilgi birikimine olduğu kadar sahip olduğu veri kümesinin yeterliliğine de bağlıdır. Diğer bir deyişle kararın başarısında, verilerin doğru depolanması, doğru sınıflanması, doğru ayıklanması, doğru işlenmesi ve doğru yorumlanması çok önemli bir rol oynar. Ancak karar süreçlerinin karmaşıklaşması, sayısal olarak daha fazla veriye gereksinimi ortaya çıkarmış, bu durum ise veri depolarının büyüklüğünü manuel olarak kontrol edilemeyecek boyutlara ulaştırmıştır.

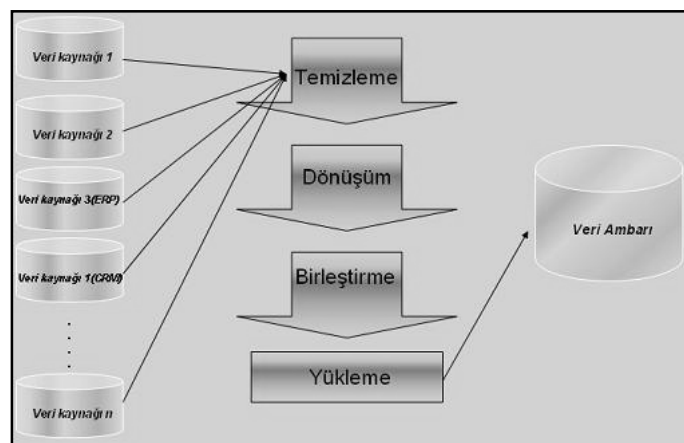
Artık ekonomik sistemde veri ya da bilgi, mal ya da hizmet üretiminin faktörlerinden birisi olarak algılanmaktadır. Bu ise karar vericileri yanlış karar riskinden uzaklaşabilmek için, mümkün olduğunca fazla veriyi depolamaya zorlamaktadır. Ayrıca internetin küreselleşmeyi körüklemesi, rekabetin kırıcı seviyelere ulaşması, kar marjlarının düşmesi ve müşteri memnuniyetinin zorlaşması, bu endişeyi daha da körüklemektedir. Bu durum ise doğru veriyi toplama ya da

doğru veriye ulaşma zorunluluğunu doğurmaktadır. Diğer bir deyişle artık, veriye erişmek en az verinin kendisi kadar önemli hale gelmiştir.

Günümüzde bilgisayarlar ve bilgisayar sistemleri daha büyük miktardaki veriyi saklamakta ve bu veriyi daha kısa sürede işlemektedir. Bunun yanında bilgisayar ağlarındaki ilerleme ile bu veriye başka bilgisayarlardan da hızla ulaşabilmek mümkün hale gelmiştir. Veri doğrudan sayısal olarak toplanmakta ve saklanmaktadır. Bunun sonucu olarak da detaylı ve doğru bilgiye ulaşılabilir hale gelmiştir.

Veritabanı düzenli bilgiler topluluğudur. Kelimenin anlamı bilgisayar ortamında saklanan düzenli verilerle sınırlı olmamakla birlikte, daha çok bu anlamda kullanılmaktadır. Veritabanı bilgisayar terminolojisinde, sistematik erişim imkânı olan, yönetilebilir, güncellenebilir, taşınabilir, birbirleri arasında tanımlı ilişkiler bulunabilen bilgiler kümesidir. Bir başka tanımı da, bir bilgisayarda sistematik şekilde saklanmış, veri analizi programlarınca istenebilecek veri yığındır.

Canitez'e göre (2007: 9) veri ambarları basit olarak veri madenciliği işleminin yapılacağı verilerin oluşturduğu özel veri tabanlarıdır. Veri ambarlarının oluşturulması işlemi verinin çeşitli kaynaklardan toplanarak, veriler içerisindeki uyumsuzluklar ve hatalardan arındırılmasından ibarettir. Günlük veri tabanlarından istenen özet bilgi seçilir ve gerekli ön işlemeden sonra veri ambarında saklanır (Şekil 5). Ardından amaç doğrultusunda gerekli veri ambardan alınarak veri madenciliği çalışması için standart bir forma çevrilir. Özetle veri ambarları, günlük kullanılan veri tabanlarının birleştirilmiş ve işlemeye daha uygun bir özetini saklamayı amaçlar.



Şekil 5. Veri Ambarlarının Oluşturulması

Veri ambarları bir kurumda gerçekleşen tüm operasyonel işlemlerin, en alt düzeydeki verilere kadar inebilen, etkili analiz yapılabilmesi için özel olarak modellenen, tarihsel derinliği olan, fiziksel olarak operasyonel sistemlerden farklı ortamdaki yapılardır (Özmen, 2002: 2).

Bayram'a (2001: 2) göre veri ambarı, organizasyonun ihtiyaçları ile uyumlu büyük miktarlardaki verinin kolay erişilebilir bir yapıda tutulmasını sağlayan bilgisayar tabanlı depolama sistemleridir. Akbulut (2006: 3) veri ambarlarını organizasyonel veriye kolay bir şekilde ulaşılmasını sağlayan yapılar olarak tanımlamaktadır. Veri ambarları veriyi kullanılabilir eğilim, ilişki ve profillerde sınıflandırmazlar; sadece potansiyel bilgiye sahip veritabanlarıdır. Veri ambarından veriyi çekebilmek için hangi verinin gerekli olduğunu ve bu verinin nerede olduğunu tespit etmek önemlidir. Çoğunlukla gerekli veri, farklı sistemler üzerinde olup, farklı formatlardadır. Bu nedenle, ilk aşamada veri temizleme ve düzenleme işlemi gerçekleştirilmelidir. Akbulut'un (2007: 3) W.H. Inmon'dan aktardığına göre veri ambarı verinin temizlendiği, birleştirildiği ve yeniden düzenlendiği merkezi ve bütünleşmiş bir depodur. Inmon veri ambarını özne-tabanlı, bütünleşmiş, zaman dilimli ve yöneticinin karar verme işlemine yardımcı olacak biçimde toplanmış olan değişmeyen veriler topluluğu olarak da tanımlamaktadır (Dolgun, 2006:11).

Tantuğ'a (2002: 8) göre veri ambarı, bir işletmenin ya da kurumun çeşitli birimleri tarafından canlı sistemler aracılığı ile toplanan verilerin, ileride değerlendirmeye alınabilecek olanlarının geri planda yer alan bir sistemde birleştirilmesinden oluşan büyük ölçekli bir veri deposudur. Canlı sistemlerde oluşan veriler periyodik olarak veri ambarına aktarılır. Bu periyodun seçimi tamamen veri ambarını kullanan işletmenin ihtiyaçları doğrultusunda belirlenir ve 1 gün, 1 hafta, 1 ay gibi çok değişken olabilir. Dolayısı ile veri ambarı çevrimdışı olarak çalışmaktadır. Yani veri ambarı içerisindeki kayıtlar güncel olmayabilir, çoğunlukla da güncel değildir.

Canitez'e (2007: 12) göre bir veri ambarı, veri madenciliği ve çok boyutlu analiz gibi günümüzün veri analizi teknikleri ve klasik sorgulama-raporlama için

sağlam bir temel oluşturmaktadır. Bu tekniklerin veri ambarlamayla kullanılması daha formel karar vermede kullanılan bilgilere çok daha kolay ulaşılmasını sağlamaktadır. Merkezi ve büyük bir veri ambarı kurulsa da, bir yâda daha fazla sayıda küçük boyutlu veri ambarları kullanılsa da, ya da bu ikisinin bir karışımı da oluşturulsa; eninde sonunda verinin yapısının ne olacağı noktası önemlidir. Bu ise veri ambarlarının geleneksel veritabanları ve karar destek uygulamalarından farklılıklarını ortaya koyar. Veri ambarlamadaki temel fark verinin yapılandırılması ve etrafına uygulamaların kurulmasıdır. Geleneksel veritabanı yönetim sistemlerinde uygulama yapılandırılır ve veri bu yapıya getirilir.

Veri ambarlama bir ürün seti değil, bir çözümdür. Tipik hareket bazlı operasyonel çevreden farklı bir bilgi ortamıdır. Temel amacımız karar verme becerisinin geliştirilmesidir. Bir veri ambarı daha iyi karar vermek için cevaplar sunmayı kendi başına saptayamaz. Veri ambarlama genel karar verme süreçlerinin içinde yer alan bir unsurdur. Bir veri ambarı kurmadan önce, "Veri ambarındaki veri nereden geldi? Nasıl korundu? Veri ambarda nasıl yapılandırıldı? Veri ambarının içinde gerçekten ne var?" gibi bazı temel sorulara cevap vermek gerekir (Canitez, 2007: 13).

Veri ambarında veri oluşturulduktan sonra bu verinin elle veya gözle analizi yapılabilir. Bunun için OLAP (Online Analytical Processing) programları kullanılır. Bu programlar veriye, her boyutu veride bir alana karşılık gelen çok boyutlu bir küp olarak bakmayı ve incelemeyi sağlar (Seyrek, 2006: 73). Böylece boyut bazında gruplama, boyutlar arasındaki korelasyonları inceleme ve sonuçları grafik veya rapor olarak sunma olanağı sağlanmış olur.

### **1.2.1. Veri Ambarı Yapısı**

Karakaş'a (2005: 2) göre bir veri ambarı yapısı; organizasyon içindeki bütün son kullanıcılara verileri ve işlem sonuçlarını sunan, en gelişmiş iletişimi sağlayan bir dizi birbiriyle bütünleşik alt bileşenlerden oluşur. Bunlar;

- Operasyonel Veri Tabanı / Harici Veri Tabanı Katmanı,
- Enformasyon Ulaşım katmanı,

- Veri Ulaşım Katmanı,
- Veri Dizini (Metadata) Katmanı,
- İşlem (process) Yönetim Katmanı,
- Uygulama Haberleşmesi Katmanı,
- Veri Ambarı Katmanı,
- Veri Sunum Katmanıdır.

#### i) Operasyonel Veri Tabanı / Harici Veri Tabanı Katmanı

Operasyonel sistemler kritik operasyonel ihtiyaçlar için verileri işlerler. Bunu gerçekleştirebilmek amacıyla uzun süreçlerdeki tecrübelerden yararlanılarak en verimli iş modelleri tanımlanmış ve operasyonel veri tabanları oluşturulmuştur. Bu yüzden, operasyonel sistemler üzerindeki limitli odaklanmalar yönetim ve enformasyonel amaçlı verilere ulaşmada zorlukların çıkmasına yol açmıştır. Operasyonel verilere ulaşmadaki bu zorluklar, çoğu 10-15 yıllık operasyonel sistemler tarafından da imkânsız gibi görülmeye başlamıştır. Ancak bahsedilen veri tabanları bu zorlukları aşarak kolay bir şekilde verilere ulaşılmasını sağlamaktadır.

Veri ambarlarının amacı, operasyonel veri tabanlarındaki verileri kullanılabilir kılmak ve diğer harici enformasyonel verilerle harmanlamaktır. Günümüzde artan bir şekilde çok büyük şirketlerin dış dünyadaki veri tabanlarından verileri kendi ihtiyaçları doğrultusunda yakalayarak bünyelerine dâhil ettiği görülmektedir. Bu veriler demografik, ekonometrik, rekabetçi ve satın alma trendleri ile ilgili olabilmektedir. “Bilgi otobanı” diye tarif edilen kavram her geçen gün daha fazla bilgiye ulaşmaya olanak sağlamaktadır.

#### ii) Enformasyon Ulaşım Katmanı

Enformasyon ulaşım katmanının bir veri ambarındaki yeri son kullanıcının direkt olarak veriye ulaşması ile ilgili kısmıdır. Bu amaca yönelik olarak enformasyon ulaşım katmanı kullanıcıların bu verilere ulaşması için araçlar sunar; bunlar yaygın olarak kullanılan Excel, Lotus 1-2-3, Focus, Access gibi yazılımlardır. Bu katman donanım ve yazılımları, grafikleri, sunuları ve sonuçları kâğıt üzerine

aktaran yazıcıları da kapsar. Özellikle geçtiğimiz son 10 yıl içinde bilgisayar ağlarının oldukça kolay kurulur ve kullanılabilir hale gelmesi ortak bir platformda çalışmayı da kolay hale getirmiştir.

### iii) Veri Ulaşım Katmanı

Bir veri ambarı yapısının veri ulaşım katmanı, enformasyon ulaşım katmanı ile operasyonel katmanın birbiriyle haberleşmesini sağlayan bir katmandır.

Günümüzdeki network dünyasında kabul edilen ortak sorgulama dili olan SQL bu sebeple çıkmıştır. İlk olarak IBM tarafından geliştirilen bu veri tabanı sorgulama dili geçtiğimiz 20 yılda, bir veri sorgulama dilinden öte bir veri değişimi dili olarak endüstriyel bir standart halini almıştır.

Veri ambarlarının en önemli stratejilerinden birisi "veriye evrensel ulaşım sağlama"dır. Veriye evrensel ulaşım sağlamak demek, son kullanıcının, teorik olarak da olsa, veriye herhangi bir ekstra araç, yazılım kullanmadan ihtiyaç duyduğu her an her yerden ulaşabilmesidir.

Veri ulaşım katmanı, enformasyon ulaşım araçlarıyla operasyonel veri tabanları arasında bir arabirim oluşturmaktan da sorumludur. Çoğu durumda bu, son kullanıcının ihtiyaç duyduğu tek şeydir. Bunlara ek olarak, günümüzdeki birçok organizasyon veri ambarını destekleyici yönde her geçen gün yeni felsefi derinliği olan fikirler üzerinde çalışmaktadırlar.

### iv) Veri Dizini (Metadata) Katmanı

Verilerin yapısı, evrensel veri ulaşımını sağlayabilmek için, incelenmesi gereken en önemli konulardan birisidir. Veri dizini bir veri ambarındaki verilerin genel özelliklerini ve yerleşim özelliklerini içeren bir tablodur. İdeal bir veri ambarına sahip olabilmek için kullanılabilir haldeki çeşitli veri dizini tiplerine ihtiyaç vardır. Mükemmel olan, kullanıcıların veri ambarına verilerin nerede tutulduğunu bile bilmeksizin ulaşabilmesidir.

### v) İşlem Yönetim Katmanı

İşlem yönetim katmanı zamanı belli olan, yapılması ve devamlılığı olan bir dizi işlemleri ihtiva eder. İşlem yönetim katmanı, yüksek seviyede işi kontrol altında

tutabilmeyi sađlayan, işleri planlayan ve büyük miktarlardaki verinin her zaman doğru ve yeni olmasını sađlayan bir ara birim olarak da düşünülebilir.

#### vi) Uygulama Haberleşme Katmanı

Uygulama Haberleşme katmanı, enformasyonun bilgisayar ağı (network) haberleşmesi teknikleri ve teknolojileriyle bütün organizasyon içinde nakledilmesinden sorumlu olan katmandır. Uygulama haberleşmesi “orta kademe katman” olarak da adlandırılabilir. Fakat sadece network haberleşmesinden öte, uygulamaların birbirinden yalıtılmasından, verilerin operasyonel ya da enformasyonel olarak ayrımı ve verilerin doğru formata çevriminden de sorumludur. Uygulama haberleşmesi katmanı, bilgisayarlar arasındaki uygulama işlemlerinin sonuçlarını kaydetmek ve doğru zamanda, doğru kimseye, en doğru verileri sađlamak amacıyla da kullanılır.

#### vii) Veri Ambarı Katmanı

Veri ambarları temel olarak enformasyonel veriyi kullanır. Çoğu zaman insanlar veri ambarlarının fiziksel olarak dataları depolamadığını, bunun sanal olarak verilerin farklı bir formatta görünümünden ibaret olduğunu düşünürler.

Operasyonel faaliyetler ve müşteri bilgileri gibi veriler, fiziksel bir veri ambarında, çok kolay ulaşılabilir ve kolaylıkla diğer formatlara çevrilebilecek şekillerde saklanır. Günümüzdeki veri ambarları web tabanlı kişiselleştirilmiş platformları kullanmaya başlamıştır; ancak temelde aynı yapıyı kullanırlar.

#### viii) Veri Sunum Katmanı

Bir veri ambarı yapısının sonuncu bileşeni de veri sunum katmanıdır. Veri sunum katmanı, “kopya yönetim katmanı” yâda “replikasyon katmanı” olarak da adlandırılır. Fakat aslında bu katman operasyonel ve/veya harici veri tabanlarına ulaşarak yapılan tespit, sonuçlandırma, uyumlu hale getirme ve yükleme işlemleri için gereken bütün aşamaları da kapsar.

Veri sunumu karmaşık programlama tekniklerine de gereksinim duyar. Ancak günümüzde hızla gelişen veri ambarcılığı araçları ile bu aşama her geçen gün daha da kolaylaşmaktadır.

Veri sunum katmanı mevcut operasyonel data içindeki kalıpları ve veri modellerini tespit etmek amacıyla da, veri kalitesini analiz eden muhtelif filtreleri ve/veya programları da içerebilir.

### 1.2.2. Veri Ambarı Tipleri

Öğüdücü (2002: 4-5) veri ambarlarını amaca yönelik, birleştirilmiş, zaman değişkenli, değişken olmayan veri ambarları olarak dörde ayırmaktadır.

- Amaca yönelik veri ambarları; müşteri, ürün, satış gibi belli konular için düzenlenebilir. Verinin incelenmesi ve modellenmesi için oluşturulurlar ve konuyla ilgili karar vermek için gerekli olmayan veriyi kullanmayarak konuya basit ve özet bir bakış sağlarlar.
- Birleştirilmiş veri ambarları; veri kaynaklarının birleştirilmesiyle oluşturulurlar. Canlı veri tabanlı dosyalardır. (En fazla 1 yıllık verilerden oluşurlar.) Veri temizleme ve birleştirme teknikleri kullanılır. Değişik veri kaynakları arasındaki tutarlılık sağlanır.
- Zaman Değişkenli veri ambarları; zaman değişkeni canlı veri tabanlarına göre daha uzundur. Geçmiş yâda 10 yıllık dönem hakkında bilgi verirler.
- Değişken olmayan veri ambarları; canlı veritabanlarından alınmış verinin fiziksel olarak başka bir ortamda saklanmasıdır. Canlı veritabanlarındaki değişim veri ambarlarını etkilemez.

Veri ambarı bir kurumun stratejik kararları verme sırasında ihtiyaç duyduğu bilgiyi depolayan ve karar destek veri modelinin fiziksel bir sunumu gibi çalışan, anlamsal olarak tutarlı bir veri deposudur. Veri ambarı aynı zamanda, yapısal ve/veya planlanmamış sorgular, analitik raporlar ve karar vermeyi desteklemek için farklı türde kaynaklardan veriyi bütünleştirerek oluşturulan bir mimari olarak da görülür (Han ve Kamber, 2001: 105-106).



### 1.2.3. Veritabanı ve Veri Ambarı Arasındaki Farklar

Dolgun'un Han ve Kamber'dan (2001:106) aktardığına göre veri ambarlarını, ilişkisel veritabanı sistemlerinden ayıran başlıca dört özellik özne-tabanlı, bütünleşmiş, zaman dilimli ve değişmeyen kavramlardır.

- **Özne-tabanlı:** Bir veri ambarı; tüketici, tedarikçi firma, ürün ve satış gibi önemli konular etrafında kurulur. Veri ambarı bir organizasyonun her güne ait işleri ve hareket işleme faaliyetleri üzerinde yoğunlaşmak yerine, karar verecek kimseler için veriye ait modelleme ve çözümleme üzerinde yoğunlaşır. Bundan dolayı, veri ambarları tipik olarak ayrıntılı konu sorunları çevresinde, karar destek sürecinde faydalı olmayan veriyi hariç tutarak basit ve öz bir bakış sağlar.
- **Bütünleşik:** Bir veri ambarı genellikle; ilişkisel veri tabanları, düz dosyalar ve OLAP kayıtları gibi çeşitli farklı türde (heterojen) kaynaklar bütünleştirilerek oluşturulur. Veri temizleme ve veri bütünleme yöntemleri; isimlendirmede, şifreleme yapılarında, nitelik ölçütlerinde ve benzeri konularda tutarlılığı garantilemek için uygulanır.
- **Zaman Dilimli:** Veriler, tarihi bir bakış açısından bilgi sağlamak için depolanır (örneğin 5-10 yıllık geçmiş içerisinden). Veri ambarındaki her anahtar yapı, zamanın bir elemanı olarak hem kesinlik hem de açıklık içerir.
- **Değişmeyen:** Bir veri ambarı, her zaman operasyonel çevre içerisinde bulunan uygulama verisinden dönüştürülen veri kaydından fiziksel olarak ayrıdır. Bu ayrıma bağlı olarak bir veri ambarı hareket işlemeyi, geri almayı ve rastlantısal kontrol mekanizmalarını gerektirmez. Veriye erişim için sadece iki işlem gerektirir; verinin ilk yüklemesi ve verinin erişimi.

Veritabanı içerisindeki bilgiler genelde anlık bilgilerdir. Yani belirli bir süre sonunda güncelliğini kaybedecek olan bilgilerdir. Ancak veri ambarı içerisinde veriler genelde yığılarak birikirler ve verilerin geçerliliği çok daha uzun sürmektedir. Veri ambarı içerisinde ne kadar çok kayıt olursa, yapılan incelemelerin sonucu da o kadar doğru olacaktır. Oysa veritabanı içerisindeki kayıt sayısının çok fazla olması durumunda, veritabanını kullanan canlı sistemlerin performansları düşecek ve bu da

verilere erişimi çok yavaşlatacaktır. Bu durum, canlı sistemlerde en istenmeyen durumdur (Tantuğ, 2002: 10).

## 2. VERİ MADENCİLİĞİ

Bilgisayar teknolojisi ve veritabanlarının kullanımı son yıllarda yaygınlaşmıştır. Bununla birlikte birçok kuruluş rutin olarak çok büyük miktarda veri toplamakta ve depolamaktadır. Verinin boyutları ve karmaşıklığı, veriyi anlamayı ve çözümlenmeyi zorlaştırmıştır.

Veri madenciliği net olarak tanımlanamayan bir alandır. Veri madenciliği literatürde pek çok farklı şekilde tanımlanmıştır. Tanımı çoğunlukla tanımlayıcının arka planına, ilgi alanına ve görüşüne dayanmaktadır.

Akbulut'a (2006: 3) göre veri madenciliği; verideki trendleri, ilişkileri ve profilleri belirlemek için veriyi sınıflandıran bir analitik araç ve bilgisayar yazılım paketidir. Bir başka tanıma göre veri madenciliği, ham verinin tek başına sunamadığı bilgiyi çıkaran veri analizi sürecidir. Canitez (2007: 20) de veri madenciliğini; önceden bilinmeyen, geçerli ve uygulanabilir bilginin veri yığınlarından dinamik bir süreç ile elde edilmesi olarak tanımlamıştır.

Ceran'a (2006: 25) göre veri madenciliği, eldeki verilerden üstü kapalı, çok net olmayan, önceden bilinmeyen ancak potansiyel olarak kullanışlı bilginin çıkarılmasıdır. Başka bir deyişle, veri madenciliği, verilerin içerisindeki desenlerin, ilişkilerin, değişimlerin, düzensizliklerin, kuralların ve istatistiksel olarak önemli olan yapıların yarı otomatik olarak keşfedilmesidir. Veri madenciliğinde keşfedilecek kurallar veritabanının özelliklerine ve kuralların kullanımına göre farklı tekniklerle bulunur.

Hand (1998: 113), veri madenciliğini istatistik, veritabanı, örüntü tanıma, makine öğrenmesi alanlarının etkileşimde olduğu yeni bir disiplin ve büyük veritabanlarında önceden tahmin edilemeyen, bilinmeyen ilişkilerin analizi olarak tanımlamıştır.

Özmen'in (2002:3) Swift'den (2000) aktardığına göre veri madenciliği; veri ambarlarında tutulan çok çeşitli verilere dayanarak daha önce keşfedilmemiş bilgileri ortaya çıkarmak, bunları karar verme ve eylem planını gerçekleştirmek için kullanma sürecidir. Bu noktada kendi başına bir çözüm değil çözüme ulaşmak için verilecek karar sürecini destekleyen, problemi çözmek için gerekli olan bilgileri sağlamaya yarayan bir araçtır.

Veri madenciliği, verilerin içerisindeki desenlerin, ilişkilerin, değişimlerin, düzensizliklerin, kuralların ve istatistiksel olarak önemli olan yapıların yarı otomatik olarak keşfedilmesidir (Vahaplar & İnceoğlu, 2007: 2).

Jacobs (1999: 8), veri madenciliğini, işlenmemiş verinin tek başına gösteremediği bilgiyi ortaya çıkaran veri analizi süreci olarak tanımlamıştır.

Kumar ve Wahidabanu'ya (2008: 44) göre veri madenciliği, veri sahibi tarafından anlaşılabilir ve veri sahibine faydalı olacak şekilde, genellikle büyük boyutlu gözlemsel veri setlerinin, var olduğu bilinmeyen ilişkilerin tespit edilmesi ve verinin metinsel bir biçimde özetlenmesi amacıyla analiz edilmesidir. Veri madenciliği işlemi sonucunda bulunan ilişkiler ve elde edilen metinler genellikle model veya örüntü olarak isimlendirilir.

Akbulut'un (2006: 4) David'den (1999) aktardığına göre veri madenciliği büyük hacimli verilerdeki örüntüleri araştıran matematiksel algoritmaları kullanmaktadır. David'e göre veri madenciliği hipotezleri keşfeder, sonuçları birleştirmek için insan yeteneğini kullanır. Veri madenciliğinin sadece bir bilim olmadığı, aynı zamanda bir sanat olduğu da söylenebilir.

Veri madenciliği ve SQL (Structured Query Language) birbirine yakın işleri yapan iki tekniktir. SQL ve veri madenciliği ile ilgili şu söylenebilir: tam olarak ne arandığı biliniyorsa SQL, ne arandığı belli belirsiz biliniyorsa veri madenciliği kullanılır. İkinci bir nokta ise geleneksel sorgu (query) veya raporlama araçlarının veri yığınları karşısında yetersiz kalmasıdır. Genellikle ne arandığı belli belirsiz olduğu durumlar, ne arandığı bilindiği durumlardan çok daha fazladır. Bu durum da son zamanlarda veri madenciliğine olan ilginin sebebini açıklamaktadır (Adriaans ve Zantinge, 1996: 159).

İşletmelerde istatistiğin kullanım amacı; nasıl ana kütle hakkında anlamlı bilgiler elde etmek ve yorum yapmaksa, veri madenciliğinin amacı da; anlamlı bilgiler elde etmek ve bunu eyleme dönüştürecek kararlar için kullanmaktır. İlgilendiği ana kütle örneğin mevcut veya potansiyel müşteriler olabilir. Müşterilerin profillerini, satın alma eğilimlerini, bir ürünü veya hizmeti kabul etme veya etmeme ihtimallerini tahmin etme veri madenciliğinde hedeflenen amaçlar arasındadır. Bu tahminler ise ana hedef olan, strateji belirlemek ve çeşitli kararlar vermek için kullanılır (Canitez, 2007: 22).

Veri madenciliği asıl olarak veritabanlarıyla ilgilidir. Veritabanlarını kullanıcı için daha kolay hale getirerek, büyük miktarlardaki veriyi yönetmeyi kolaylaştırır. Veri madenciliğinde girdi ve çıktı kolay anlaşılır ve kullanıma uygun olmalıdır. Bu yüzden görselleştirme tekniklerine büyük bir gereksinim vardır, çünkü karmaşık sonuçlar daha iyi anlaşılabilmesi için görselleştirilebilir (Baykal, 2003: 38-39).

Veri madenciliği kendi başına bir çözüm değil çözüme ulaşmak için verilecek karar sürecini destekleyen, problemi çözmek için gerekli bilgileri sağlamaya yarayan bir araçtır. Veri madenciliği; analistine, iş yapma aşamasında oluşan veriler arasındaki şablonları ve ilişkileri bulması konusunda yardım etmektedir. Özetle veri madenciliği:

- Anlamlı örüntüler ve kurallar bulmak için büyük miktardaki verinin analiz ve keşfidir.
- Verinin faydalı özetinin ortaya çıkarılmasıdır.
- Veri madenciliği, verinin analizi ve veri kaynaklarındaki benzerlik ve örüntüleri bulmak için kullanılan yazılım teknikleri ile ilgilidir.
- Veri madenciliği analizleri, veriden anlamlı bilgiye, güvenilir karar ve sonuçlara varmak için verinin dönüşümüdür.

Veri madenciliğinin genellikle kullanıldığı yerler büyük veritabanlarıdır. Küçük ölçekli veritabanlarında veri madenciliğine çok fazla gerek yoktur. Örneğin 200-400 adet üyeye sahip bir web sitesi için, üye portföyündeki insanların bilgilerini tahmin etmek veya onlara ilişkin istatistikler saptamak için bir veritabanı uzmanının yapabileceği sorgular yeterli olabilecektir. Fakat 1-2 milyon üyeye sahip, günlük 1

milyona yakın giriş yapılan bir web sitesindeki üyelere ait bilgilerin incelenmesi için bir veritabanı uzmanı ve geleneksel sorgu yöntemleri yeterli olmayacaktır.

Veri madenciliğinden önce, verilerin toplanması ve veri tabanları oluşturulması ilk olarak 1960'larda başlamıştır. Bu dönemde veri toplama sistemleri popülerleşmiş ve 1970'li yıllarda "ilişkisel veri modelleri", buna bağlı olarak "ilişkisel veritabanı yönetim sistemleri" ortaya çıkmıştır. Daha sonra 1980'lerde ise ilişkisel veri modelleri yaygınlaşarak "ileri veri modelleri" ve uygulama kaynaklı veritabanı yönetim sistemleri eldeki verilerin aktif hale dönüşebileceği düşüncesinin ürünü olarak gelişmiştir. 1990 ve 2000'li yıllarda ise veri madenciliği ve veri ambarlama, çoklu ortam veritabanları ve web veritabanları teknolojilerin gelişimine paralel olarak ortaya çıkıp günümüze kadar gelmiştir. Tablo 3'de teknolojik gelişmeler ışığında çeşitli veri işleme teknikleri ile veri madenciliğinin karşılaştırılmasını göstermektedir (Aldana, 2000: 11).

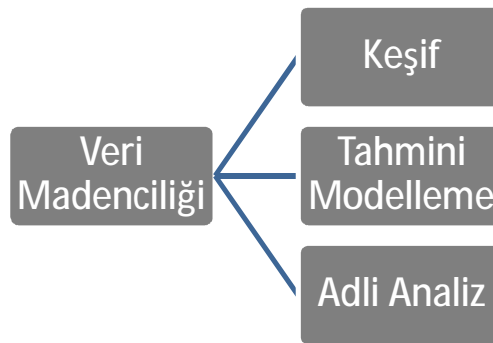
*Tablo 3. Teknolojik Gelişmeler Işığında Veri Analiz Tekniklerinin Gelişimi*

<b>Gelişim Adımları</b>	<b>Cevaplanan Karar Problemi</b>	<b>Kullanılabilen Teknolojiler</b>	<b>Ürün Sağlayıcıları</b>	<b>Karakteristikler</b>
<b>Veri Toplama (1960'lar)</b>	Benim toplam kârım geçen 5 yılda ne kadardı?	Bilgisayarlar, Teypler, Diskler	IBM, CDC	Geriye dönük, statik veri dağıtımı
<b>Veri Erişimi (1980'ler)</b>	İngiltere'de geçen Mart ayında birim satışlar ne kadardı?	İlişkisel veritabanı, SQL, ODBC	Oracle, Sybase, Informix, IBM, Microsoft	Kayıt düzeyinde geriye dönük, dinamik veri dağıtımı
<b>Veri Ambarlama ve Karar Destek Sistemleri (1990'lar)</b>	İngiltere'de geçen Mart ayında birim satışlar ne kadardı?	OLAP, Çok boyutlu veritabanı sistemleri, Veri ambarları	Pilot, Comshare, Arbor, Cognos, Microstrategy	Çoklu düzeylerde geriye dönük dinamik veri dağıtımı
<b>Veri Madenciliği (Bugün)</b>	Gelecek ay Boston'daki birim satışlar muhtemelen ne olabilir, niçin?	İleri düzeyde algoritmalar, çok işlemcili bilgisayarlar, büyük veritabanları	Pilot, Lockheed, IBM, SGL, SPSS, SAS, Microsoft	Geleceğe dönük, proaktif enformasyon dağıtımı

Spesifik veri madenciliği yazılımları; kümeleme, doğrusal regresyon, sinir ağları, bayes ağları, görselleştirme ve ağaç tabanlı modeller gibi pek çok modeli içerir. Veri madenciliği uygulamalarında yıllar boyu istatistiksel yöntemler kullanılmıştır. Bugünün veri madenciliği teknolojisinde eski yöntemlerin yanında, çeşitli yapay zekâ algoritmaları da işin içine katılarak, büyük veri kümelerindeki eğilim ve ilişkileri kısa zamanda saptayabilmek için yüksek hızlı bilgisayarlar kullanılmaktadır. Böylece veri madenciliği, gizli trendleri minimum çaba ve emekle ortaya çıkarmaktadır (Akbulut, 2006: 3).

## 2.1. VERİ MADENCİLİĞİNİN İŞLEVLERİ

Veri madenciliği işlevleri 3 başlık altında toplanabilir: Keşif (discovery), tahmini modelleme (predictive modeling) ve adli analiz (forensic analysis) (Şekil 6) (Vahaplar ve İnceoğlu, 2007: 5).



Şekil 6. Veri Madenciliği Aktiviteleri

### 2.1.1. Keşif

Keşif, ne olabileceği konusunda önceden belirlenmiş bir fikir ya da hipotez olmadan, veri tabanı içerisinde gizli desenleri arama işlemidir. Geniş veri tabanlarında kullanıcının pratik olarak aklına gelmeyecek ve bulmak için gerekli doğru soruları bile düşünemeyeceği birçok gizli desen olabilir. Buradaki asıl amaç, bulunacak desenlerin zenginliği ve bunlardan çıkarılacak bilginin kalitesidir.

Örneğin, bir ülkenin nüfus kayıtlarını incelendiğini varsayalım, Kullanıcı “yaş”la ilgili ilginç veriler bulmak istediğinde, sistem bir analist gibi hareket edecek

ve kurallar çıkarmaya çalışacaktır. Örneğin “Eğer Meslek=Sporcu ise, Yaşı %71 olasılıkla 30’dan küçüktür.” kuralının anlamı, eğer veri tabanından 100 adet Sporcu seçilirse, bunların 71 adedinin yaşı, 30’dan küçüktür demektir. Benzer olarak sistem, “Eğer Meslek=Sporcu ise, Yaşı %97 olasılıkla 60’dan küçüktür” sonucunu da çıkarabilir. Bu da 100 sporcudan en az 97’sinin 60 yaşından küçük olduğunu belirtir.

### **2.1.2. Tahmini modelleme**

Tahmini modellemede, veri tabanından çıkarılan desenler, geleceği tahmin için kullanılır. Bu model, kullanıcının bazı alan bilgilerini bilmeseydi kayıt etmesine izin verir. Sistem, bu boşlukları, önceki kayıtlara bakarak tahmin yoluyla doldurur. Keşif, verideki desenleri bulmaya yönelikken, tahmini modelleme, bu desenleri yeni veri nesnelere bulmak için uygundur.

Önceki örneği baz alırsak, artık mesleği sporcu olan birinin yaşı yaklaşık olarak tahmin edilebilir. Kayıtlara yaşı bilinmeyen fakat mesleği sporcu olan biri eklendiğinde, bu kişinin yaşının %71 oranıyla 30’dan küçük, hatta %97 oranında 60’dan küçük olduğu tahmin edilebilir. Burada keşif, genel bilgiyi bulmamıza yardımcı olur ama tahmini modelleme, daha spesifik bilgileri tahmin etmekte kullanılır.

### **2.1.3. Adli analiz**

Adli analiz, normal olmayan ya da sıra dışı veri elemanlarını bulmak için, çıkarılmış desenleri uygulama işlemidir. Sıra dışı olanı bulmak için ilk önce sıradan kısmı tespit etmek gerekir. Örnekte 60 yaşından sonra hâlâ spor yapan %3’lük bir kesimin olduğunu bilmekte ancak sebebi bilinmemektedir. Bunlar sıra dışı eleman olarak kabul edilmektedirler. Kimisi normalin dışında sağlıklı olabilir ya da yaş ile ilgisi olmayan sporlarla (örneğin golf) uğraşarak olabilirler. Ya da bu veri tabanındaki bilginin yanlış olabileceğini de gösteriyor olabilir. Görüldüğü gibi adli analiz, keşifte aranan genel bilginin tersine, sıra dışı ve özel durumları araştırır.

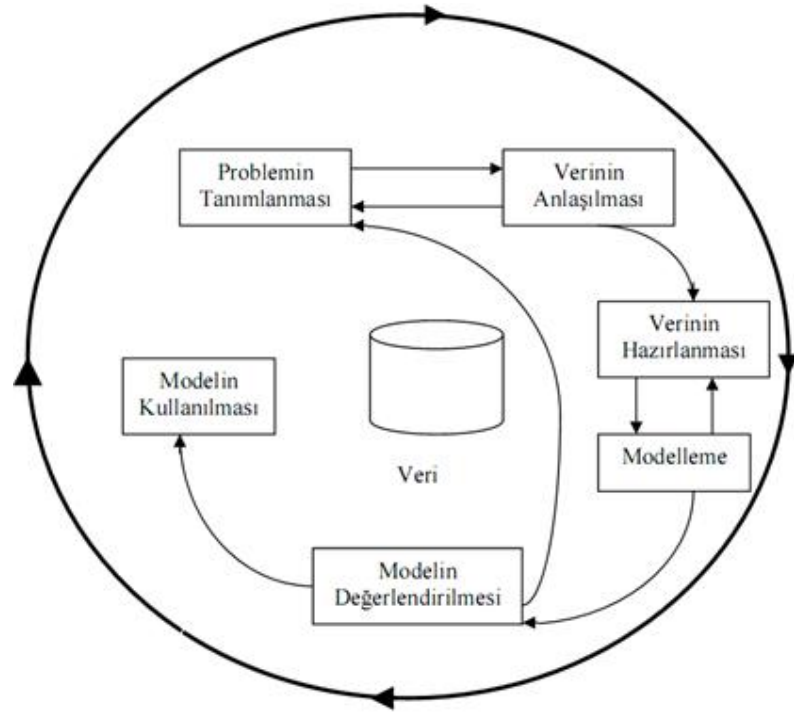
## 2.2. VERİ MADENCİLİĞİ SÜRECİ

İş ortamında, karmaşık veri madenciliği projelerinde bir kuruluştaki birçok uzman ve departman koordineli olarak çalışmaktadır. Dolayısıyla veri madenciliği çalışmalarının başarılı bir şekilde sonuçlanabilmesi için sistematik bir yaklaşım zorunludur. Veri madenciliği literatüründe verilerin toplanması, verilerin analizi, sonuçların yorumlanması, düzenlemelerin yapılması ve sonuçların izlenmesi aşamalarında bu projelerde kullanılabilecek bazı genel çerçeveler sunulmuştur.

Veri madenciliği yazılımları sunan birçok satıcı, danışma firması ve araştırmacı kullanıcılara yönelik birçok süreç modelleri oluşturmuşlardır. Bunlardan en yaygın olarak kullanılanı, veri madenciliği yazılımlarını programlayan ve satan firmalardan bazılarının ve veri madenciliği uygulamalarını işletme faaliyetlerine uyarlayan ilk kuruluşların (Daimler–Chrysler-Almanya, NCR Systems Engineering Copenhagen-Danimarka, SPSS/Integral Solutions Ltd –İngiltere ve OHRA Verzegeringen en Bank Groep B.V.-Hollanda) oluşturduğu bir konsorsiyum tarafından geliştirilen CRISP-DM (Cross-Industry Standard Process for Data Mining - Veri Madenciliği İçin Çapraz Endüstri Standart Süreci)'dir (Fernandez, Stelios ve Walczak, 2002: 788).

CRISP-DM, bilgi keşfi için yapılacak veri madenciliğinin temel adımlarını tanımlayan hiyerarşik bir süreçtir. Veri madenciliği projesi yaşam döngüsünün gözden geçirilmesini sağlamaktadır. CRISP-DM'de veri madenciliği projesinde neler yapılması gerektiği üzerine vurgu yapılmaktadır. CRISP-DM sürecine göre veri madenciliği süreci altı aşamadan oluşan etkileşimli ve yinelemeli bir süreçtir. Aşamalar arasındaki sıra katı ve kesin değildir. Şekil 7'de gösterilen akış şemasının (Şimşek, 2006: 8) herhangi bir aşamasında elde edilen sonuçlara göre, sonraki aşamaya geçilip ya da önceki bir aşamaya dönülüp yeni belirlenen problemlere, ilgi alanlarına göre iyileştirmeler ya da farklı işlemler yapılabilir (Larose, 2005: 5-6).





Şekil 7. CRISP-DM Metodoloji Döngüsü

CRISP-DM modelinde veri madenciliği süreci aşağıdaki adımlardan oluşmaktadır (Larose, 2005: 6-7).

### 2.2.1. Çalışmanın, Araştırmanın Planlanması (İşi anlama)

Veri madenciliği çalışmalarında başarılı olmanın en önemli şartı, projenin hangi amaç için yapılacağına açık bir şekilde tanımlanmasıdır. Proje ilgili 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.

Bu aşamada mevcut problemini nasıl bir sonuç üretilmesi durumunda çözüleceğinin, üretilecek olan sonucun fayda - maliyet analizinin başka bir deyişle üretilen bilginin değerinin doğru analiz edilmesi gerekmektedir. Analistin üretilen sayısal verilerin boyutlarının proje için yeterlilik düzeyinin iyi analiz edilmesi

gerekmektedir. Ayrıca analistin işletmenin iş süreçlerini de iyi analiz etmesi gerekmektedir.

Özetle bu aşamada araştırmanın amaçları ve gerekleri ayrıntılı olarak belirlenmelidir. Daha sonra bu amaçlar ve sınırlılıklar veri madenciliği probleminin tanımını oluşturmaktadır.

### **2.2.2. Verilerin Anlama**

Bu aşamada veriler üzerinde bir ön inceleme yapılmalıdır. Firmanın mevcut bilgi sistemleri üzerinde ürettiği sayısal bilginin iyi analiz edilmesi, veriler ile mevcut iş problemi arasında ilişki olması gerektiği unutulmamalıdır. Proje kapsamında kullanılacak sayısal verilerin, hangi iş süreçleri ile yaratıldığı da bu veriler kullanılmadan analiz edilmelidir; bu sayede analist veri kalitesi hakkında fikir sahibi olabilir. Bu aşamada verinin yapısını anlamak için temel istatistiksel işlemler uygulanabilir. Verinin kalitesi de bu aşamada değerlendirilir.

### **2.2.3. Verilerin Hazırlanması**

Araştırmanın daha sonraki aşamalarında kullanılacak veri seti bu aşamada hazırlanır. Analize uygun veriler ve değişkenler bu aşamada seçilir. Veri madenciliğinin en önemli aşamalarından bir tanesi olan verinin hazırlanması aşaması analistin toplam zaman ve enerjisinin %50 - %75'ini harcamasına neden olmaktadır. Verilerin hazırlanması aşaması kendi içerisinde toplama, birleştirme ve temizleme, dönüştürme adımlarından meydana gelmektedir.

a. Toplama: Tanımlanan problem için gerekli olduğu düşünülen verilerin ve bu verilerin toplanacağı veri kaynaklarının belirlenmesi adımdır. Verilerin toplanmasında kuruluşun kendi veri kaynaklarının dışında, nüfus sayımı, hava durumu, merkez bankası kara listesi gibi veritabanlarından veya veri pazarlayan kuruluşların veri tabanlarından faydalanılabilir.

b. Birleştirme ve Temizleme: Bu adımda toplanan verilerde bulunan farklılıklar giderilmeye çalışılır. Hatalı veya analizin yanlış yönlendirilmesine sebep olabilecek verilerin temizlenmesine çalışılır. Genellikle yanlış veri girişinden veya

bir kereye özgü bir olayın gerçekleşmesinden kaynaklanan verilerin, önemli bir uyarıcı enformasyon içerip içermediği kontrol edildikten sonra veri kümesinden atılması tercih edilir. Ancak basit yöntemlerle ve baştan savma olarak yapılacak sorun giderme işlemlerinin, ileriki aşamalarda daha büyük sorunların kaynağı olacağı unutulmamalıdır.

c. Dönüştürme: Kullanılacak model ve algoritma çerçevesinde verilerin tanımlama veya gösterim şeklinin de değiştirilmesi gerekebilir. Örneğin kredi riski uygulamasında iş tiplerinin, gelir seviyesi ve yaş gibi değişkenlerin kodlanarak gruplanması faydalı olacaktır.

#### **2.2.4. Modelleme aşaması**

Bu aşamada uygun modelleme teknikleri seçilir ve uygulanır. Sonuçların iyileştirilmesi için model ayarlarıyla oynanabilir. Gerekliyse verilerin hazırlanması aşamasına tekrar dönülüp veri tekniğe daha uygun hale getirilebilir.

Bir modelin doğruluğunun test edilmesinde kullanılan en basit yöntem basit geçerlilik (Simple Validation) testidir. Bu yöntemde tipik olarak verilerin bir kısmı 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. (Doğruluk Oranı = 1 - Hata Oranı)

Önemli diğer bir değerlendirme kriteri de modelin anlaşılabilirliğidir. Bazı uygulamalarda doğruluk oranlarındaki küçük artışlar çok önemli olsa da, birçok uygulamasında ilgili kararın niçin verildiğinin yorumlanabilmesi çok daha büyük önem taşıyabilir. Çok ender olarak yorumlanamayacak kadar karmaşık olsalar da, genel olarak karar ağacı ve kural temelli sistemler model tahmininin altında yatan nedenleri çok iyi ortaya koyabilmektedir.

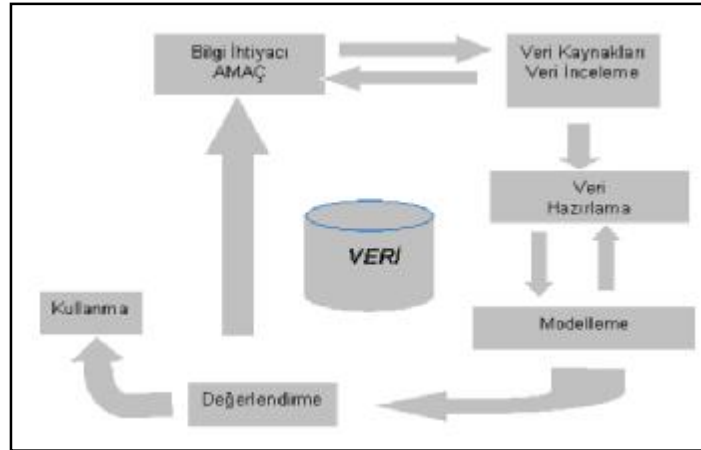
### **2.2.5. Değerlendirme aşaması**

Sonuçlarını kullanmadan önce, modelleme aşamasındaki modeller kaliteye ve verimliliğe göre değerlendirilir. Modelin ilk aşamada belirlenen amaca uygun olup olmadığına bakılır. Araştırma problemine cevap alınamayan noktalar belirlenir. En sonunda bu veri madenciliği sonuçlarının kullanılıp kullanılmayacağına karar verilir.

Kurulan ve geçerliliği kabul edilen model doğrudan bir uygulama olabileceği gibi, bir başka uygulamanın alt parçası olarak kullanılabilir. Kurulan modeller risk analizi, kredi değerlendirme, dolandırıcılık tespiti gibi işletme uygulamalarında doğrudan kullanılabilir gibi, promosyon planlaması simülasyonuna entegre edilebilir veya tahmin edilen envanter düzeyleri yeniden sipariş noktasının altına düştüğünde, otomatik olarak sipariş verilmesini sağlayacak bir uygulamanın içine gömülebilir.

### **2.2.6. Uygulama aşaması**

Bu aşamada kurulan modellerden yararlanır. En basit uygulama bir rapor oluşturulmasıdır. Daha karmaşık bir uygulama olarak ise değişik departmanlarda paralel bir veri madenciliği uygulaması yapılması düşünülebilir. Bu arada, 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. Bu aşamalar arasındaki ilişki Şekil 8'deki gibi özetlenebilir.



Şekil 8. Veri Madenciliği Aşamaları Arasındaki İlişki

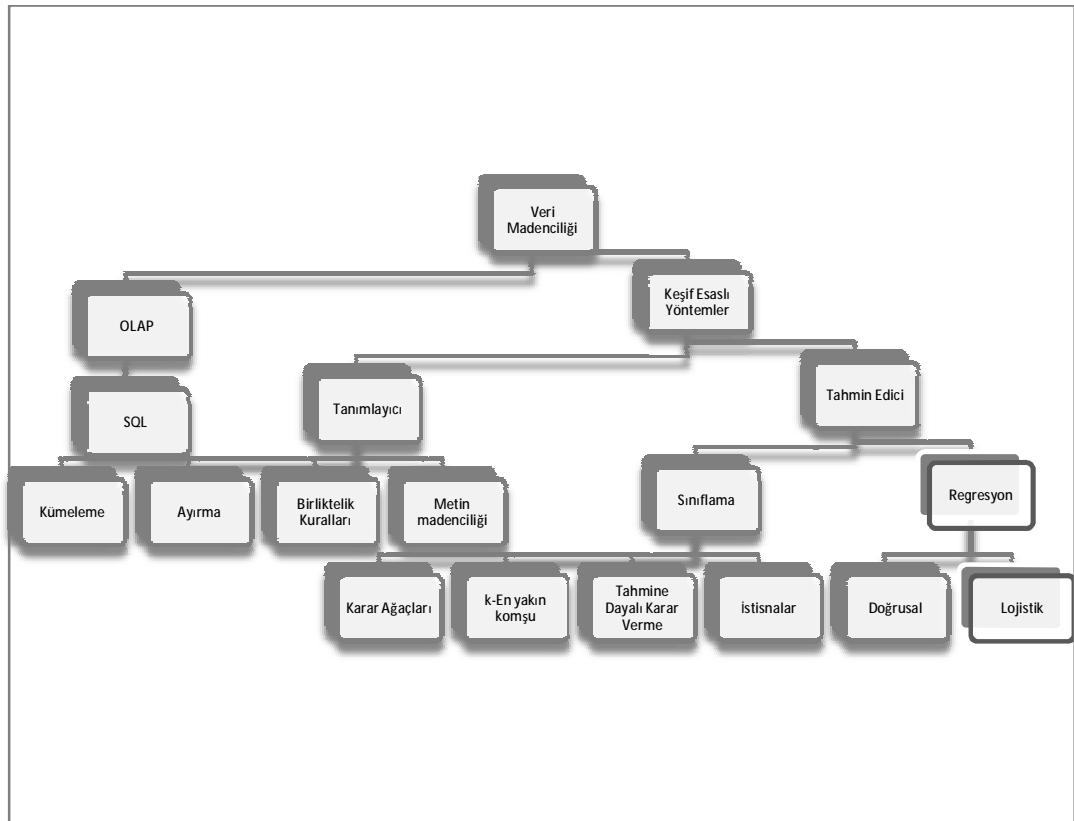
### 2.3. VERİ MADENCİLİĞİ MODELLERİ

Akpınar'a (2000: 5) göre veri madenciliğinde kullanılan modeller, tahmin edici (predictive) ve tanımlayıcı (descriptive) olmak üzere iki ana başlık altında incelenmektedir.

Tahmin edici modellerde, sonuçları bilinen verilerden hareket edilerek bir model geliştirilmesi ve kurulan bu modelden yararlanılarak, sonuçları bilinmeyen veri kümeleri için sonuç değerlerin tahmin edilmesi amaçlanmaktadır. Örneğin bir banka önceki dönemlerde vermiş olduğu kredilere ilişkin tüm verilere sahip olabilir. Bu verilerde bağımsız değişkenler kredi alan müşterinin özellikleri, bağımlı değişken ise kredinin geri ödenip ödenmediğidir. Bu verilere uygun olarak kurulan model, daha sonraki kredi taleplerinde müşteri özelliklerine göre verilecek olan kredinin geri ödenip ödenmeyeceğinin tahmininde kullanılmaktadır.

Tanımlayıcı modellerde ise karar vermeye rehberlik etmede kullanılacak mevcut verilerdeki örüntülerin tanımlanması sağlanmaktadır. X-Y aralığında geliri, evi ve arabası olan, ayrıca çocukları okul çağında olan aileler ile çocuğu olmayan ve geliri X-Y aralığından düşük olan ailelerin satın alma örüntülerinin birbirlerine benzerlik gösterdiğinin belirlenmesi tanımlayıcı modellere bir örnektir (Akpınar, 2000: 6).

Şekil 9’da veri madenciliği modellerinin şematik gösterimi görülmektedir.



Şekil 9. Veri Madenciliği Modelleri

Çeşitli kaynaklarda bu modellerden başka modellere de rastlanmaktadır. Karar verme, ayırma, istisnalar (fark saptaması), metin madenciliği bunlardan bazılarıdır.

### 2.3.1. Sınıflama ve Regresyon

En sık kullanılan veri madenciliği işlerinden biri olan sınıflandırma (Classification), aslında insanlar için bir zorunluluktur. İnsanlar dünyayı anlamak ve iletişime geçmek için sürekli olarak sınıflandırır, kategorilere ayırır ve derecelendirirler. Örneğin canlıları sınıflara, türlere; maddeleri elementlere, insanları ırklara, vb. ayırırlar (Berry ve Linoff, 2004: 8).

Sınıflandırma yeni bir nesnenin özelliklerinin incelenmesi ve bu nesnenin daha önceden özellikleri belirlenmiş bir gruba atanmasıdır. Sınıflandırılması istenilen

nesneler genellikle bir veritabanındaki kayıtlar şeklindedir ve sınıflandırma işi sonunda bir sınıf kodu içeren yeni bir sütun tabloya eklenir. Sınıflandırma işleminde sınıf tanımlarının iyi yapılması ve bu sınıflara örnek olarak verilen bir eğitim veri seti olması önemlidir. Yapılacak iş sınıflandırılmamış bir veri girildiğinde bu veriyi sınıflandırabilecek bir model inşa etmektir.

Bu modelde istenilen bir değişken bağımlı değişken, diğerleri tahmin edici (bağımsız) değişkenler olarak adlandırılır. Amaç, girdi olarak tahmin edici değişkenlerin yer aldığı modelde, çıktının yani bağımlı değişkenin değerinin bulunduğu anlamlı bir model kurmaktır. Bağımlı değişken sayısal değil ise problem sınıflama problemidir. Eğer bağımlı değişken sayısal ise problem regresyon (Regression) problemi olarak adlandırılır (Ganti, Gehrke ve Ramakrishnan, 1999: 40).

Mevcut verilerden hareket ederek geleceğin tahmin edilmesinde faydalanılan ve veri madenciliği yöntemleri içerisinde en yaygın kullanıma sahip olan sınıflama ve regresyon modelleri arasındaki temel fark, tahmin edilen bağımlı değişkenin kategorik veya süreklilik gösteren bir değere sahip olmasıdır.

Aşağıda birkaç sınıflandırma örneği verilmiştir:

- Kredi için başvuran kişilerin düşük, orta, yüksek riskli olarak sınıflandırılması
- Bir web sayfasında yer alacak içeriğin seçilmesi
- Hileli sigorta başvurularının belirlenmesi

Tüm bu örneklerde sınırlı sayıda sınıf vardır ve kayıtların bu sınıflardan birine girebileceği beklenmektedir (Berry & Linoff, 2004: 9).

Sınıflandırma kullanarak pazarlama alanında da araştırmalar yapılabilir. Örneğin: “Genç kadınlar küçük araba satın alır, yaşlı, zengin erkekler büyük, lüks araba satın alır.” Amaç bir malın özellikleri ile müşteri özelliklerini eşlemektir. Böylece bir müşteri için ideal ürün veya bir ürün için ideal müşteri profili çıkarılabilir. Örneğin bir otomobil satıcısı şirket geçmiş müşteri hareketlerinin analizi ile yukarıdaki gibi iki kural bulursa, genç kadınların okuduğu bir dergiye reklam verirken küçük modelinin reklamını verir.

Regresyon modelleri ise bir ya da daha çok deęişkenin başka deęişkenler cinsinden tahmin edilmesini olanaklı kılan ilişkiler bulmak amacıyla geliştirilirler. Örnek olarak, “ev sahibi olan, evli, aynı iş yerinde beş yıldan fazladır çalışan, geçmiş kredilerinde geç ödemesi bir ayı geçmemiş bir erkeğin kredi skoru 825’dir.” Başvuru puanlamada (application scoring) bir finans kurumuna kredi için başvuran kişi ile ilgili finansal güvenilirliğini notlayan, örneğin 0 ile 1000 arasında bir skor hesaplanır. Bu skor kişinin özellikleri ve geçmiş kredi hareketlerine dayanılarak hesaplanır. Sonucu bir regresyon ilişkisidir (Alpaydın, 2000: 2).

Akpınar’a (2000: 5) göre sınıflandırma ve regresyon modellerinde en çok kullanılan teknikler aşağıda sıralanmıştır.

- Karar Ağaçları (Decision Trees),
- Yapay Sinir Ağları (Artificial Neural Networks),
- Genetik Algoritmalar (Genetic Algorithms),
- K-En Yakın Komşu (K-Nearest Neighbor),
- Bellek Tabanlı Yöntemler (Memory Based Reasoning),
- Naive-Bayes,
- Lojistik Regresyon.

### 2.3.2. Kümeleme

İnsanlar önemli kararlar almadan önce genellikle bir adım geri atar ve büyük resmi görmeye çalışır. Ancak zaman zaman bu büyük resim anlaşılacak kadar karmaşıktır. Büyük bir veritabanı çok miktarda boyut, alan içerebilir ve çok karmaşık bir yapıya sahip olduğundan en iyi uygulanabilen veri madenciliği yöntemleri bile bu veri yığını içerisinde anlamlı sonuçlar üretmeyebilir (Dolgun, 2006: 26).

Çok karmaşık ve büyük sorunları çözmekte izlenen yöntem genellikle büyük sorunu daha küçük ve tek başına daha rahat çözülebilecek alt sorunlara bölmek ve her bir alt sorunu çözdükten sonra çözümleri birleştirerek sonuca gitmek şeklindedir.



Kümeleme, verideki benzer kayıtların gruplandırılmasını sağlayan bir tekniktir. Kümeleme analizi hangi yöntem ile yapılırsa yapılsın, süreç aynı şekilde işler. Her kayıt var olan kümelerle karşılaştırılır. Bir kayıt kendisine en yakın kümeye atanır ve bu kümeyi tanımlayan değeri değiştirir. Optimum çözüm bulununcaya kadar kayıtlar yeniden atanır ve küme merkezleri ayarlanır (Akbulut, 2006: 24).

Akpınar da (2000) kümelemeyi sınıfların içerisindeki benzerliklerin maksimize edilip, diğer sınıflar ile olan benzerliklerinin ise minimize edilmesi prensibi ile verilerin gruplandırılması olarak tanımlamaktadır. Kümeleme analizine denetimsiz sınıflama da denebilir. Kümeleme analizi, büyük ölçekli sistemleri daha küçük bileşenlere ayırarak dizayn ve uygulamayı basitleştirmek amacıyla, “böl ve yönet” metodolojisi üzerine kurulmuştur. Küme üyeleri birbirine çok benzer fakat özellikleri birbirinden farklıdır.

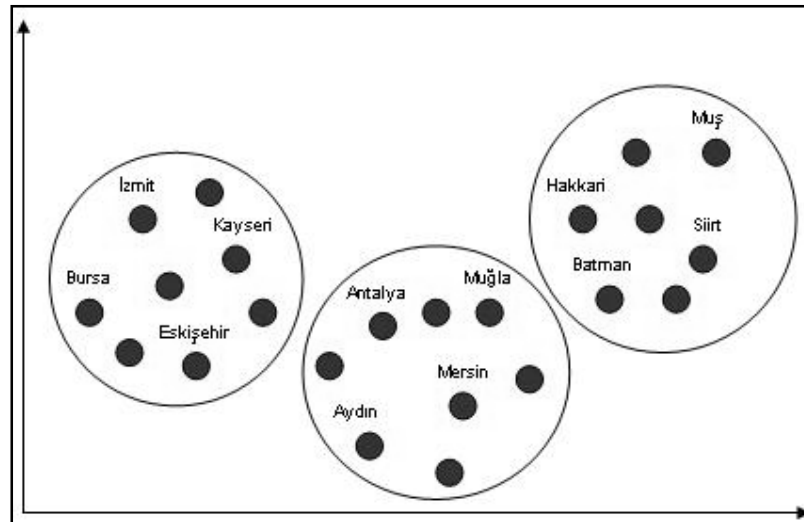
Kümeleme (Clustering) işlemi, heterojen yapıya sahip bir kitleyi daha homojen birkaç alt gruba ya da kümeye bölme işlemidir. Sınıflama ile kümelemeyi birbirinden ayıran en önemli fark, kümeleme işleminin sınıflama işleminde olduğu gibi önceden belirlenmiş bir takım sınıflara göre bölme yapmamasıdır. Sınıflamada her bir veri, önceden sınıflandırılmış bir takım sınıflar üzerinde yapılan bir eğitim neticesinde ortaya çıkan bir modele göre, önceden belirlenmiş olan bir sınıfa atanmaktadır. Kümeleme işleminde ise önceden tanımlanmış sınıflar ya da örnek sınıflar bulunmamaktadır. Verilerin kümelenebilmesi işlemi, verilerin birbirlerine olan benzerliklerine göre yapılmaktadır. Oluşan sınıfların hangi anlamları taşıdığına belirlenmesi tamamen çözümlenmeyi yapan kişiye kalmıştır.

Kümeleme modellerinde amaç, küme üyelerinin birbirlerine çok benzediği, ancak özellikleri birbirlerinden çok farklı olan kümelerin bulunması ve veri tabanındaki kayıtların bu farklı kümelere bölünmesidir. Başlangıç aşamasında veri tabanındaki kayıtların hangi kümelere ayrılacağı veya kümelemenin hangi değişken özelliklerine göre yapılacağı bilinmemekte, konunun uzmanı olan bir kişi tarafından kümelerin neler olacağı tahmin edilmektedir. Örnek olarak bir süpermarketin müşteri bilgileri ve satış kayıtları incelenecek olursa, müşterilerin büyük bir kısmının düzenli

olarak Cuma akşamları kredi kartıyla alışveriş yaptıkları şeklinde bir sonuca ulaşılabilir.

Akpınar da (2000: 12) kümeleme analizinin amacını, varolan verileri içsel olarak homojen, kümeler arası da heterojen olarak ayırmak olarak belirtmiştir. Kümeleme analizi genellikle market araştırmaları, gen araştırmaları gibi konularda sıklıkla kullanılır. Bu veriler arasında nitelik bakımından ilginç korelasyonlar (ilişkiler, benzerlikler) elde edilebilir. Örneğin tıp alanında benzer özelliklere sahip genler aynı küme içine alınabilir.

Örneğin, Türkiye'deki şehirlerin profilini çıkarmak üzere yapılacak bir araştırmada tarıma dayalı bir gelir sistemi olan bir şehri, geliri sanayiye dayalı olan şehirlerle kıyaslamak çok güvenilir sonuçlar vermeyebilir. Aynı şekilde nüfusu milyonlarla ölçülen şehirleri nüfusu yüzbinlerle ölçülen şehirlerle kıyaslamak da yanlıştır. Belirlediğimiz kriterlere göre benzer özellik gösteren şehirler bir gruba toplanır ve kendi aralarında analiz yapılır. Söz gelimi Hakkari'yi Ankara ile kıyaslamak yerine profili benzer olabilecek Siirt, Batman, Muş vb. şehirlerle kıyaslamak çok daha güvenilir sonuçlar elde etmemizi sağlayacaktır (Şekil 10) (İnal, 2008:26).



Şekil 10. Kümeleme Analizi Örneği (İnal,2008:26)

Literatürde birçok kümeleme algoritması bulunmaktadır. Kullanılacak olan kümeleme algoritmasının seçimi, veri tipine ve çalışmanın amacına bağlıdır. Genel olarak başlıca kümeleme yöntemleri şu şekilde sınıflandırılabilir;

- Bölme Yöntemleri (Partitioning Methods)
- Hiyerarşik Yöntemler (Hierarchical Methods)
- Yoğunluk Tabanlı Yöntemler (Density-based Methods)
- Izgara Tabanlı Yöntemler (Grid-based Methods)
- Model Tabanlı Yöntemler (Model-based Methods) (Özekes, 2002)

### **2.3.3. Birliktelik Kuralları ve Ardışık Zamanlı Örüntüler**

Birliktelik kuralları aynı işlem içinde çoğunlukla beraber görülen nesnelere içeren kurallardır.

Birliktelik kuralları (Association Rules) ve ardışık zamanlı örüntüleri (Sequential Patterns) birbirinden ayıran özellik zaman kavramının uygulamada olmasıdır. Belli bir dönem boyunca nesnelere arasındaki birlikteliklerin incelenmesi "ardışık zamanlı örüntü çözümlemesi" olarak da isimlendirilir (Goebel ve Gruenwald, 1999'dan aktaran Dolgun, 2006: 28).

Kumar ve Wahidabanu'ya (2008: 44) göre birliktelik kuralları, bir veritabanındaki bazı olayların birlikte görülmesinin olasılık yapılarıyla gösterilmesidir. Birliktelik kuralları bulunması işlemi, daha önce belirlenen frekans ve doğruluk kriterlerine uyan kuralların belirlenmesidir.

Birliktelik kurallarının bulunması ile pazar sepeti çözümlemesi yapılmaktadır. Pazar sepeti çözümlemesinde, nesnelere müşteriler tarafından satın alınan ürünlerdir ve bir işlem (kayıt) ise birçok nesneyi içinde bulunduran tek bir satın almadır. Pazar sepeti çözümlemesinde sıklıkla beraber alınan nesnelere üzerine çalışılır. Bulunan kurallar ile nesnelere birbiri ile nasıl ilişkili olduğu bilgisine ulaşılır (Dolgun, 2006:28).

Bir alışveriş sırasında veya birbirini izleyen alışverişlerde müşterinin hangi mal veya hizmetleri satın alma eğiliminde olduğunun belirlenmesi, müşteriye daha

fazla ürün satma yollarından birisidir. Bu analiz sonucunda bulunan ilişkiler katalog tasarımlarında, çapraz satışlarda ve müşteri davranışını anlama konusunda yardımcı olabilir. (Han ve Kamber, 2001: 228).

Birliktelik kuralları analizinde önemli bir problem, bir eşik değeri belirlemektir. Bu eşik değerini belirlemek, dolayısıyla bulunabilecek birçok birliktelik kuralı arasından değerli bilgiyi ayırabilmek çok zordur. Bu yüzden ilginç birliktelik kurallarından, ilginç olmayanları ayırt edebilmek için bazı ölçütlerin belirlenmesi gereklidir. Bu ölçütler destek ve güven değerleridir (Adriaans ve Zantinge, 1996: 64).

Birliktelik kuralı madenciliğinin amacı, kullanıcı tarafından belirlenen minimum destek ve güven değerlerini sağlayan kuralların bulunmasıdır. Anlamlılığı destek ve güven değerleri ile ölçülen birliktelik kuralları, "X nesnesini alan bir müşterinin muhtemelen Y nesnesini de alması" tipindeki kuralların tanımlanmasını amaçlamaktadır (Brin, Motwani ve Silverstein, 1997: 3).

Birliktelik kuralları aşağıda sunulan örneklerde görüldüğü gibi eş zamanlı olarak gerçekleşen ilişkilerin tanımlanmasında kullanılır.

- Müşteriler bira satın aldığında, % 75 ihtimalle patates cipsi de alırlar,
- Düşük yağlı peynir ve yağsız yoğurt alan müşteriler, %85 ihtimalle diyet süt de satın alırlar.

Ardışık zamanlı örüntüler ise aşağıda sunulan örneklerde görüldüğü gibi birbirleri ile ilişkisi olan ancak birbirini izleyen dönemlerde gerçekleşen ilişkilerin tanımlanmasında kullanılır.

- X ameliyatı yapıldığında, 15 gün içinde % 45 ihtimalle Y enfeksiyonu oluşacaktır,
- İMKB endeksi düşerken A hisse senedinin değeri % 15'den daha fazla artacak olursa, üç iş günü içerisinde B hisse senedinin değeri % 60 ihtimalle artacaktır,

- Çekiç satın alan bir müşteri, ilk üç ay içerisinde % 15, bu dönemi izleyen üç ay içerisinde % 10 ihtimalle çivi satın alacaktır.

#### **2.3.4. Tahmine Dayalı Karar Verme**

Sınıflandırmada kesin sonuçlar vardır. Eldeki bilgi ve sınıf tanımlarına göre kayıtlar sınıflara ayrılırlar. Karar vermede ise sürekli değer biçilen çıktılar vardır. Bazı giriş verilerinden, bilinmeyen değişkenler hesaplanır. Örneğin bir kredi kartı şirketi bir reklam kampanyası için, müşterilerini kayak yapanlar ve kayak yapmayanlar diye iki sınıfa ayırabilir. Ya da tüm müşterilere “kayak yapma puanı” diye bir puan verebilir. Bu puan müşterinin kayak yapma olasılığını 0 ile 1 arasında bir değerle belirtmektedir. Belli bir değerün üstünde puan alanlar kayak yapanlar yâda yapma ihtimali bulunanlar, altında puan alanlar ise kayak yapmayanlar yâda kayak yapma ihtimalleri bulunmayanlar olarak değerlendirilebilir (Berry ve Linoff, 2004: 9-10).

#### **2.3.5. Ayırma**

Ayırma analizi, iki veya daha fazla sayıdaki grubun ayırımı ile ilgilenen çok değişkenli ilgi analizidir (Şimşek, 2006: 25). Amaçları arasında, analiz öncesi tanımlanmış iki veya daha fazla sayıda grubun ortalama nitelikleri arasında önemli farkların olup olmadığının test edilmesi, gruplar arasındaki farka her bir değişkenin katkısının saptanması ve grup içi değişime oranla gruplar arasındaki ayırımı maksimize eden tahmin değişkenleri kombinasyonunun belirlenmesi sayılabilir. Örneğin, bira içenleri bira içmeyenlerden ayırt etmenin bir pazarlama sorunu olduğu kabul edilirse, büyük bir bira üreticisinin yaptığı araştırma ayırma analizine örnek olarak gösterilebilir. Bu nedenle, tesadüfî olarak seçilen 500 kişilik bir tüketici bölümünü örnek olarak alınmış ve bu kişilerin bira içip içmedikleri, cinsiyetleri ve sporla ilgilenme dereceleri saptanmıştır. Cinsiyet ve sporla ilgilenmenin tahmin değişkenleri olarak kullanılmalarının nedeni, daha önceki çalışmaların bu değişkenlerle bira içme arasında kuvvetli bir ilginin olduğunu göstermiş olmasıdır. Ayırma analizi sonuçlarının test edilme olanağının bulunması sonuçların geçerliliğini ve güvenilirliğini ve dolayısıyla analizin gücünü artıran önemli bir etmendir.

### 2.3.6. İstisnalar (Fark Saptaması)

“Normalden farklı davranış gösteren müşterilerim var mı?” Amaç önceki uygulamaların aksine kural bulmak değil, kurala uymayan istisnai hareketleri bulmaktır. Bu da örneğin olası sahtekârlıkların saptanmasını (fraud detection) sağlar. Örneğin Visa kredi kartı için yapılan CRIS sisteminde, bir yapay sinir ağı kredi kartı hareketlerini takip ederek müşterinin normal davranışına uymayan hareketler için müşterinin bankası ile temasa geçerek, müşteri onayı istenmesini sağlar (Yılmaz, 2006: 83).

### 2.3.7. Metin Madenciliği

“Arşivimde (veya internet üzerinde) bu dokümana benzer hangi dokümanlar var?” Amaç dokümanlar arasında ayrıca elle bir tasnif gerekmeden benzerlik hesaplayabilmektir (text mining). Bu genelde otomatik olarak çıkarılan anahtar sözcüklerin tekrar sayısı sayesinde yapılır (Alpaydın, 2000: 2).

### 2.3.8. Web Madenciliği

Özakar’a (2002: 2) göre Web madenciliği temel olarak üçe ayrılır. Web içerik madenciliği, web’deki veri ve belgelerden faydalı bilgi bulmaya odaklanır. Web yapı madenciliği ise bağlantı (köprü-link) yapılarını bulmaya ve bunların arasındaki ilişkileri modellemeye çalışır. Ama bu iki alan arasındaki fark her zaman net değildir. Üçüncü alan olan Web kullanım madenciliği diğer iki alana göre daha bağımsızdır. Web kullanım madenciliği konusunda şu çalışmalar sürdürülmektedir: Genel Web Kullanım Madenciliği, Site Değişikliği (Modifikasyonu), Sistem İyileştirme ve Kişiselleştirme. Büyük sitelerde her bir çalışma çok önemli sonuçlar elde etmeye yardımcı olabilir.

Genel web kullanım madenciliği log dosyalarını inceleyerek içindeki desenleri bulmaya çalışır. Site Değişikliği (Modifikasyonu) sistemleri sitelerin içerik ve konu yapısında iyileştirmeler önererek, site tasarımını iyileştirmeye çalışır. Sistem İyileştirme, web trafiğini iyileştirmek için web madenciliği yapar. Son olarak,

kişiselleştirme sistemleri kullanıcılarının tercihlerini anlamayı ve web sitelerini kişisel hale getirmeyi amaçlar.

## 2.4. VERİ MADENCİLİĞİNDE KULLANILAN YÖNTEMLER

Literatürde veri madenciliği modellerinde kullanılan yöntemler karar ağaçları, yapay sinir ağları, genetik algoritmalar, k-en yakın komşu algoritması, bellek tabanlı yöntemler, geleneksel sorgu araçları, görsel yöntemler, sapan değer çözümlenmesi ve evrimsel çözümlenme gibi başlıklar altında incelemiştir.

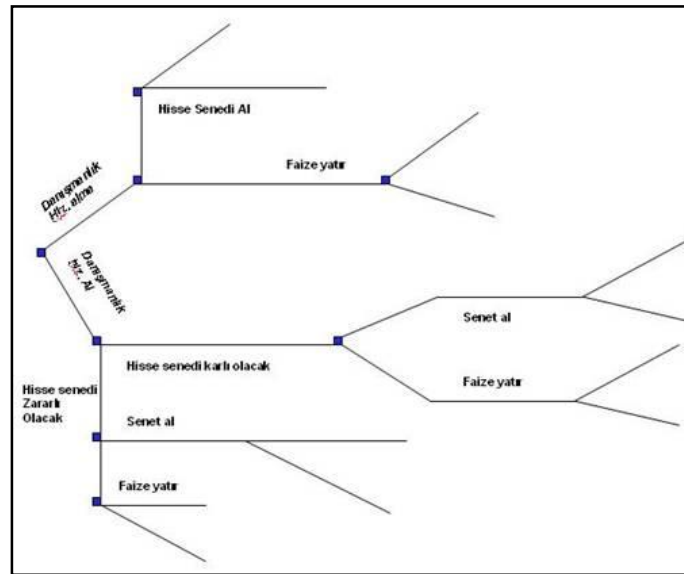
### 2.4.1. Karar Ağaçları

İstatistiksel yöntemlerde veya yapay sinir ağlarında, veriden bir fonksiyon öğrenildikten sonra bu fonksiyonun insanlar tarafından anlaşılabilir bir kural olarak yorumlanması zordur. Karar ağaçları ise veriden oluşturulduktan sonra ağaçta kökten yaprağa doğru gidilerek kurallar (IF-THEN rules) yazılabilir (Eker, 2005). Karar ağacı, adından da anlaşılacağı gibi, ağaç görünümünde, tahmin edici bir yöntemdir.

Karar ağaçlarında her dal kendinden sonra evet veya hayır şeklinde en az iki dala ayrılmaktadır. Karar ağacında bulunan her bir dalın belirli bir olasılığı mevcuttur. Bu sayede son dallardan köke veya istediğimiz yere ulaşana dek olasılıkları hesaplamak mümkündür. Bütün ağacın hata oranı, ayrı ayrı dallar için ağırlıklı hata oranlarının toplamı olarak hesaplanabilir. Hata oranları ve hesaplamaların verimliliği ağacın belirli dalları kesilerek yani daha az yararlı kurallardan kurtularak geliştirilebilir.

Örneğin bir şirket yöneticisi elindeki şirkete dair yüklü bir miktar parayı en yüksek getiriye sağlayacak şekilde faiz veya senet olarak değerlendirmek istiyor. Bu konuda istenirse, bir danışmandan yardım alabilir veya kendi karar verebilir. Olasılıklar çıkartılacak olursa; ilk olarak danışmana başvurulduğunda, danışman “senet al veya faize yatır” seçeneklerini sunacaktır. Bu seçenekler de kendi aralarında başarılı veya başarısız olarak ikiye ayrılacaktır. İlk etapta finans ile ilgili yeterli bilgiye sahip olunmadığı düşünülerek danışmana başvurmak mantıklı

gelecektir. Ama bu seçenek sonucunda danışmana da bir miktar ödeme yapmamız gerekecektir. Diğer yandan danışmanlık hizmeti alınmazsa kendisi karar vermek zorundadır. Hisse senedi kârlı olacak (kârlı olmasına karşın faize yatır veya senet al) veya hisse senedi zararlı olacaktır. Bu olayların karar ağacı Şekil 11'deki gibi olacaktır.



Şekil 11. Karar Ağacı Örneği

Karar ağacı yöntemini kullanarak verinin sınıflanması iki basamaklı bir işlemdir. İlk basamak öğrenme basamağıdır. Öğrenme basamağında önceden bilinen bir eğitim verisi, model oluşturmak amacı ile sınıflama algoritması tarafından çözümlenir. Öğrenilen model, sınıflama kuralları veya karar ağacı olarak gösterilir. İkinci basamak ise sınıflama basamağıdır. Sınıflama basamağında test verisi, sınıflama kurallarının veya karar ağacının doğruluğunu belirlemek amacıyla kullanılır. Eğer doğruluk kabul edilebilir oranda ise, kurallar yeni verilerin sınıflanması amacıyla kullanılır (Özekes, 2002'den aktaran Dolgun, 2006: 29).

Veri madenciliğinde karar ağaçları; kuruluşlarının ucuz olması, yorumlanmalarının kolay olması, güvenilir olması ve veritabanı sistemlerine kolayca eklenebilmeleri nedenleri ile sınıflama modelleri içerisinde en yaygın kullanıma sahiptir.



1970'li yılların başlarında kullanılan AID (Automatic Interaction Detector) karar ağacı temelli ilk algoritma ve yazılımından sonra karar ağacı modellerinin gelişimi çeşitli algoritmalar ile devam etmiştir. Geliştirilen bu algoritmalar içerisinde CHAID (Chi-Squared Automatic Interaction Detector), C&RT (Classification and Regression Trees), ID3, Exhaustive CHAID, C4.5, MARS (Multivariate Adaptive Regression Splines), QUEST (Quick, Unbiased, Efficient Statistical Tree), C5.0, SLIQ (Supervised Learning in Quest), SPRINT (Scalable Parallelizable Induction of Decision Trees) başlıcalarıdır (Akpınar, 2000).

#### 2.4.2. Yapay Sinir Ağları

İnsan beyni çok karmaşık bir yapıya sahiptir. İnsan beyni, yaklaşık  $10^{11}$  sayıda sinir hücresinden (nöron) oluşur. Bu sinir hücreleri çok sayıda birleşme noktasına (synapse) sahiptir. Bir sinir hücresi diğer bir sinir hücresiyle bu birleşme noktaları yoluyla bağlantı kurar.

Yapay sinir ağları (YSA), basit biyolojik sinir sisteminin çalışma şekli simüle edilerek tasarlanan programlama yaklaşımıdır. Simüle edilen sinir hücreleri (nöronlar) içerirler ve bu nöronlar çeşitli şekillerde birbirlerine bağlanarak ağı oluştururlar. Bu ağlar öğrenme, hafızaya alma ve veriler arasındaki ilişkiyi ortaya çıkarma kapasitesine sahiptirler. Diğer bir ifadeyle, YSA'lar, normalde bir insanın düşünme ve gözlemlemeye yönelik doğal yeteneklerini gerektiren problemlere çözüm üretmektedir. Bir insanın, düşünme ve gözleme yeteneklerini gerektiren problemlere yönelik çözümler üretebilmesinin temel sebebi ise insan beyninin ve dolayısıyla insanın sahip olduğu yaşayarak veya deneyerek öğrenme yeteneğidir.

İnsan beyninin karmaşık yapısı daha karmaşık öğrenme makinelerinin oluşturulmasında model olarak alınmış ve yapay sinir ağları olarak adlandırılmıştır. Yapay sinir ağları için farklı yapılar vardır ve bunların her biri verilen işleri yapmak için farklı yol ve öğrenme yöntemleri kullanırlar. Yapay sinir ağının, belirli bir işi yapmak için eğitildiği evreye “şifreleme evresi”, sınıflama ya da kestirim yapma evresine ise “şifre çözme evresi” adı verilir.

Frank Rosenblatt 1958 yılında ilk yapay sinir ağı olarak bilinen algılayıcıyı tasarlamıştır. Ama algılayıcının karmaşık mantıksal fonksiyonlarda işe yaramayacağı

ispatlanınca, yapay sinir ağıları üzerine yapılan çalışmalar hemen hemen durma noktasına gelmiştir. 1980'li yıllara kadar kayda değer bir gelişme olmamasına rağmen, bu dönemden sonra araştırmacılar daha ileri düzeyli yapay sinir ağıları geliştirmeye başlamışlardır. Bu dönemdeki en büyük gelişme, geri yayılım ağılarındaki gizli tabakaların bulunmasıdır (Adriaans and Zantinge, 1996: 68-69).

Yapay sinir ağıları, verideki eğilimleri ve örüntüleri belirlemek için çok uygundur. Bu yüzden tahmin yapmak için de kullanılır (Dolgun, 2006: 31).

Yapay sinir ağıları, insanların önceki deneyimlerinden bir takım bilgiler çıkartması gibi, kendisine verilen örneklerden bir takım bilgiler çıkartma yeteneğine sahiptir. Yapay sinir ağıları karmaşık sorunları çözebilir, ayrıca temel uygulamalardan "öğrenebilir". Yani, soruna kötü bir çözüm bulduysa, ağ bu soruna bir dahaki sefer daha iyi bir çözüm bulacak biçimde değiştirilir. Yapay sinir ağıları öncelikle sonuçları bilinen belirli bir veri kümesi üzerinde öğrenme algoritmaları çalıştırılarak eğitilir. Bu eğitim neticesinde yapay sinir ağının içerisindeki ağırlıklar belirlenir. Bu ağırlıklar kullanılarak yeni gelen veriler işlenir ve bir sonuç üretilir (Tantuğ, 2002).

Yapay sinir ağlarının en olumsuz tarafı ise bu ağırlıkların neden ilgili değeri aldığının bilinmemesi ya da çıkan sonucun neden geçerli bir sonuç olduğunu açıklayamamasıdır. Yapay sinir ağlarını kullanmak için en iyi yaklaşım onları içi bilinmeyen bir şekilde çalışan kara kutular olarak düşünmek olacaktır.

Yapay sinir ağıları; sınıflama, kümeleme ve tahmin analizler ile kolaylıkla kullanılacak genel amaçlı ve güçlü bir tekniktir. Ekonomik alanlardan tıbbi konulara, değerli müşterilerin belirlenmesi için yapılan kümeleme işlemlerinden kredi kartlarında sahtekârlıkların belirlenmesine kadar çok geniş bir alanda uygulanabilmektedir.

Yapay sinir ağlarının veri madenciliği açısından kuvvetli yönleri şunlardır (Tantuğ, 2002);

- Çok geniş açıdaki (spektrum) sorunların çözümünde kullanılabilirler,
- Çok karmaşık durumlarda dahi iyi sonuçlar üretmektedirler,
- Hem sayısal hem de kategorik veriler üzerinde işlem yapabilirler.

Yapay sinir ağlarının veri madenciliği açısından zayıf yönleri de şunlardır;

- 0 ile 1 arasında giriş verileri olması zorunludur,
- Ürettikleri sonuçların açıklamasını yapamazlar,
- Varılan sonucun olası en iyi sonuç olduğunun garantisi yoktur.

### 2.4.3. Genetik Algoritmalar

Genetik algoritmalar, biyoloji alanındaki teorilerden yararlanılarak geliştirilmiş makine öğrenimi yöntemlerinden birisidir. İlk çalışmalar 1960-1970’li yıllarda gerçekleşse de, 1980’li yılların sonuna doğru kabul görmüş bir yöntemdir (Adriaans ve Zantinge, 1996: 72-73).

Genetik algoritmalar açıklanabilir sonuçlar üretirler. Değişik tiplerdeki verileri işleme özelliğine sahip olan genetik algoritmalar en iyileme (optimization) amacı ile kullanılabilirler. Ayrıca genetik algoritmalar yapay sinir ağları ile ortaklaşa çalışarak başarılı sonuçlar üretmektedirler. Genetik algoritmalar yapay sinir ağlarının eğitilmesi, bellek tabanlı yöntemlerde kombinasyon fonksiyonunun oluşturulması gibi işlerde de kullanılmışlardır (Tantuğ, 2002).

Genellikle genetik öğrenme şöyle olur: Rastgele oluşturulmuş kuralları içeren ilk popülasyon (sayı kümesi) oluşturulur. Her kural, bir bit dizisi şeklinde gösterilir. Basit bir örnekle açıklamak gerekirse; eğitim kümesinde A1 ve A2 boolean (evet veya hayır) niteliklerinin ve C1 –C2 sınıflarının verildiğini varsayalım. Kural “IF A1 AND NOT A2 THEN C2” bit dizisiyle 100 olarak ifade edilir. A1 ve A2 soldaki 2 bitle, sınıf da sağdaki bitle gösterilir. Benzer şekilde “IF NOT A1 AND NOT A2 THEN C1” kuralıda 001 sekline kodlanır. Eğer bir nitelik k ( $k > 2$ ) değerlerine sahipse, niteliklerin değerleri k bitleri kullanılarak kodlanabilir. Sınıflar benzer şekilde kodlanabilir. Yeni nesil (offspring), mutasyon ve crossover gibi genetik operatörleri uygulanarak yaratılabilir. Crossover’da kural çiftlerinden yarı diziler, yeni kural çiftleriyle yer değiştirilir. Mutasyonda da, kural dizisinden rastgele seçilmiş bitlerin tersi alınır (invert) (Han ve Kamber, 2001: 351).

Tüm bu olumlu yönlerine rağmen genetik algoritmaların kullanımlarında bazı sıkıntılar da yaşanmaktadır. Bunlardan en belirgin olanı karmaşık sorunların

genetik kodlanmasının çok zor olmasıdır. Ayrıca en iyi (optimal) sonucun üretildiğine dair bir garanti de bulunmamaktadır (Tantuğ, 2002).

#### **2.4.4. k - En Yakın Komşu**

Veri uzayında birbirine yakın olan aynı tip kayıtlar, birbirlerinin komşusu konumundadırlar. Bu tanım doğrultusunda, çok kolay fakat güçlü olan k-en yakın komşu algoritması geliştirilmiştir. k-en yakın komşu algoritmasının temel mantığı “komşunun yaptığını yap”tır. Belirli bir birey veya kayıtın davranışı tahmin edilmek istenirse, veri uzayında o bireye yakın olan örneğin 10 bireyin davranışına bakılır. Bu 10 komşunun davranışının ortalaması hesaplanır ve bu hesaplanan ortalama araştırılan bireyin davranış tahmini olur. K - en yakın komşudaki k harfi incelenen komşu sayısıdır. Örneğin, 5 - en yakın komşuda 5 komşuya bakılır (Adriaans ve Zantinge, 1996: 56-57).

#### **2.4.5. Bellek Tabanlı Yöntemler**

İnsanlar kararlarını genellikle daha önce yaşadıkları deneyimlere göre verirler. Örneğin doktorlar bir hastayı incelerken, elde ettiği bulguları daha önce tedavi ettiği benzer hastalığa yakalanmış hastalar üzerindeki deneyimlerini kullanarak değerlendirirler. Bellek tabanlı yöntemler de benzer şekilde deneyimleri kullanmaktadır. Bu yöntemlerde, bilinen kayıtların bulunduğu bir veritabanı oluşturulur ve sistem yeni gelen bir kayda komşu olan diğer kayıtları belirler ve bu kayıtları kullanarak tahminde bulunur ya da bir sınıflama işlemi uygular. Bellek tabanlı yöntemlerin en önemli özelliği veriyi olduğu gibi kullanabilme yeteneğidir. Diğer veri madenciliği yöntemlerinin aksine bellek tabanlı yöntemler, kayıtların şekli (format) yerine sadece iki işlemin varlığı ile ilgilenir. Bu işlemler, iki kayıt arasındaki uzaklığı belirleyen bir uzaklık fonksiyonu ve komşu kayıtları işleyerek bir sonuç üreten kombinasyon fonksiyonudur (Tantuğ, 2002).

Bellek tabanlı yöntemler sahtekârlık tespiti ve klinik işlemler gibi alanlarda kullanılmaktadır.

Bellek tabanlı yöntemlerin güçlü olduğu noktalar şunlardır:

- Kolayca anlaşılabilir sonuçlar üretir,
- Rastgele seçilen, hatta birbiri ile ilgisiz olabilen verilere bile uygulanabilir,
- Çözümleme alanlarının çok olduğu durumlarda dahi etkili olarak çalışabilir,
- Eğitim kümesinin oluşturulması basittir.

Bellek tabanlı yöntemlerin zayıf olduğu noktalar şunlardır;

- Sınıflama ve tahmin işlemleri için kullanıldığında işlem maliyeti yüksektir,
- Eğitim kümesi için büyük miktarlarda alana ihtiyaç vardır,
- Üretilen sonuçlar; seçilen uzaklık fonksiyonuna, kombinasyon fonksiyonuna ve komşu sayısına doğrudan bağlıdır (Tantuğ, 2002).

#### **2.4.6. Geleneksel Sorgu Araçları**

Bir veri madenciliği projesinde genellikle ilk aşamada geleneksel sorgu araçları kullanılarak veri kümesinin yüzeysel bir çözümlemesi yapılır. Bilinen basit istatistiksel yöntemlerle temel bilgilerin elde edilmesi başlangıç için en kolay ve en iyi yoldur. Veri kümesine sadece basit SQL sorguları uygulanarak, önemli başlangıç bilgileri elde edilebilir. Daha gelişmiş çözümleme algoritmaları (veri madenciliği yöntemleri, vb.) uygulamadan önce veri kümesinin basit görünüşü ve temel yapısının bilinmesi, uygulamadaki bir sonraki adım için yararlıdır. SQL ile veri kümesinden kolayca ulaşılabilecek olan basit, yüzeysel bilgiler ortaya çıkarılabilir ancak saklı bilgi SQL ile bulunamaz. Adriaans ve Zantinge (1996: 48) SQL ile ilgilenilen bilginin yaklaşık %80'inin elde edilebildiğini; kalan %20'lik kısım olan saklı bilgilerin keşfi için daha gelişmiş yöntemler gerektiğini belirtmişlerdir.

#### 2.4.7. Görsel Yöntemler

Görsel yöntemler, veri kümesindeki örüntülerin keşfinde çok kullanışlı bir yöntemdir. Bu yöntemler veri tabanında bilgi keşfi sürecinin başlangıcında kullanılarak verinin kalitesi konusunda ve örüntülerin nerede bulunacağı konusunda bilgi verir (Dolgun, 2006: 34).

#### 2.4.8. Sapan Değer Çözümlemesi

Bir veritabanındaki tüm veriler, tek bir model davranışına uygunluk göstermeyebilir. Bu tür veriler, sapan değer (outlier) olarak adlandırılırlar. Birçok veri madenciliği yönteminde sapan değerler gürültü ya da istisna olarak isimlendirilir ve analizlerde değerlendirilmez. Buna rağmen, bazı uygulamalarda, özellikle sahtekârlık tespitinde az olmuş olaylar (genel yapıya uymayan olaylar), sık oluşmuş olanlara (genel yapıya uyan olaylara) göre daha ilginç ve önemli olabilirler. Sapan verinin çözümlemesi, sapan değer çözümlemesi olarak adlandırılır. Sapan değer çözümlemesinde, veri için genel bir dağılım veya olasılık modeli oluşturup, sonra bu modele uymayan verileri tespit eden istatistiksel testler kullanılabilir (Han ve Kamber, 2001: 26-27).

#### 2.4.9. Evrimsel Çözümleme

Evrimsel çözümleme, zamanla davranışları değişen nesnelerin yapılarını ya da eğilimlerini ortaya çıkarmayı amaçlar. Evrimsel çözümleme; tanımlama, ayırlama, birliktelik, sınıflama ve kümeleme yöntemlerini içerse de; asıl amacı zaman serileri ardışıklık ve periyodiklik örüntüsü bulma, benzerlik çözümlemesi gibi yöntemleri kullanarak verinin zaman ile olan ilişkisini ortaya çıkarmaktır. (Han ve Kamber, 2001: 27).

#### 2.4.10. Naive-Bayes

Naive Bayes, hedef değişkenle bağımsız değişkenler arasındaki ilişkiyi analiz eden tahminci ve tanımlayıcı bir sınıflama algoritmasıdır (Hudairy, 2004'den

aktaran Akbulut, 2006: 21). Naive Bayes, sürekli veri ile çalışmaz. Bu nedenle sürekli değerleri içeren bağımlı ya da bağımsız değişkenler kategorik hale getirilmelidir. Örneğin; bağımsız değişkenlerden biri yaş ise, sürekli değerler “<20” “21-30”, “31-40” gibi yaş aralıklarına dönüştürülmelidir.

Naive Bayes, modelin öğrenilmesi esnasında, her çıktının öğrenme kümesinde kaç kere meydana geldiğini hesaplar. Bulunan bu değer, öncelikli olasılık olarak adlandırılır. Örneğin; bir banka kredi kartı başvurularını “iyi” ve “kötü” risk sınıflarında gruplandırmak istemektedir. İyi risk çıktısı toplam 5 vaka içinde 2 kere meydana geldiyse iyi risk için öncelikli olasılık 0,4’tür. Bu durum, “Kredi kartı için başvuran biri hakkında hiçbir şey bilinmiyorsa, bu kişi 0,4 olasılıkla iyi risk grubundadır” olarak yorumlanır Naive Bayes aynı zamanda her bağımsız değişken / bağımlı değişken kombinasyonunun meydana gelme sıklığını bulur. Bu sıklıklar öncelikli olasılıklarla birleştirilmek suretiyle tahminde kullanılır (Akbulut, 2006: 21-22).

## 2.5. VERİ MADENCİLİĞİNİN UYGULAMA ALANLARI

Veri madenciliği her geçen gün yeni ve farklı alanlarda kullanılmaya başlamakla birlikte, günümüzde yaygın olarak kullanıldığı alanlar birkaç kategoride toplanabilir. Bu kategoriler aşağıda belirtilmiştir (Akbulut, 2006; Canitez, 2007; Dolgun, 2007; Şimşek, 2007).

**Pazarlama:** Müşterilerin satın alma örüntüleri, kampanya ürünleri belirleme, müşterilerin demografik özellikleri arasındaki bağlantıların bulunması, mevcut müşterileri kaybetmeden yeni müşteriler kazanma, posta kampanyalarında cevap verme oranının artırılması, Pazar Sepeti Çözümlemesi, CRM (Customer Relationship Management), müşteri değerlendirme (Customer Value Analysis) ve satış tahmini alanları en yaygın Veri madenciliği uygulama alanlarıdır.

**Banka ve Sigortacılık:** Farklı finansal göstergeler arasında korelasyon tespiti, kredi kartı dolandırıcılıklarının tespiti, kredi taleplerinin değerlendirilmesi, kredi kartı harcamalarına göre müşteri profili belirlenmesi, sigorta dolandırıcılıklarının tespiti ve yeni poliçe talep edecek müşterilerin tahmini gibi alanlarda yoğun olarak kullanılmaktadır.

**Biyoloji, Tıp ve Genetik:** Gen haritasının çözümlenmesi, genetik hastalıkların ve kanserli hücrelerin tespiti, yeni virüs türlerinin keşfi ve sınıflandırılması, fizyolojik parametrelerin çözümlenmesi ve değerlendirilmesinde kullanılmaktadır.

**Kimya:** Yeni kimyasal moleküllerin keşfi ve sınıflandırılması, yem ve ilaç türlerinin keşfinde kullanılmaktadır.

**Yüzey Çözümlemesi ve Coğrafi Bilgi Sistemleri:** Bölgelerin coğrafi özelliklerine göre sınıflandırılması, kentlerde yerleşim yerleri belirleme, kentlerde suç oranı, köken belirleme, kentlere yerleştirilecek posta kutusu, otomatik para makineleri, otobüs durakları gibi hizmetlerin konumlarının tespitinde kullanılmaktadır.

**Telekomünikasyon:** Kalite ve iyileştirme analizleri, hisse tespitleri ve hatların yoğunluk tahminlerinde kullanılmaktadır.

**Meteoroloji ve Atmosfer Bilimleri:** Bölgesel iklim ve yağış haritaları oluşturma, hava tahminleri ve okyanus hareketlerinin belirlenmesinde kullanılmaktadır.

**Metin Madenciliği:** Çok büyük ve anlamsız metin yığınları arasından anlamlı ilişkiler elde etmekte kullanılmaktadır.

**Web Verileri:** İnternet ve web üzerindeki veriler hem hacim hem de karmaşıklık olarak hızla artmaktadır. Sadece düz metin ve resimden başka akan (streaming) sayısal veriler de web verileri arasında yer almaktadırlar (Han ve Kamber, 2001). Bu verilerin çözümlenmesi e-ticaret, web sayfalarının tasarımı ve düzenlenmesi gibi alanlarda kullanılmaktadır.

Çeşitli kaynaklarda bahsedilen, çeşitli sektörlerde, veri madenciliğinin karar verici için olası yararları aşağıdaki gibi sıralanabilir:

- Mevcut müşterilerin karar verici tarafından daha iyi tanınmasını sağlayabilir.
- Özellikle finans sektöründe mevcut müşterileri bölümlere ayırıp, kredi risk davranış modelleri oluşturarak, yeni başvuruda bulunan müşterilere karşı riskin minimize edilmesini sağlayabilir



- Mevcut müşterilerin ödeme performansları incelenip kötü ödeme performansı gösteren müşterilerin ortak özellikleri belirlenerek, benzer özelliklere sahip tüm müşteriler için yeni risk yönetim politikaları oluşturulabilir.
- En iyi müşteriler veya müşteri bölümlerinin bulunmasında kullanılabilir. Bulunan bu iyi müşteri bölümlerine yönelik yeni pazarlama stratejileri oluşturulabilir.
- Kuruluşlar tarafından düzenlenecek çeşitli kampanyalarda mevcut müşteri kitlesinin seçimi ve bu müşterilerin davranış özelliklerine yönelik kampanya şartlarının oluşturulması sağlanabilir.
- Bankacılık faaliyetlerinde, küçük işletmelere yönelik olarak makine ve ekipman satışı yapan dağıtıcı firmalarla ortak hareket ederek oluşturulacak satış paketleri ile pazarlama stratejileri geliştirilebilir.
- Mevcut müşteriler üzerinde firma ürünlerinin çapraz satış kapasitesinin artırılması sağlanabilir.
- Veri madenciliği ile mevcut müşteriyi tanıyarak kuruluşların müşteri ilişkileri yönetimlerinde düzenleme ve geliştirmeler yapılabilir. Bu sayede kuruluşun müşterilerini daha iyi tanıyarak müşteri gibi düşünme kapasitelerinin artırılması sağlanabilir.
- Günümüzde var olan yoğun rekabet ortamında kuruluşların hızlı ve kendisi için en doğru kararı almalarını sağlayabilir.
- Kuruluşlar veri analizi ile müşterilerini kişiselleştirilmiş ürün ve hizmetler hakkında bilgilendirebilirler.
- Veri madenciliği ile kuruluşların müşteriyle bütünleşmiş satış politikaları oluşturması sağlanabilir.

## 2.6. LİTERATÜR TARAMASI

Veri madenciliğinin çok miktarda veri içeren hemen hemen her alanda uygulanabilmesi, kullanım alanını genişletmektedir. Literatür incelendiğinde veri

madenciliğinin birçok bilim dalında çok yoğun bir şekilde uygulandığı göze çarpmaktadır. Örneğin veri madenciliği ve özellikle de bu çalışmada kullanılan birliktelik ve kümeleme analizleri, daha önce **sağlık sektöründe** (Chae, Kima, Tarkb, Parkb, ve Ho, 2003; Özdemir, 2004; Bellazi, Larizza ve Magni, 2005; Phillips-Wren, Sharkey ve Dy, 2008; Karabatak ve İnce, 2008; Kianmehr ve Alhaji, 2008), **bilgisayar ve bilişim sektöründe** (Facca ve Lanzi, 2005), **coğrafya alanında** (Cervone, Kafatos, Napoletani ve Singh, 2006), **trafik mühendisliğinde** (Pande ve Abdel-Aty, 2008), **eğitim alanında** (Özçınar, 2006; Romero, Ventura ve Garcia, 2008; Shih, Chiang, Lai ve Hu, 2008; Chiang, Keh, Huang ve Chyr, 2008), **elektrik mühendisliğinde** (Aydın, 2006) kullanılmıştır.

Birçok bilim dalında yoğun bir şekilde kullanılan veri madenciliği çözümleri aynı zamanda işletmelerdeki değişik problemlerin çözümünde de kullanılmıştır. **Üretim yönetiminde** (Kusiak, 2000; Caskey, 2001; Cox ve Lewis, 2002; Lian, Lai, Lin, ve Yao, 2002; Last ve Kandel, 2004; Ceran, 2006; Jiao, Zhang ve Helande, 2006; Liao ve Perng, 2008), **finans sektöründe** (Lin ve McClean, 2001; Çerkez, 2003; Lee, Chiu, Chou ve Lu, 2004; Boginski ve Butensko, 2006; Lee, Chiu, Chou ve Lu, 2006; Koyuncugil, 2006; Sanchez, Vila, Cerda ve Serrano, 2008; Liao, Ho ve Lin, 2008b), **lojistik yönetiminde** (Hong, Park, Jon ve Rho, 2005; Chen, Huang, Chen ve Wu, 2005; Huang, Chen ve Wu, 2005; Lin ve Chuang, 2008), **insan kaynakları yönetiminde** (Kolay, 2006; Chien ve Chen, 2008), şirket birleşmeleri analizlerinde (Liao, Chang ve Lee, 2008a) çeşitli araştırmalara rastlanmaktadır.

Perakende sektöründe de veri madenciliği son yılların gözde çalışma alanlarından birisidir. Çalışmalar incelendiğinde karar ağacı analizi, sınıflama analizi, lojistik regresyon, birliktelik kuralları analizi, kümeleme analizi gibi yöntemlerin bu çalışmalarda sıklıkla kullanıldığı görülmektedir. **Müşteri ilişkileri** (Hsieh, 2004; Hwang, Jung ve Suh, 2004; Meriç, 2004; Yen ve Lee, 2006; Ahn ve Sohn, 2008; Sohn ve Kim, 2008), **e-ticaret** (Kim, Kim ve Lee, 2002; Aydoğan, 2003; Suh, Lim, Hwang ve Kim, 2004; Chang, Hung ve Ho, 2008; Lazcorreta, Botella ve Fernandez-Caballero, 2008), **şirket değiştirme ve sahtekârlık tespiti** (Bayram, 2001; Akbulut, 2006; Hung, Yen ve Wang, 2006; Kuşaksızoğlu, 2006), **katalog tasarımı, ürün önerme** (Liu ve Shih, 2005; Liao, Hsieh ve Huang, 2008c; Lin ve Hong, 2008) gibi alanlarda çalışmalar gerçekleştirilmiştir. Seyrek (2006),

Canitez (2007) gibi arařtırmacılar da sektörde gerekleřtirilen eřitli veri madencilięi uygulamalarını incelemiřlerdir.

Kaya (2003), Koyuncu (2004), Cavique (2007) gibi arařtırmacılar da birliktelik kuralları algoritmaları üzerinde alıřmıřlar ve bu algoritmaların daha hızlı ve daha verimli alıřabilmeleri iin eřitli yntemler nermiřlerdir.

Ulař (2001) Gima Trk A.ř. marketler zincirine baęlı bir markette, 2000 yılının Haziran, Temmuz ve Aęustos aylarında yapılan alıřveriřleri ieren bir veri kmesi üzerinde veri madencilięi alıřması gerekleřtirmiřtir. Kullandıęı Apriori algoritması ile rn satıřlarına iliřkin analiz yapmıřtır. Ancak sonuları rn isimleri olarak deęil de rn kodları olarak verdięi iin, sonular hakkında bir yorum yapılabilmesi mmkn olmamıřtır.

Russell ve Petersen (2000) rn kategorileri arasındaki talep iliřkisine (tamamlayıcı, baęımsız ve ikame) dayanarak birliktelik analizleriyle bir pazar sepeti analizi gerekleřtirmiřlerdir. Aynı kategorileri ieren ve aynı marka rnleri satan 4 kk market üzerinde yaptıkları alıřmada, apraz kategori fiyat elastikiyetinin dřk olduęunu tespit etmiřlerdir. Dolayısıyla pazar sepeti tahmininde, maęaza trafik dzenlerinin, talep baęlılıklarına gre daha etkili olabileceęi yorumunu yapmıřlardır.

Liao, Chen ve Wu (2008d) perakende sektrnde rn ve marka geniřletme aktiviteleri sırasında, karar alıcıya yol gsterme amacıyla bir veri madencilięi alıřması gerekleřtirmiřlerdir. Birliktelik kuralları ve kmeleme analizlerini kullanarak, mřteri alıřveriř davranıřları hakkında bilgiler elde etmiřler ve bu bilgileri kullanarak rn ve marka geliřtirme sırasında kullanılacak bilgi rnt, kural ve kmelerine ulařmıřlardır.

Chen, Tang, Shen ve Hu (2005) birok maęazadan oluřan bir iřletmede yaptıkları pazar sepeti analizi ile rnlerin maęazadaki yerleřimleri ve satıř istatistikleri arasındaki baęlantıları bulmuřlardır. Geliřtirdikleri yntemle daha byk boyutlu, rn karmasının ok sık deęiřtięi, ok sayıda maęazadan oluřan byk iřletmelerde bile pazar sepeti analizi uygulanabileceęini belirtmiřlerdir.

Dolgun (2006) byk bir alıřveriř merkezinde satılan rnlere ynelik bir sepet analizi alıřması gerekleřtirmiřtir. 8 gnlk bir sre zarfında, toplam 8687 fiř hareketi elde edilmiřtir. Bu fiřler SAS 8.2 paket programının VM iin geliřtirilen

Enterprise Miner modülü ile incelenerek bir veri madenciliği uygulaması yapılmıştır. Çalışmada öncelikle ürünlerin sıklıkları (toplam kaç adet alındıkları) tespit edilmiştir. Daha sonra ise birliktelik analizleri yapılmıştır. Bu analizlerde hangi ürünlerin birlikte satıldığı tespit edilmeye çalışılmıştır. Bunun sonucunda ürünlerin birlikte görülme olasılıkları ve bir ürünü alan bir müşterinin, diğer ürünü de alma olasılıkları hesaplanmıştır. Bu araştırmadaki tespit edilen birkaç birliktelik kuralı aşağıda verilmektedir:

- Yeşillik  $\bar{\cup}$  Sebze kuralı için; yeşillik ve sebzenin toplam fiş hareketlerinde birlikte görülme olasılıkları % 11,29'dur. Yeşillik alan müşterinin % 59,56 olasılıkla sebze de aldığı söylenebilir.
- Zeytin  $\bar{\cup}$  Peynir kuralı için; zeytin ve peynirin toplam fiş hareketlerinde birlikte görülme olasılıkları % 3,67'dir. Zeytin alan müşterinin % 70,11 olasılıkla peynir de aldığı söylenebilir.
- Yeşillik & Soğan  $\bar{\cup}$  Sebze kuralı için; yeşillik, soğan ve sebzenin toplam fiş hareketlerinde birlikte görülme olasılıkları % 6,20'dir. Yeşillik ve soğan alan müşterinin % 71,11 olasılıkla sebze de aldığı söylenebilir.
- Zeytin & Süt  $\bar{\cup}$  Peynir kuralı için; zeytin, süt ve peynirin toplam fiş hareketlerinde birlikte görülme olasılıkları % 1,82'dir. zeytin ve süt alan müşterinin % 77,83 olasılıkla peynir de aldığı söylenebilir.

Şimşek (2006) de Türkiye'de perakende sektöründe önde gelen işletmelerin bir tanesinin Şubat-Mayıs 2004 ayları arasındaki verileri üzerinde, bir veri madenciliği çalışması gerçekleştirmiştir. 628 adet müşterinin toplam 27924 alışveriş hareketi incelenmiştir. Bu veriler üzerinde kümeleme analizi ve birliktelik analizi teknikleri uygulanarak çeşitli bulgulara ulaşılmıştır. Birliktelik kuralları ile ilgili elde ettiği bulgular şunlardır:

- Meşrubat ve kuruyemiş ile unlu mamul alanların %92,464'ü, şekerli ürün de satın almaktadır. Bu kişiler, toplam müşterilerin %54,936'sını oluşturmaktadır.
- Temizlik malzemesi ile şekerli ürün alanların %90,407'si unlu mamul de almaktadır. Bu kişiler, toplam müşterilerin %54,777'sini oluşturmaktadır.
- Süt alanların %91,496'sı, şekerli ürün de almaktadır. Bu kişiler, toplam müşterilerin %54,299'unu oluşturmaktadır.

Bu çalışmada elde edilen bu bulguların yorumlanması sırasında bir hata olduğu düşünülebilir. Gelen müşterilerin yarıdan fazlasının bahsedilen ürün gruplarını alması pek akla yatkın gözükmemektedir. Zaten araştırmacının yorumladığı, Clementine yazılımında hesaplanan “support” değeri, literatürde bahsedilen terim değildir. Literatürde geçen “support” değeri yerine, Clementine yazılımında “rule support” değeri kullanılmaktadır.

Literatür taraması sonucunda ürün kategorileri ve ürün sınıfları arasındaki ilişkileri, uzun süreli verileri analiz ederek kapsamlı bir şekilde inceleyen bir çalışmaya rastlanmamıştır. Dolayısıyla bu çalışmada ürün kategorileri ve ürün sınıfları arasındaki birlikte satış ilişkileri ve ürün sınıflarının birlikte satıldıkları sırada oluşturdukları kümeler, bir işletmenin 1 yıllık satış fişleri verileri incelenerek belirlenecektir.

## ÜÇÜNCÜ BÖLÜM

### BİRLİKTELİK KURALLARI ANALİZİ VE KÜMELEME ANALİZİ

Bu çalışmada veri madenciliği modellerinden birliktelik kuralları analizi ve kümeleme analizi modellerinden yararlanılacaktır. Bu bölümde bu iki analiz tekniğiyle ilgili bilgiler verilmiştir.

#### 1. BİRLİKTELİK KURALLARI ANALİZİ

Kumar ve Wahidabanu'ya (2008: 44) göre birliktelik kuralları, bir veritabanındaki bazı olayların birlikte görülmesinin olasılık yapılarıyla gösterilmesidir. Birliktelik kuralları verinin potansiyel ilişkilerini tanımlar. Amaç birlikte gerçekleşmesi muhtemel olayların kurallarının ortaya konmasıdır.

Birliktelik kurallarının kullanışlı olması için hem konu ile ilgili hem de anlaşılabilir olması gerekir. Birliktelik kuralları sembolik ve sezgisel yapıda olduğundan anlaşılabilirlik her zaman birliktelik kurallarının güçlü yönüdür. Birliktelik kurallarında, kullanıcının kuralların tipini ve sayısını kontrol edebileceği çeşitli yollar vardır. Şimşek'e (2007:51) göre kurallar destek ve güvenilirlik seviyelerine göre karakterize edilirler. Güçlü bir kural çok büyük bir destek ve yüksek seviyeli bir güvenilirliğe sahiptir. Bu yöntemde sadece kullanıcı tarafından belirlenen eşik değerlerinden büyük olan destek ve güven değerlerine sahip kurallar bulunur ve kullanılır. Diğer bir yöntem de kullanıcının sınırlanmış nesne tanımlamasıdır. Sınırlanmış nesne, kuralların içeriğinin sınırlanmasında kullanılan mantıksal bir ifadedir. Örneğin sınırlanmış nesne cips veya hamburger olsun. Bu model sadece cips veya hamburger içeren kurallar ile ilgilenilir. Agrawal ve Srikant (1994:4) sınırlanmış nesne ile kurallar için çeşitli etkin çözümleme yöntemleri geliştirmişlerdir.

Birliktelik kurallarında sıklıkla kullanılan birkaç önemli terim vardır. Bunlar; kuralın sol tarafını ifade eden öncül (antecedent), kuralın sağ tarafını ifade eden izleyen (consequent), destek değeri, güven değeri, “min\_des” olarak gösterilen minimum destek değeri, “min\_güv” olarak gösterilen minimum güven değeri, nesne küme ve sık geçen nesne kümedir.

Pazar sepeti verisi üzerinde birliktelik kuralları çıkarımı problemi ilk olarak Agrawal, Imelinski ve Swami (1993) tarafından ele alınmıştır. Çalışmada birliktelik kuralları, X ve Y'nin nesne kümeler olduğu  $X \Rightarrow Y$  (X birliktelik Y) şeklinde ifade edilmiş ve birliktelik kurallarının matematiksel şekli belirlenmiştir. Bu çalışmada kuralların kullanılabilirliğinin ve gücünün ifadeleri olan destek ve güven değerleri belirlenmiştir. Amaç, kullanıcı tarafından belirlenen minimum destek ve minimum güven değerlerini sağlayan bütün birliktelik kurallarını bulmaktır.

Srikant ve Agrawal'a (1996: 1) göre birliktelik kuralları analizi bütün özelliklerin mantıksal olduğu ilişkili tablolarda "1" değerleri arasında birliktelik bulma olarak düşünülebilir. Tablo her nesneye uygun bir özelliğe ve her işleme uygun bir kayda sahiptir. Kayda uygun işlemde özelliğe uygun nesne var ise, yani özelliğe uygun nesne işlemde var ise bu özelliğin değeri "1", değil ise "0" olarak kayıt edilmektedir.

Eğer elimizdeki veride mallar için sadece satın alındı/alınmadı bilgisi varsa, sepet analizinde mallar arasındaki bağıntı, destek ve güven kıstasları aracılığıyla hesaplanır. İki mal, X ve Y, için destek (support) ve güven (confidence) tanımları şöyledir:

Destek:  $P(X \text{ ve } Y) = X \text{ ve } Y \text{ mallarını satın almış müşteri sayısı} / \text{Toplam müşteri sayısı}$

Güven:  $P(X|Y) = P(X \text{ ve } Y) / P(Y)$

= X ve Y mallarını satın almış müşteri sayısı / Y malını satın almış müşteri sayısı

Destek veride bu bağıntının ne kadar sık olduğunu, güven de Y malını almış bir kişinin hangi olasılıkla X malını alacağını söyler. Bağıntının önemli olması için her iki değer de olabildiğince büyük olması gerekir.

Eğer bir nesnekümede 2 adet nesne var ise karşımıza en fazla iki adet birliktelik kuralı çıkacaktır.  $A \rightarrow B$  Ve  $B \rightarrow A$ . Eğer A, B ve C gibi 3 nesne var ise ortaya çıkabilecek birliktelik kuralı sayısı Tablo 4’de görüldüğü gibi 12’ye çıkmaktadır (Teknomo, 2008: 2).

*Tablo 4. A, B ve C Nesneleri Arasındaki Birliktelik Kuralları*

	<b>X</b>	<b>→</b>	<b>Y</b>
1	A	→	B
2	A	→	C
3	A	→	B, C
4	B	→	A
5	B	→	C
6	B	→	A, C
7	C	→	A
8	C	→	B
9	C	→	A, B
10	A, B	→	C
11	A, C	→	B
12	B, C	→	A

Nesne sayısı 4’e çıktığında ortaya çıkması muhtemel kural sayısı 50’ye çıkmaktadır. Nesne sayısına (d) göre ortaya çıkması muhtemel birliktelik kuralı sayısı (R) aşağıdaki formüle göre hesaplanabilir (Tan, Steinbach & Kumar, 2006: 331).

$$R = 3^d - 2^{d+1} + 1$$

*Tablo 5. Nesne Sayısına Göre Ortaya Çıkabilecek Birliktelik Kuralları Sayısı*

<b>Toplam nesne sayısı, d</b>	<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>	<b>5</b>	<b>10</b>	<b>100</b>	<b>500</b>
<b>Ortaya çıkması muhtemel Birliktelik kuralı sayısı, R</b>	0	2	12	50	180	57002	5.15378E+47	3.636E+238

Tablo 5 nesne sayısına göre ortaya çıkabilecek olası birliktelik kuralları sayısını vermektedir.



Örneğin bir markette alışveriş yapan müşterilerden 5 tanesinin alışveriş sepetlerinin Tablo 6'daki gibi olduğu varsayılırsa (Özkan, 2008: 159);

*Tablo 6. Örnek Alışveriş Sepetleri*

Müşteri	Alışveriş sepetindeki ürünler
1	Makarna, yağ, meyve suyu, peynir
2	Makarna, Ketçap
3	Ketçap, yağ, meyve suyu, bira
4	Makarna, ketçap, yağ, meyve suyu
5	Makarna, ketçap, yağ, bira

Bu verilerden yararlanarak birliktelik kuralları analizi gerçekleştirildiğinde aşağıdaki biçimde sonuçlar elde edilir:

- {ketçap, meyve suyu} → {yağ} (Destek=0,4 Güven=1,0)
- {ketçap, yağ} → {meyve suyu} (Destek=0,4 Güven=0,67)
- {Yağ, meyve suyu} → {ketçap} (Destek=0,4 Güven=0,67)
- {meyve suyu} → {ketçap, yağ} (Destek=0,4 Güven=0,67)
- {yağ} → {ketçap, meyve suyu} (Destek=0,4 Güven=0,5)
- {ketçap} → {yağ, meyve suyu} (Destek=0,4 Güven=0,5)

Bu sonuçlar da aşağıdaki şekilde yorumlanabilir:

- Ketçap ve meyve suyunu birlikte alanlar mutlaka yağ da almaktadır.
- Ketçap ve yağ satın alan müşteriler %67 olasılıkla meyve suyu da almaktadır.
- Yağ ve meyve suyunu birlikte alan müşteriler %67 olasılıkla ketçap da almaktadır.
- Meyve suyu alanlar %67 olasılıkla ketçap ve yağ da satın almaktadır.
- Yağ alanlar %50 olasılıkla ketçap ve meyve suyu da almaktadır.
- Ketçap alanlar %50 olasılıkla yağ ve meyve suyu da almaktadır.

Birliktelik kuralının matematiksel modeli Agrawal, Imielinski ve Swami (1993:2) tarafından ifade edilmiştir. Bu modele göre;  $I = \{i_1, i_2, i_3, \dots, i_m\}$  nesnelerin kümesi ve  $D$  işlemler kümesi olarak ifade edilir. Her  $i$ , bir nesne (ürün) olarak adlandırılır.  $D$  veritabanında her işlem  $T$ ,  $T \subseteq I$  olacak şekilde tanımlanan nesnelerin kümesi (nesne küme) olsun. Her işlem bir tanımlayıcı alan olan TID ile temsil edilir.  $A$  ve  $B$  nesnelerin kümeleri olsun. Bir  $T$  işlemler kümesi ancak ve ancak  $A \subseteq T$  ise yani  $A$ ,  $T$ 'nin alt kümesi ise  $A$ 'yı kapsıyor denir. Bir birliktelik kuralı  $A \Rightarrow B$  formunda ifade edilir.  $A$  öncül (andecent) ve  $B$  izleyen (consequent) olarak adlandırılır. Burada,  $A \subset I$ ,  $B \subset I$  ve  $A \cap B = \emptyset$  dır.

İlk olarak,  $A \Rightarrow B$  kuralı için  $d$  olasılığı ile kuralın destek değeri tanımlanır. Destek,  $T$  işleminin  $A \cup B$ 'yi içermeye olasılığıdır. İkinci olarak,  $A \Rightarrow B$  kuralının  $g$  ile gösterilen güven değeri tanımlanır. Bu olasılık,  $T$  işleminin  $A$ 'yı ve aynı zamanda  $B$ 'yi içermesidir.

Matematiksel ifade ile kuralın destek ve güven değerleri aşağıdaki şekilde ifade edilir.

$$\text{Destek } (A \Rightarrow B) = P(A \cup B)$$

$$\text{Güven } (A \Rightarrow B) = P(B \setminus A) \text{ veya}$$

$$\text{Güven } (A \Rightarrow B) = \text{Destek } (A \Rightarrow B) / \text{Destek } (A)$$

$$\text{Burada Destek } (A) = \text{Destek } (A \Rightarrow A) \text{ dır.}$$

Başka bir ifade ile destek ve güven değerleri aşağıdaki şekilde tanımlanır:

$$\text{Destek } (A) = |A| / |D|$$

$$\text{Destek } (A \Rightarrow B) = |A.B| / |D|$$

$$\text{Güven } (A \Rightarrow B) = \text{Destek } (A \Rightarrow B) / \text{Destek } (A)$$

Burada;  $|A|$ , incelenen kayıtlardaki  $A$  ürününü içeren işlemlerin sayısı.  $|A.B|$ , incelenen kayıtlardaki  $A$  ve  $B$  ürünlerini birlikte içeren işlemlerin sayısı.  $|D|$ , veritabanındaki bütün işlemlerin sayısını ifade etmektedir.

Kuralın destek ve güven değerleri, kuralın ilginçliğini ifade eden iki ölçüdür. Bu değerler sırasıyla keşfedilen kuralların yararlılığını (kullanışlılığını) ve

kesinliğini (doğruluğunu) ifade eder. Destek değeri, A ve B nesnelerinin birlikte bulunma olasılıklarını ifade eder. Güven değeri ise, A'yı içeren kayıtların B'yi de içereceğini ifade eder. Yüksek güven ve destek değerine sahip kurallara güçlü (strong) kurallar adı verilir (Agrawal vd., 1993: 1).

Kullanıcı tarafından minimum destek eşik değeri (min\_des) ve minimum güven eşik değeri (min\_güv) belirlenebilir. Bu belirlenen eşik değerlerini aşan birliktelik kuralları dikkate alınır ve ilginç olarak ifade edilir. İlginç bir örüntü, bilgi (knowledge) olarak ifade edilir. Genel olarak bu değerler 0-1 aralığından çok %0-%100 aralığında ifade edilmektedir (Han ve Kamber, 2001: 230).

Nesnelerin kümesi, nesneküme olarak ifade edilmektedir. Eğer bir küme k tane nesne içeriyorsa bu küme "k-nesneküme" olarak ifade edilir. Eğer bir nesneküme minimum destek değerini sağlıyor ise, bu nesneküme sık geçen nesneküme olarak adlandırılır. k-nesnekümenin sık geçen kümeleri  $L_k$  şeklinde ifade edilmektedir (Han ve Kamber, 2001: 235).

Minimum güven ve destek değerlerini sağlayan birliktelik kuralları çıkarım problemi iki adıma bölünmüştür (Agrawal vd., 1993; Agrawal ve Srikant, 1994).

1. Sık geçen nesnekümelere bulunur: Kullanıcı tarafından belirlenmiş olan minimum destek eşik değerini sağlayan nesnekümelere sık geçen nesneküme adı verilmektedir. Bu adımda sık geçen nesnekümelere bulan etkili yöntemler kullanılmalıdır.

2. Sık geçen nesnekümelere güçlü birliktelik kuralları oluşturulur: Sık geçen nesnekümelere kullanarak minimum güven eşik değerini sağlayan birliktelik kurallarının bulunmasıdır. Bu adımdaki işlem oldukça basittir. Minimum güven eşik değerine göre taranarak bulunan birliktelik kuralları kullanıcının ilgilendiği ve potansiyel olarak önemli bilgiyi içeren kurallardır. Birliktelik kuralı algoritmasının performansını belirleyen adım birinci adımdır. Sık geçen nesnekümelere belirlendikten sonra, birliktelik kurallarının bulunması kolay bir adımdır.

Büyük veri kümelerini ortaya çıkarma algoritmalarında, veri üzerinden çok sayıda geçiş yapılır. İlk geçişte her bir verinin destekleri sayılır ve desteği en az olan büyük veri kümeleri ortaya çıkarılır. Sonraki geçişlerde bulunmuş büyük veri kümeleri ile başlanır. Her geçişte bulunmuş kümeler yeni aday veri kümesi siye

adlandırılan daha da büyük kümeleri yaratmakta kullanılır ve bunların destekleri sayılır. Geçiş sonunda en büyük aday veri küme ortaya çıkarılır ve sonraki geçişte kullanmak için hazır tutulur. Bu işlem yeni büyük veri kümesi bulununcaya kadar devam edilir (Agrawal ve Srikant, 1994: 5-6).

## 1.1. BİRLİKTELİK KURALLARINDA KULLANILAN ALGORİTMALAR

Birliktelik kuralları analizinde bu işlemlerin yapılıp, kuralların çıkartılabilmesi için çeşitli algoritmalar geliştirilmiştir. AIS algoritması (Agrawal, 1993:3) 1993'te ortaya atılmış, daha sonra SETM algoritması (Houtsma ve Swami, 1995: 28-30) ve 1994'de ortaya atılan ve bugün de çok kullanılan Apriori ve AprioriTid (Agrawal ve Srikant, 1994: 5) algoritmaları geliştirilmiştir.

### 1.1.1. AIS Algoritması

AIS algoritması, geniş nesne kümeleri üretmek için geliştirilmiş bir algoritmadır. 1993 yılında geliştirilmiştir. Veritabanındaki isimlerin, yani nesne isimlerinin A'dan Z'ye sıralanması kısıtını taşır.

AIS algoritması veritabanını birçok kez tarar ve her tarama esnasında tüm işlemleri okur. İlk tarama esnasında veritabanındaki nesnelere, teker teker sayarak hangilerinin geniş nesnelere (çok rastlananlar) olduğunu belirler. Bunlardan geniş olanlar aday nesne kümeleri olarak işaretlenir. Bir işlem tarandıktan sonra, bir önceki taramada geniş oldukları belirlenen nesne kümeleriyle, o işlemin nesnelere arasındaki ortak nesne kümeleri belirlenir. Belirlenen bu ortak nesne kümeleri işlemde mevcut olan diğer nesnelere birleştirilerek yeni aday kümeler oluşturulur. Herhangi bir nesne kümesi, bir işlemdeki nesnelere birleşip aday kümelerden birini oluşturabilmesi için, birleşeceği nesnenin hem geniş olması hem de harf sırası açısından nesne kümesi içindeki tüm nesnelere sonra geliyor olması gerekir.

AIS algoritması bu adımı gerçekleştirebilmek için, bir budama tekniği kullanır. Budama tekniğinin özünde, aday kümeler içindeki gereksiz kümelerin silinmesi vardır. Bu adımdan sonra, her aday kümesinin desteği hesaplanır. Destek seviyeleri minimum destek seviyesine eşit veya bu seviyeden büyük çıkanlar geniş

nesne kümesi olarak işaretlenirler. Bir sonraki taramada bu geniş işareti taşıyan kümeler, yukarıda anlatıldığı gibi bir sonraki aday kümelerin belirlenmesi için kullanılır (Agrawal, 1993: 5-6).

### 1.1.2. SETM Algoritması

Bu algoritmada,  $L_k$  geniş nesne kümesinin herbir elemanı iki parametreden oluşur; bunların bir tanesi nesnenin ismiyken diğeri bu nesneyi ayırt etmeye yarayan bir özellik numarasıdır. Algoritma içinde bu numara TID (Transaction Identification) olarak kullanılır. Benzer şekilde her bir aday nesne kümeleri,

$C_k, \langle TID, Nesne Kümesi İsmi \rangle$

formatında tutulur. (Houtsma ve Swami, 1995: 28:30)

AIS algoritmasında olduğu gibi SETM algoritması da veritabanını birçok kez tarar, ilk tarama esnasında veritabanındaki nesnelere, teker teker sayarak hangilerinin geniş nesnelere olduğunu belirler. Sonraki taramalarda, bir önceki geçişte geniş olarak işaretlenmiş nesne kümelerini kullanarak aday kümeleri belirler. Farklı olarak, SETM algoritması aday kümelerle birlikte üzerinde çalışılan işlemin TID bilgisini de tutar. Bundan sonra, aday nesne kümeleri nesne ismine göre sıraya dizilir ve küçük nesne kümeleri silinir. Eğer veritabanı TID numarasına göre sıralanmışsa, bir sonraki tarama esnasında herhangi bir işlemdeki geniş nesne kümeleri  $L_k$  'nin TID numarasına göre sıralanmasıyla elde edilir. Bu şekilde veritabanı bir kaç kez taranır. Artık başka herhangi bir geniş nesne kümesi bulunamadığında algoritma sonlandırılır.

SETM algoritmasında TID bilgisinin de tutulması, algoritmanın yer karmaşıklığını arttıracaktır, bu dezavantajın dışında başka bir eksi nokta ise, aday nesne kümesinin desteği hesaplanırken  $C_k$  sıralanmış halde değildir, bunun için nesne kümelerinin bir kez daha sıraya dizilmesi gerekecektir. Bu da zaman karmaşıklığını arttıran bir unsurdur.

### 1.1.3. Apriori Algoritması

Apriori algoritması bağlantı analizlerinin yapılp bağlantı kurallarının ortaya çıkartılması konusunda en çok bilinen ve kullanılan algoritmadır. Geniş nesne kümelerinin ortaya çıkartılması işlemleri için kullanılır. Algoritmanın ismi, sık geçen nesnelerin önsel bilgilerini kullanmasından yani bilgileri bir önceki adımdan almasından “önceki (prior)” anlamında aprioridir.

Geniş nesne kümelerini ortaya çıkartan algoritmalar eldeki tüm verileri birçok kez tararlar. İlk taramada, her bir nesnenin destek seviyesi, hesaplanarak kullanıcı tarafından başlangıçta girilen minimum destek seviyesi ile karşılaştırılır ve her bir nesnenin geniş olup olmadığına bakılır. Bundan sonraki her tarama bir önceki taramada geniş olarak tespit edilmiş nesnelere başlar ve geniş nesne kümeleri oluşturulur. Bu geniş nesne kümelerine aday nesne kümeleri denir. Taramanın sonunda ise hangi aday nesne kümesinin gerçekten geniş olduğu kontrol edilir. Daha önce de belirtildiği gibi bir nesne kümesinin geniş olarak adlandırılabilmesi için o nesne kümesinin kullanıcı tarafından verilen minimum destek seviyesinin üzerine bir destek seviyesine sahip olması gerekir. Bir sonraki taramada, yine bir önceki taramada geniş olarak seçilen nesne kümelerinden başlanır ve veritabanının sonuna kadar bu nesne kümelerinin destekleri hesaplanır. Bu işlem, başka yeni geniş nesne kümeleri bulunamayana kadar sürer. (Agrawal ve Srikant, 1994: 5-6)

Apriori algoritması daha önceden ortaya atılmış olan AIS ve SETM algoritmalarından her bir geçişte aday nesne kümelerinin sayılma ve bu aday kümelerinin üretilme şekliyle ayrılır. Hem AIS algoritmasında, hem de SETM algoritmasında, tarama esnasında, veriler okunurken aday nesne kümeleri üretilir. Bir işlem (T) (transaction) okunduktan sonra, geniş nesne kümelerinin bu işlemlerde olup olmadığına da bakılır. Yeni aday nesne kümelerinin üretilmesi ise işlemlerdeki diğer nesnelerle elde edilen geniş nesne kümelerinin birleştirilmesiyle üretilir (Agrawal, 1993). Tabi bu da, gereksiz yere, aslında küçük nesne kümesi olan birçok aday nesne kümesinin sanki geniş nesne kümesiymiş gibi üretilmesi ve sayılması sonucunu doğurur. Bu da algoritmanın zaman karmaşıklığını arttırır.

Apriori algoritması ise aday nesnelere üretirken veritabanındaki işlemleri hiç işin içine sokmadan, yalnızca bir önceki taramada geniş olduğu tespit edilmiş nesne

kümelerini kullanarak oluşturur. Apriori algoritması geniş bir nesne kümesinin herhangi bir alt kümesinin de geniş olacağı kabulüne dayanır. Böylece k adet nesneden oluşmuş bir nesne kümesi, k-1 adet nesneye sahip geniş nesne kümelerinin birleştirilmesi ve alt kümeleri geniş olmayanların silinmesiyle elde edilebilir. Bu birleşme ve silme işlemi sonunda daha az sayıda aday nesne kümeleri oluşacaktır.

Apriori, seviye mantığı (level-wise) arama olarak bilinen yinelemeli bir yaklaşım kullanır. Bu yaklaşımda k-nesnekümeler, (k+1)-nesnekümelerin araştırılması için kullanılır. İlk olarak, sık geçen 1-nesnekümelerin kümesi bulunur. Bulunan bu küme  $L_1$  olarak adlandırılır.  $L_1$ ,  $L_2$ 'nin (sık geçen 2-nesnekümelerin kümesi) bulunmasında kullanılır.  $L_2$ ,  $L_3$ 'ün bulunmasında kullanılır ve algoritma bu şekilde daha fazla sık geçen k-nesnekümeler bulamayınca kadar yinelemeli bir şekilde devam eder. Her  $L_k$ 'nin bulunması bütün veritabanınının taranması anlamına gelmektedir (Han ve Kamber, 2001: 235).

Sık geçen nesnekümelerin seviye mantığı (level-wise) oluşturma ile etkinliğini artırmak için önemli bir özellik olan “apriori özelliği” kullanılır. Bu özellik ile arama alanı küçültülür.

Apriori algoritması sırasında birleştirme (join) ve budama (prune) basamakları gerçekleşmektedir.

1. Birleştirme Adımı:  $L_k$ 'yi bulmak için kullanılan aday k-nesne kümelerinin kümesi ( $C_k$ ),  $L_{k-1}$ 'in kendi arasında birleştirilmesi ile oluşturulur.  $I_1$  ve  $I_2$ ,  $L_{k-1}$ 'de bulunan nesnekümeler olsun.  $I_1[j]$  gösterimi,  $I_i$  'deki j'nci nesneyi temsil eder (örneğin;  $I_1[k-2]$ ,  $I_1$ deki sondan 2. nesneyi ifade eder). Geleneksel olarak Apriori, bir işlem veya nesnekümedeki nesnelerin sözcük sırasında sıralandığını varsayar.  $L_{k-1}$ 'in ilk (k-2) nesnelere ortak ise  $L_{k-1}$ 'in elemanları birleştirilebilir. Ancak bu durum altında  $L_{k-1} \infty L_{k-1}$  birleştirme işlemi gerçekleştirilebilir. Bu durum;  $L_{k-1}$ 'in  $I_1$  ve  $I_2$  nesnekümelere eğer,

$$(I_1[1] = I_2[1]) \cap (I_1[2] = I_2[2]) \cap \dots \cap (I_1[k-2] = I_2[k-2]) \cap (I_1[k-1] = I_2[k-1])$$

ise birleştirme işlemi gerçekleştirilebilir anlamına gelmektedir.  $I_1[k-1] < I_2[k-1]$  koşulu aynı adayların oluşmamasını garanti etmektedir.  $I_1$  ve  $I_2$ 'nin birleştirilmesiyle oluşan nesnekümesinin sonuçları  $I_1[1]I_1[2] \dots I_1[k-1]I_2[k-1]$  dir.

2. Budama Adımı:  $C_k$ ,  $L_k$ 'nin bir üstkümesidir, aday kümenin ( $C_k$ ) elemanları sık geçen olabilirler veya olmayabilirler, fakat sık geçen k-nesnekümelerinin hepsi  $C_k$  da yer alır. Veritabanında yapılan bir taramada  $C_k$ 'daki her adayın sayısının saptanması  $L_k$ 'nin belirlenmesi ile sonuçlanır (örneğin, minimum destek sayısından küçük olmayan sayıya sahip bütün adaylar tanım gereği sık geçendir ve bundan dolayı  $L_k$  'da yer alır).  $C_k$  çok büyük boyutta olabilir, bu durum da taramanın çok ağır ilerlemesine sebep olur.  $C_k$ 'nin boyutunun azaltılması için "Apriori Özelliği" aşağıdaki gibi kullanılır. Herhangi bir sık geçen olmayan (k-1)-nesneküme (k-1 ögeli küme), sık geçen k-nesnekümenin altkümesi olamaz. Bundan dolayı, aday k-nesnekümenin (k-1) ögeli altkümeleri  $L_{k-1}$  kümesinin ögesi değilse, aday sık geçen değildir ve  $C_k$  aday kümesinden çıkartılır. Bu çıkarma işleminin sonucunda  $C_k$  aday kümesinin ögeleri azalır, dolayısıyla boyut da azalır. Bu altküme testi hash ağacı yöntemi ile daha hızlı bir şekilde yapılabilir (Han ve Kamber, 2001).

Yukarıda anlatılan algoritmanın Tablo 7'de görülen alışveriş fişleri üzerinde uygulanışı aşağıda açıklanmaktadır (Dolgun, 2006: 43-47).

*Tablo 7. Alışveriş Numaraları ve Alınan Ürünlerin Tablosu*

Alışveriş No	Alınan Ürünler
ANO100	$I_1, I_2, I_5$
ANO200	$I_2, I_4$
ANO300	$I_2, I_3$
ANO400	$I_1, I_2, I_4$
ANO500	$I_1, I_3$
ANO600	$I_2, I_3$
ANO700	$I_1, I_3$
ANO800	$I_1, I_2, I_3, I_5$
ANO900	$I_1, I_2, I_3$

Tablo 7'te bir marketten yapılan alışverişlerin bilgilerini içeren bir veritabanı görülmektedir. Bu veritabanında yapılan alışverişlerin numaraları Alışveriş No sütununda, her alışverişte satın alınan ürünler de Alınan Ürünler sütununda ifade edilmiştir. Apriori algoritması aşağıdaki şekilde çalışacaktır:



1. Algoritmanın ilk adımında; her nesne, 1-nesnekümeler (1 ögeli kümeler) aday kümesi olan  $C_1$  aday kümesinin bir elemanıdır. Algoritma, her ürünün sayısını bulmak için bütün veritabanını (D) tarar ve bu sayıları destek sayıları olarak belirler. Elde edilen sonuçlar Şekil 12'de Destek Sayısı sütununda görülmektedir. Tablo 8'de görülebileceği gibi D'de  $I_1$  ürününden 6 adet,  $I_2$  ürününden 7 adet,  $I_3$  ürününden 6 adet,  $I_4$  ürününden 2 adet ve  $I_5$  ürününden de 2 adet satılmıştır.

2. Minimum destek sayısının 2 olarak belirlendiği varsayılırsa, 1 ögeli sık geçen nesnelere (sık geçen 1-nesnekümeler)  $L_1$  kümesinde gösterilmektedir.  $C_1$  kümesindeki bütün ürünlerin destek sayısı, minimum destek eşik değeri olan 2'den fazla olduğundan  $C_1$ 'deki bütün nesnelere sık geçen nesnekümeler olarak değerlendirilir ve  $L_1$  kümesine aktarılır.

3. Sık geçen 2-nesnekümelerin kümesi olan  $L_2$ 'yi belirlemek için algoritma  $L_1 \times L_1$  kartezyen çarpımını kullanarak yeni bir aday küme oluşturur.  $C_2$  olarak adlandırılan bu aday küme,  $L_1$  kümesindeki ürünlerin ikili kombinasyonlarından oluşur.

4.  $C_2$  kümesindeki ürünlerin destek sayılarını bulmak amacıyla D taranır ve bulunan değerler destek sayısı sütununda belirtilir.

5.  $C_2$  kümesindeki ürünlerden minimum destek eşik değerini aşan ürünler  $L_2$  kümesine aktarılır.

6. Sık geçen 3-nesnekümelerin kümesi olan  $L_3$ 'ü belirlemek için algoritma  $L_2 \times L_2$  kartezyen çarpımını kullanarak yeni bir aday küme olan  $C_3$  kümesini, yani

$$C_3 = L_2 \times L_2 = \{\{I_1, I_2, I_3\}, \{I_1, I_2, I_5\}, \{I_1, I_3, I_5\}, \{I_2, I_3, I_4\}, \{I_2, I_3, I_5\}, \{I_2, I_4, I_5\}\}$$

oluşturur. Apriori özelliği temel alınarak, algoritma çarpım sonucunda ortaya çıkan son dört adayın sık geçen olmayacağını belirler. Bu son dört aday sık geçen olmadıklarından  $C_3$ 'den çıkartılır. Böylece çıkarılan bu adayların sayıları,  $L_3$ 'ün testi için gerekli tarama ve sonraki taramalar boyunca gereksiz yere sayılmayarak fazladan işlem yükü ortadan kaldırılmış olur. Burada dikkat edilmesi gereken bir başka nokta ise şudur;  $C_3$  aday kümesine ait altkümeler incelenirken sadece 2 ögeli altkümelere bakılır. Yani k ögeli (k-nesneküme) aday kümelerinin altkümeleri incelenirken, o aday kümeyle ait (k-1) ögeli altkümelere bakılır. O halde, k ögeli bir aday

küme verildiğinde, bu kümenin (k-1) ögeli altkümelerine bakmak yeterli olacaktır. Sonuç olarak yeni  $C_3$  kümesi  $C_3 = \{I_1, I_2, I_3\}, \{I_1, I_2, I_5\}$  olur.

7.  $C_3$  kümesindeki ürünlerin destek sayılarını bulmak amacıyla D taranır ve bulunan değerler destek sayısı sütununda belirtilir.

8.  $C_3$  kümesindeki ürünlerden minimum destek esik değerini asan ürünler  $L_3$  kümesine aktarılır.

9. Algoritma  $L_3 \times L_3$  kartezyen çarpımını  $C_4$  aday kümesini belirlemek üzere kullanır.  $C_4 = \{I_1, I_2, I_3, I_5\}$  olarak belirlenir. Ancak bu nesnekümenin alt kümesi olan  $\{I_2, I_3, I_5\}$  sık geçen nesneküme olmadığı için bu nesneküme budanır ve  $C_4$  kümesi boş küme olur. Apriori bütün sık geçen nesnelere bulunduğu algoritma sonlanmış olur.

6. adımdaki işlemler:

1. Birleştirme:

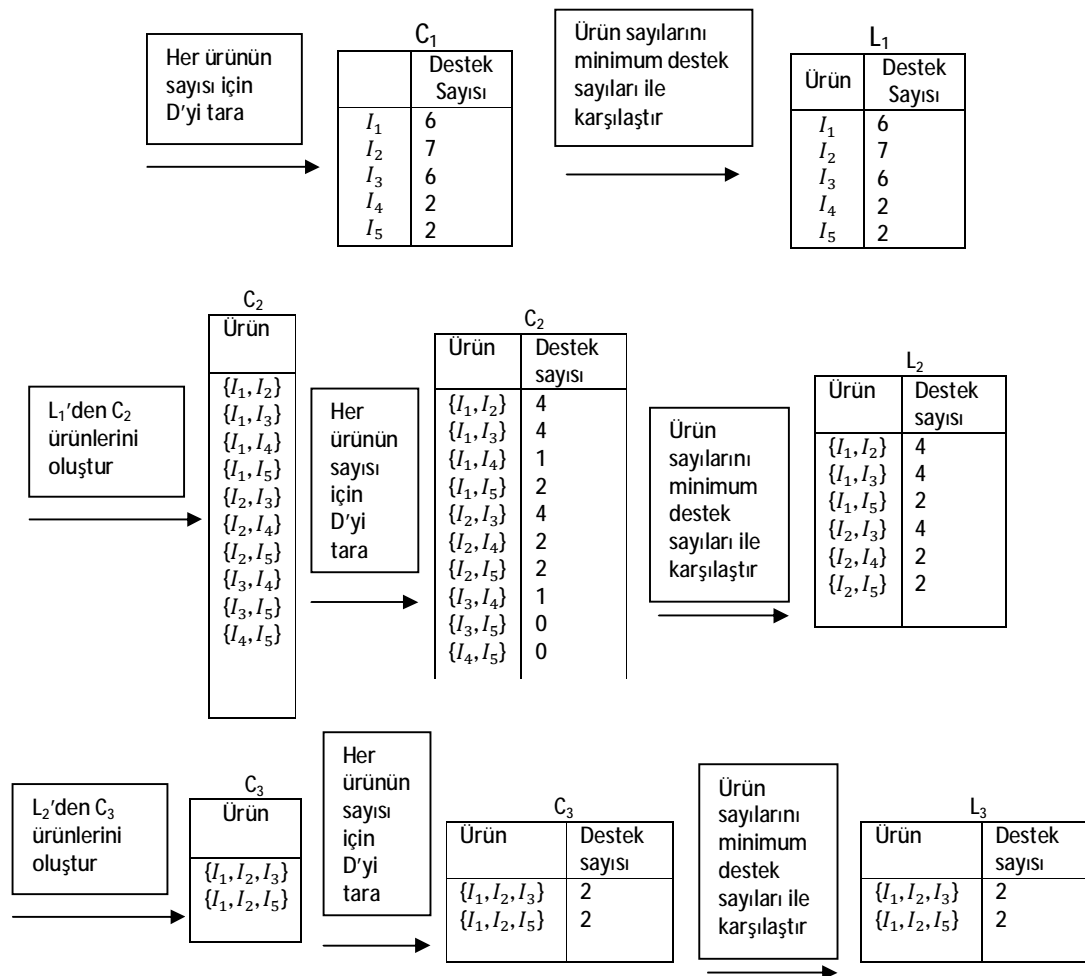
$$C_3 = L_2 \times L_2 = \{I_1, I_2, I_3\}, \{I_1, I_2, I_5\}, \{I_1, I_3, I_5\}, \{I_2, I_3, I_4\}, \{I_2, I_3, I_5\}, \{I_2, I_4, I_5\}$$

2. Apriori özelliğini kullanarak budama: Sık geçen nesnekümenin boş küme olmayan bütün altkümeleri de sık geçendir.

- $\{I_1, I_2, I_3\}$ 'ün 2 ögeli altkümeleri;  $\{I_1, I_2\}, \{I_1, I_3\}, \{I_2, I_3\}$ 'dür. Bu altkümelerin hepsi  $L_2$ 'nin ögeleri olduklarından yani sık geçen nesneküme olduklarından  $\{I_1, I_2, I_3\}$  nesnekümesi  $C_3$ 'te yer alır.
- $\{I_1, I_2, I_5\}$ 'in 2 ögeli altkümeleri;  $\{I_1, I_2\}, \{I_1, I_5\}, \{I_2, I_5\}$ 'dir. Bu altkümelerin hepsi  $L_2$ 'nin ögeleri olduklarından yani sık geçen nesneküme olduklarından  $\{I_1, I_2, I_5\}$  nesnekümesi  $C_3$ 'te yer alır.
- $\{I_1, I_3, I_5\}$ 'in 2 ögeli altkümeleri;  $\{I_1, I_3\}, \{I_1, I_5\}, \{I_3, I_5\}$ 'dir.  $\{I_3, I_5\}$   $L_2$ 'nin ögesi değildir, yani sık geçen değildir. Bundan dolayı,  $\{I_1, I_3, I_5\}$  nesnekümesi  $C_3$ 'ten çıkartılır.
- $\{I_2, I_3, I_4\}$ 'in 2 ögeli altkümeleri;  $\{I_2, I_3\}, \{I_2, I_4\}, \{I_3, I_4\}$ 'dir.  $\{I_3, I_4\}$   $L_2$ 'nin ögesi değildir, yani sık geçen değildir. Bundan dolayı,  $\{I_2, I_3, I_4\}$  nesnekümesi  $C_3$ 'ten çıkartılır.

- $\{I_2, I_3, I_5\}$ 'in 2 ögeli altkümeleri;  $\{I_2, I_3\}$ ,  $\{I_2, I_5\}$ ,  $\{I_3, I_5\}$ 'dir.  $\{I_3, I_5\}$   $L_2$ 'nin ögesi değildir, yani sık geçen değildir. Bundan dolayı,  $\{I_2, I_3, I_5\}$  nesnekümesi  $C_3$ 'ten çıkartılır.
- $\{I_2, I_4, I_5\}$ 'in 2 ögeli altkümeleri;  $\{I_2, I_4\}$ ,  $\{I_2, I_5\}$ ,  $\{I_4, I_5\}$ 'dir.  $\{I_4, I_5\}$   $L_2$ 'nin ögesi değildir, yani sık geçen değildir. Bundan dolayı,  $\{I_2, I_4, I_5\}$  nesnekümesi  $C_3$ 'ten çıkartılır.

3. Bundan dolayı, budamadan sonra  $C_3 = \{\{I_1, I_2, I_3\}, \{I_1, I_2, I_5\}\}$  olur.



Şekil 12. Apriori Algoritmasının Adımları

Sık geçen nesneküme bulunduktan sonra, sıra birliktelik kurallarını oluşturmaya gelir. Örneğin sık geçen bir nesne olan  $I = \{I_1, I_2, I_5\}$  için boş olmayan

bütün alt kümeler şunlardır:  $\{I_1, I_2\}$ ,  $\{I_1, I_5\}$ ,  $\{I_2, I_5\}$ ,  $\{I_1\}$ ,  $\{I_2\}$ ,  $\{I_5\}$ . Bu durumda şu birliktelik kuralları çıkarılabilir;

$$1. I_1 \& I_2 \Rightarrow I_5, \text{güven} = \frac{2}{4} = \%50$$

$$2. I_1 \& I_5 \Rightarrow I_2, \text{güven} = \frac{2}{2} = \%100$$

$$3. I_2 \& I_5 \Rightarrow I_1, \text{güven} = \frac{2}{2} = \%100$$

$$4. I_1 \Rightarrow I_2 \& I_5, \text{güven} = \frac{2}{6} = \%33$$

$$5. I_2 \Rightarrow I_1 \& I_5, \text{güven} = \frac{2}{7} = \%29$$

$$6. I_5 \Rightarrow I_1 \& I_2, \text{güven} = \frac{2}{2} = \%100$$

Eğer minimum güven eşik değeri %70 olarak belirlenmişse sadece; ikinci, üçüncü ve sonuncu kurallar dikkate alınır. Çünkü diğer kurallar eşik değerini aşamamıştır (Han ve Kamber, 2001: 240).

#### 1.1.4. AprioriTid Algoritması

Birliktelik kurallarında, algoritmalar desteği hesaplamak için tüm veritabanını tarar; ancak her aşamada veritabanının tamamının taranmasına gerek olmayabilir. Bu yaklaşımla Agrawal ve Srikant (1994: 4-5), Apriori algoritmasıyla birlikte AprioriTid algoritmasını da sunmuştur.

AprioriTid algoritması da taramadan önce aday nesne kümelerini belirlemek için apriori-gen fonksiyonunu kullanır. Apriori'den en büyük farkı ilk geçişten sonra veritabanının destek seviyesini bulmak için taranmamasıdır. Bu iş için  $C_k$  kullanılır. SETM algoritmasında olduğu gibi  $C_k$  nin her elemanı  $\langle TID, \{X_k\} \rangle$  formundadır. Burada  $X_k$ , TID numaralı işlemde bulunan potansiyel geniş k-nesne kümesidir,  $k=1$  iken  $C_1$  veritabanına karşılık gelir. Bununla beraber her nesne, nesne kümesiyle yer değiştirir,  $k>1$  olduğu durumlarda  $C_k$  algoritmanın onuncu adımında olduğu gibi üretilir,  $t$  işlemindeki  $C_k$  bir elemanı  $\langle TID, c \rangle$  şeklindedir. Burada  $c$ ,  $t$  işlemindeki  $C_k$  ya ait bir aday elemanıdır,  $\{c \in C_k | c\}$ . Eğer bir işlemin, herhangi bir  $k$  nesne

kümesi adayı yoksa, bu durumda  $C_k$ 'nın bu işlem için herhangi bir girdisi, elemanı olmayacaktır. Daha doğrusu bu işlemin TID numarasını taşıyor olacaktır. Böylece  $C_k$  daki girdi sayısı, özellikle bu  $k$  değerleri için, veritabanındaki işlem sayısından daha küçük olabilir. Bunun dışında yine büyük  $k$  değerleri için her girdi kendisine karşılık gelen işlemden daha küçük olabilir. Çünkü o işlemde çok az sayıda aday barınıyor olabilir. Ancak, küçük  $k$  değerleri için bunun tersi olacaktır; yani girdiler kendilerine karşılık gelen işlemlerden daha büyük olabileceklerdir (Agrawal ve Srikant, 1994:5).

### 1.1.5. Diğer Algoritmalar

Bu konuda bugüne kadar birçok algoritma geliştirilmiştir. Bazı algoritmalar eş zamanlı olarak birbirinden bağımsız ve habersiz birden fazla grup tarafından da geliştirilmiştir. Literatürdeki diğer birliktelik kuralları algoritmaları önceki algoritmalara benzer mantık yürütmektedirler. Aşağıda bu algoritmaların kısa bir özeti verilmiştir (Silahtaroglu, 2008: 90-91):

- Bu algoritmalarından bir tanesi Apriori ve AprioriTid algoritmalarının bir karışımı olan Apriori-Hybrid algoritmasıdır.
- Geniş nesne kümelerini belirlemek için veritabanından alınmış küçük örneklerin çok iyi sonuçlar verebileceği fikrine dayanan OCD (Off-line Candidate Determination - Sıradışı Aday Belirleme) algoritması (Mannila, 1994).
- Veritabanını küçük parçalara bölerek, bellekte işgal edilen yeri azaltıp daha hızlı sonuca ulaşma sağlayan bölümlenme (partitioning) (Toivonen, 1996) tekniği.
- 1996'da Toivonen tarafından ortaya atılan ve veritabanındaki tarama sayısını azaltan örnekleme (Savasere, 1995) tekniği.
- Kullanıcıya her taramadan sonra oluşan kuralları gösterip, minimum destek ve güven seviyelerini değiştirme olanağı veren CARMA (Continuous Association Rule Mining Algorithm - Sürekli Bağlantı Kuralı Madenciliği) (Hidber, 1999)

- Veri paralelliğine dayanan CD (Count Distribution - Sayım Dağılımı) (Agrawal ve Shafer, 1996).
- PDM (Parallel Data Mining - Paralel Veri Madenciliği) (Park, 1995), DMA (Distributed Mining Algorithm - Dağıtılmış Madencilik Algoritması) (Cheung, 1996).
- CCPD (Zaki, 1996) (Common Candidate Partitioned Database - Ortak Aday Bölünmüş Veritabanı).
- Görev paralelliğine dayanan DD (Data Distribution - Veri Dağılımı) (Agrawal ve Shafer, 1995).
- IDD (Intelligent Data Distribution) (Han, 1997).
- HPA (Hash-based Parallel Mining of Association Rules - Bağlantı Kurallarının Çırpı Temelli Paralel Madenciliği) (Shintani ve Kitsuregawa, 1996)
- PAR (Parallel Association Rules - Paralel Bağlantı Kuralları) (Zaki, 1997)

## 1.2. BİRLİKTELİK KURALLARI HESAPLAMA YÖNTEMLERİ

Piatetsky-Shapiro (1991: 229)'de ilişkişel tablolardan elde edilen kuralların  $C_1 \rightarrow C_2$  biçiminde olduğunu belirtmişlerdir. Bu kurallar kesin olabilir, yani  $C_1$ 'in olduğu her durumda  $C_2$  de vardır veya %100 doğru olmayabilir, yani  $C_1$ 'in olduğu bazı yerlerde  $C_2$  vardır. Piatetsky, eğer "*destek* ( $X \rightarrow Y$ )  $\approx$  *destek* ( $X$ )  $\times$  *destek* ( $Y$ )" değilse,  $X \rightarrow Y$  kuralının ilginç olmadığını belirtmektedir. Bu tartışma birliktelik kurallarının madenciliğinde önemli bir eleştiri olarak ele alınmıştır.

Agrawal vd. (1993) birliktelik kurallarının belirlenmesinde destek-güven çerçevesini ortaya atmışlardır. Bu çerçeve birliktelik kuralının belirsizliğini, destek ve güven değerlerini hesaplayarak ölçer. Ama bu ölçüm tüm birliktelik kurallarının belirsizliğini ölçmek için yeterli değildir. Bu çerçeveyi geliştirmek için, destek ve güven değerleri üzerinde çeşitli hesaplamaların yapıldığı ki-kare testi modeli (Brin

vd., 1997) ve toplanmış güç tabanlı ölçüm (Agrawal ve Yu, 1998) gibi teknikler geliştirilmiştir.

### 1.2.1. Piatetsky-Shapiro

Bu ölçüme göre eğer “*destek* ( $X \rightarrow Y$ )  $\approx$  *destek* ( $X$ )  $\times$  *destek* ( $Y$ )” değilse,  $X \rightarrow Y$  kuralı ilginç değildir.

$$\text{Destek}(X \cup Y) = p(X \cup Y) \text{ ve Güven}(X \rightarrow Y) = p(Y|X) = p(X \cup Y)/p(X)$$

kurallarına göre Piatetsky-Shapiro ölçümü aşağıdaki şekilde ifade edilebilir.

$$p(X \cup Y) \approx p(X)p(Y)$$

Bu ölçüme göre eğer  $p(X \cup Y) \approx p(X)p(Y)$  değilse,  $X \rightarrow Y$  kuralı ilginç bir kural olmayacaktır. Aynı zamanda olasılık teorisine göre,  $p(X \cup Y) \approx p(X)p(Y)$  olması  $X$ 'in hemen hemen  $Y$ 'den bağımsız olduğunu göstermektedir.

### 1.2.2. Brin-Motwani-Silverstein

Brin, Motwani ve Silverstein (1997)  $X$  ve  $Y$  kümelerinin istatistiksel olarak bağımlılık tanımlamasını aşağıdaki şekilde yapmıştır.

$$\text{İlgi}(X, Y) = \frac{p(X \cup Y)}{p(X)p(Y)}$$

Verilen  $X$ 'e  $Y$ 'nin önemini gösteren bu formül, birliktelik kurallarının değerlendirilmesinde kullanılan ana hesaplamalardan biridir. Bu durumda bulunan sayı 1'den uzaklaştıkça bağımlılığın arttığı söylenebilir.

$X \rightarrow Y$  nin ilginç bir kural olabilmesi için ( $1 > \text{minilgi} > 0$  için);

$$\left| \frac{p(X \cup Y)}{p(X)p(Y)} - 1 \right| \geq \text{minilgi}$$

olmalıdır.

Piatetsky-Shapiro hesaplamasına göre  $\text{İlgi}(X, Y)$  üç farklı durumda değerlendirilebilir:

1. Eğer  $p(X \cup Y)/(p(X)p(Y)) = 1$  ise,  $p(X \cup Y) = p(X)p(Y)$ 'dir veya X ve Y birbirinden bağımsızdır.

2. Eğer  $p(X \cup Y)/(p(X)p(Y)) > 1$  veya  $p(X \cup Y) > p(X)p(Y)$  ise Y pozitif olarak X'e bağımlıdır.

3. Eğer  $p(X \cup Y)/(p(X)p(Y)) < 1$  veya  $p(X \cup Y) < p(X)p(Y)$  ise Y negatif olarak X'e bağımlıdır veya  $\neg Y$  pozitif olarak X'e bağımlıdır.

Bu bilgilere göre, kurallara ilginçlik durumu için şöyle bir yorum getirebilir: (1>minilgi>0 için)

(a) Eğer

$$\frac{p(X \cup Y)}{p(X)p(Y)} - 1 \geq \text{minilgi}$$

ise  $X \Rightarrow Y$  ilginç bir kuraldır.

(b) Eğer

$$-\left(\frac{p(X \cup Y)}{p(X)p(Y)} - 1\right) \geq \text{minilgi}$$

ise  $X \Rightarrow Y$  ilginç bir kural değildir.

Brin vd.(1997)'nin ileri sürdüğü ilgi değeri literatürde kaldırma oranı (lift) olarak da kullanılmaktadır (Han ve Kamber, 2001: 261). Bramer'a (2007: 216) göre  $L \rightarrow R$  bağıntısındaki kaldırma oranı, L ve R nesnekümleri istatistiksel olarak bağımsız varsayılırsa, işlemlerde L ve R nesnelerinin bir arada görülmesinin, beklenen değerden kaç kat fazla olduğudur.

Güven-destek çerçevesi ve Brin, Motwani ve Silverstein'in hesaplamaları birlikte ele alındığında,  $X \Rightarrow Y$ 'nin ilginç bir kural olabilmesi için aşağıdaki şartları taşıması gerekmektedir.

$$(1) X \cap Y = \mathbf{0}$$

$$(2) p(X \cup Y) \geq \text{mindestek}$$

$$(3) p(Y|X) \geq \text{mingüven yani } (\text{conf}(X \rightarrow Y) \geq \text{mingüven})$$

$$(4) \left| \frac{p(X \cup Y)}{p(X)p(Y)} - 1 \right| \geq \text{minilgi}$$



Burada mindestek (minimum destek), mingüven (minimum güven) ve minilgi (minimum ilgi değeri) kullanıcı veya konu uzmanları tarafından belirlenmelidir (Zhang ve Zhang, 2002: 33).

### 1.2.3. Ki-Kare Analizi İle Birliktelik Kurallarının Değerlendirilmesi

Olasılık teorisi ve istatistiksel yöntemler çoğu uygulamada belirsizliği ölçmek için kullanılan en eski ve yaygın yöntemlerdir. Bu yöntemler birliktelik kurallarındaki belirsizliği tahmin etmek için de kullanılabilir. Brin, Motwani ve Silverstein (1997: 3-5) klasik istatistikteki ki-kare analiziyle, birliktelik kurallarının anlamlılığını ölçmüşlerdir.

Bir marketten yapılmış  $n$  adet alışverişe ait bilgimiz olduğu varsayalım. Çay ve kahvenin satışlarını inceleyelim. Tablo 8'de  $t$  satırı çay alanları,  $\rightarrow t$  satırı çay almayanları, aynı şekilde  $c$  sütunu kahve alanları,  $\rightarrow c$  sütunu da kahve almayanları göstermektedir. Değerler % olarak verilmiştir.

Tablo 8. Çay-Kahve Satışları Tablosu

	Kahve	Kahve almayanlar	Toplam
Çay	20	5	25
Çay almayanlar	70	5	75
Toplam	90	10	100

Potansiyel  $t \rightarrow c$  kuralı için bilinen güven-destek çerçevesi uygulanabilir. Bu kural için destek %20 gibi yüksek bir rakamdır. Çay alan bir müşterinin kahve de alma olasılığı yani güven değeri de  $p(t \cap c)/p(t) = 20/25 = 0,8$  yani %80 gibi yine yüksek bir rakamdır. Bu noktada  $t \Rightarrow c$  kuralının geçerli bir kural olduğunu söyleyebiliriz.

Ama herhangi bir müşterinin kahve alma olasılığının %90 olduğuna dair bilgiyi dikkate almamız gerekir. Başka bir deyişle çay aldığı bilinen bir müşterinin kahve alma olasılığı %80 iken, markete yeni giren; daha hiç bir şey almamış

müşterinin kahve alma olasılığı %90'dır. Tabii ki bu kadar çok müşterinin çayın yanında kahve de aldığını bilmek önemlidir ama tek başına bu kural eksik bilgi vermektedir ve kullanıcıyı yanlış yönlendirebilir. Doğrusu çay almak ve kahve almak arasında negatif bir korelasyon da vardır; en azından  $t \rightarrow c$  kuralından bahsederken bu bilgi de verilmelidir. Bu korelasyonun ölçülmesinde aşağıdaki yöntem kullanılabilir.

$$\frac{p(t \cap c)}{p(t)p(c)} = \frac{0,2}{0,25 \times 0,9} = 0,89$$

Bu değer 1'den anlamlı derecede düşük olması negatif korelasyonu gösterir. Burada pay kahve ve çayı aynı zamanda alan kişilerin görülme oranını, payda ise bu iki satınalmanın tamamen bağımsız olma derecesini verir.

Burada çay ve kahve arasında bir korelasyon verilmiştir ama bu korelasyonun anlamlı olup olmadığı belirlenememiştir. Anlamlılığını ölçmek için uygulanması kolay olan ve bir iki varsayım kabul edildiğinde güvenilir olan ki-kare testi kullanılabilir. Bu testle hem korelasyon ölçülebilir, hem de gizli kalmış negatif anlamlar tespit edilebilir.

$p(A)$ , A olayının olma olasılığı;  $p(\neg A) = 1 - p(A)$  A olayının olmama olasılığı olarak kabul edilsin. Aynı şekilde  $p(A \cap B)$ , A ve B olaylarının aynı anda olma olasılığı;  $p(\neg A \cap B)$  B'nin olması ama A'nın olmaması olasılığı olarak kabul edilsin. Bu kabullere göre aşağıdaki tanımlamaları yapabiliriz.

- Eğer  $p(A \cap B) = p(A)p(B)$  ise A ve B olayları bağımsızdır.
- Eğer iki olay bağımsız değilse, bağımlıdır yani korelasyon içerir.

Eğer n denemelik bir seri varsa, A olayının olma sayısı  $O_n(A)$  veya  $O(A)$  olarak gösterilir.  $p(A)$ ,  $O_n(A)/n$  şeklinde hesaplanabilir. Bu şekilde  $p(A \cap B) \neq p(A)p(B)$  olup olmadığı belirlenebilir.

Bu tanımları birliktelik kurallarına yerleştirmek istersek;  $I$ 'yi nesnelerin kümesi,  $B$ 'yi de  $I$ 'nin alt kümesi kabul edelim. Eğer  $i_{a1}, \dots, i_{am}$  nesnelerinin oluşları korelasyonlu ise,  $\{i_{a1}, \dots, i_{am}\} \subset I$  bir korelasyondur.  $R, \{i_1, \neg i_1\} \times \dots \times \{i_k, \neg i_k\}$ ,  $r = r_1 \dots r_k \in R$  olsun. Burada R olabilecek tüm sepet değerlerinin kümesi, r de tek bir sepetin değeridir. r'nin her değeri bir hücreyi belirtir. Bu terminoloji, olumsuzluk

tablosu (contingency table) diye bilinen, R kümesini k-boyutlu bir tablo halinde gösterilmesinden gelir.  $O(r)$  r hücresine düşen sepet sayısını gösterebilir. Bir hücrenin bağımlı olup olmadığını test etmek için, r hücresindeki sayının beklenen değerden yeterince farklı olup olmadığını belirlemek gerekir.

Ki-Kare testinde beklenen değer, bağımsızlık varsayımına göre hesaplanır. Tek bir nesne için  $E[i_j] = O_n(i_j)$  olduğundan  $E[\neg i_j] = n - O_n(i_j)$  ve  $E[r] = n \times E[r_1]/n \times \dots \times E[r_k]/n$  'dir. Buna göre ki-kare istatistiği aşağıdaki şekilde hesaplanır:

$$\chi^2 = \sum_{r \in R} (O(r) - E[r])^2 / E[r]$$

Ki-kare testi tüm k nesnelere k-şekilde bağımsız olup olmadığını belirleyecektir. Eğer değişkenler gerçekten bağımsızsa, hesaplanan ki-kare değeri 0 olacaktır. Eğer kesim frekansından yüksekse (%95 anlamlılık düzeyi için 3,84), bağımsızlık varsayımı reddedilir.

Tablo 9'da  $i_2$  askerlik görevini yapmayanların sayısını,  $i_7$  ise 40 yaşından küçük olanların sayısını göstermektedir.

*Tablo 9. Ki-kare Analizi İçin Örnek Veri Dağılım Tablosu*

	$i_2$	$\neg i_2$	Toplam
$i_7$	17918	911	18829
$\neg i_7$	9111	2430	11541
<b>Toplam</b>	27029	3341	30370

Yukarıdaki teoremler Tablo 10 üzerinde şu şekilde açıklanabilir. Bu tabloda satır ve sütundaki hücreler için beklenen değerler toplandığında ki-kare değeri olarak 1922 değeri bulunur. Bu değer %95 anlamlılık derecesinde anlamlıdır. Ki-kare değeri hesaplanırken en yüksek katkı sağ alttaki hücreden gelmektedir. Bu da en yüksek bağımlılığın yaşı 40'dan büyük olan ve askerliğini yapmış olanlar arasında olduğunu gösterir.

Olumsuzluk tablolarında r hücresinin bağımlılığı hücrenin ilgisi (interest) olarak tanımlanır ve  $I(r) = O(r)/E[r]$  formülüyle hesaplanır.  $I(r)$  değeri 1'den

uzaklaştıkça  $r$  hücresindeki nesnelerin bağımlılığı artar.  $1$ 'in üzerindeki değerler pozitif bağımlılığı,  $1$ 'in altındaki değerler negatif bağımlılığı gösterir.

Başka bir deyişle  $1$ 'in üzerindeki ilgi değerleri  $X$  ve  $Y$  arasında tamamlayıcı etkiyi,  $1$ 'in altındaki ilgi değerleri  $X$  ve  $Y$  arasında ikame edilebilir (yerine geçici) etkiyi gösterir. İlgi değerinin  $1$  olması veya  $1$ 'e çok yakın olması  $X$  ve  $Y$  arasında bağımsızlığa işaret eder (Brijs ve ark., 1999: 5).

Yukarıdaki örnek için  $I(r)$  değerleri hesaplanırsa Tablo 10 elde edilir.

*Tablo 10.  $I(r)$  Değerleri Tablosu*

	$i_2$	$\rightarrow i_2$
$i_7$	1,07	0,44
$\rightarrow i_7$	0,89	1,91

Sağ alttaki hücre en büyük ilgi değerini göstermektedir. Yani 40 yaşından büyük olmak ve askerliğini yapmış olmak arasında pozitif bir bağımlılık vardır. Diğer hücrelerdeki değerler de anlamlıdır; örneğin 40 yaşından küçük olmak ve askerliğini yapmış olmak arasında büyük bir ters yönlü ilişki (0,44 ) vardır. Daha önce hesaplanan ki-kare değeri anlamlı çıktığından bu ilişkilerin istatistiksel olarak anlamlı olduğu söylenebilir.

Ki-kare istatistiksel analizi yukarıdaki şekilde kullanılarak destek-güven çerçevesine göre bulunan birliktelik kurallarının istatistiksel olarak da anlamlı olup olmadığı belirlenebilir.

### 1.3 . PAZAR SEPET ANALİZİ

Birliktelik kuralları analizi, pazarlama alanında, özellikle perakende sektöründe alışveriş sepetlerindeki ürünlerin birlikte satış ilişkisinin bulunması için kullanıldığında genellikle pazar sepet analizi (market basket analysis) olarak adlandırılmaktadır.

Geçmiş tarihli hareketleri çözümlmek, karar destek sistemlerinde karar verme aşamasında verilen kararın kalitesini artırmak için izlenen bir yaklaşımdır. 1990'lı yılların başına kadar teknik yetersizlikten dolayı, kurumlara veya müşterilere satış yapıldığı ana ait değil, belirli bir zaman aralığında (günlük, haftalık, aylık,

yıllık) gerçekleşen satış hareketlerinin tamamına ilişkin genel veriler elektronik ortamda tutulmaktaydı. Barkot uygulamalarındaki gelişme, bir harekete ait verilerin satış hareketi olduğu anda toplanmasına ve elektronik ortama aktarılmasına olanak sağlamıştır. Genellikle süper marketlerin satış noktalarında bu tür veriler toplandığından, toplanan bu veriye “pazar sepeti verisi” adı verilmiştir. Pazar sepeti verisinde yer alan bir kayıta, tekil olan hareket numarası, hareket tarihi ve satın alınan ürünlere ilişkin ürün kodu, miktarı, fiyatı gibi bilgiler yer alır (Agrawal ve diğ., 1993: 1; Han ve Kamber, 2001: 228).

Günümüzde değişen alışveriş alışkanlıkları sonucu alışverişler, bakkallardan süpermarketlere doğru bir kayma gerçekleştirmiştir. İnsanlar ihtiyaç duydukları bir ürünü bakkaldan almak yerine, alışverişlerini süpermarketlerin çeşitli avantajlarından (bol çeşit, uygun fiyat, değişik alışveriş ve eğlence olanaklarının bir arada olması, vb.) dolayı toplu olarak süpermarketlerden gerçekleştirmektedir. Süpermarketlere ulaşımın genelde bakkallar kadar kolay olmaması sebebiyle de, alışverişlerin genelde toplu olarak yapıldığı görülmektedir.

Her müşteri değişik ürünleri, değişik zamanlarda ve değişik sayılarda alabilir. Herhangi bir müşterinin alışveriş sepetinde birden fazla ürün de yer alabilir. Tüm müşterilerin alışveriş sepetleri incelendiğinde işletmeye daha önemli bilgiler verebilir. Hangi müşteri, ne zaman, hangi ürünlerden, kaç tane aldı gibi... Pazar sepeti analizi (Market Basket Analysis) müşterilerin hangi ürünleri aldığı bilgisini kullanarak, müşterilerin kim olduğu ve belli bazı alışverişleri niye yaptığını açıklayabilir. Pazar sepeti analizi ile hangi ürünlerin birlikte satıldığı ve hangi ürünlerin promosyona en uygun olduğu belirlenebilir. Bu sayede yeni mağaza yerleşimleri planlanabilir, hangi ürünlerde kampanya yapılabileceği belirlenebilir. Bu bilgiler müşteri kartları yardımıyla bireysel olarak müşterilerle eşleştirilebilirse daha değerli olabilir (Berry & Linoff, 2004: 289).

Alpaydın (2000: 9) sepet analizini amaç alanlar arasındaki ilişkileri bulmak olarak tanımlamaktadır. Bu ilişkilerin bilinmesi şirketin kârını arttırmak için kullanılabilir. Eğer X malını alanların Y malını da çok yüksek olasılıkla aldıklarını biliyorsanız ve eğer bir müşteri X malını alıyor ama Y malını almıyorsa o potansiyel bir Y müşterisidir. Örneğin internet üzerinden kitap satan Amazon şirketi

(<http://www.amazon.com>) BookMatcher adlı programıyla müşterilerine okudukları ve sevdikleri kitaplara göre satın almaları için kitap tavsiye etmektedir.

Pazar sepeti analizi ile perakende sektöründe kullanılabilecek aşağıdaki bilgiler elde edilir (Svetina ve Zupancic, 2005: 5).

- Müşteri davranışı hakkında daha ayrıntılı bilgiler elde edilir.
- Raf yerleşimi (tanzim ve teşhir), fiyatlandırma, promosyon ve kârlılık konularında daha verimli, daha bilgiye dayalı kararlar verilebilir.
- Birbirine benzerlik gösteren ürünler belirlenebilir.
- Hangi ürünlerin birbirinin yakınına yerleştirileceği belirlenebilir.
- Hangi ürünlerin çapraz satışı yapılabileceği belirlenebilir.
- Belirli bir etkileyiciye sahip olmayan başarılı ürünler belirlenebilir.

Sepet çözümlemesinin başarılı olduğu noktalar şunlardır (Tantuğ, 2002: 38):

- Kolay ve anlaşılır sonuçlar üretir,
- Değişik boyutlardaki veriler üzerinde çalışabilir,
- Her ne kadar kayıtların sayısı ve kombinasyon seçimine göre işlem adedi artsa da sepet çözümlemesi için her adımda gerekli olan hesaplamalar diğer yöntemlere göre (genetik algoritmalar, yapay sinir ağları vb.) çok daha basittir

Sepet çözümlemesinin başarısız olduğu noktalar da şunlardır (Tantuğ, 2002: 39; Rushing, 1997: 5):

- Sorunun boyutu büyüdükçe, gerekli hesaplamalar üstel olarak artmaktadır,
- Kayıtlarda çok az rastlanan ürünler analiz sonuçlarına yansımaz. Sepet çözümlemesi yönteminin en doğru sonucu, tüm ürünlerin kayıtlar içinde yaklaşık aynı frekansta görüldüğü durumlarda üretmektedir.

- Destek ve güven eşik değerleri üretilen kural sayısında sınırlama getirirler fakat eşik değerlerinin çok düşük belirlendiği durumda kullanıcı gerçekten ilgilendiği kuralları kaybetme tehlikesi ile karşı karşıya kalır.

Pazar sepeti uygulamalarında birliktelik kuralları kullanılmaktadır. Bu işlem, müşterilerin yaptıkları alışverişlerdeki ürünler arasındaki birliktelikleri bularak müşterilerin satın alma alışkanlıklarını çözümler. Bu tip birlikteliklerin keşfedilmesi, müşterilerin hangi ürünleri bir arada aldıkları bilgisini ortaya çıkarır ve market yöneticileri de bu bilgi ışığında raf düzenlerini belirleyerek satış oranlarını artırabilir ve etkili satış stratejileri geliştirebilirler. Pazar sepeti çözümlemesinin son zamanlarda çok büyük ilgi ile karşılanmasının sebebi kullanım kolaylığı ve anlaşılabilirliğidir.

#### 1.4. ÖRNEK BİR PAZAR SEPETİ ANALİZİ

Horizon şirketi internet üzerinden elektronik-bilgisayar ürünleri satan bir e-ticaret firmasıdır. Şirket yeni stoğa giren LCD Flat Panel HD televizyonların satışını nasıl arttırabileceğini araştırmıştır. Bu aşamada yeni müşteri edinmektense, eski müşterilere bu ürünün çapraz satışının daha az maliyetli ve daha karlı olacağını düşünmüşlerdir (Chiu ve Tavella, 2008: 236-240).

İşletme temel müşteri bilgilerini (müşteri numarası, adı, adresi) ve alışveriş detaylarını (alışveriş zamanı, alınan ürünler, satış fiyatları) içeren 2 yıllık bir veritabanına sahiptir. Analizi daha hızlı hale getirmek için bu iki yıllık veriden, rastgele olarak 69671 adedi (%5 örneklem) seçilmiştir. Analiz sırasında sadece iki değişken (müşteri numarası ve alınan ürünler) kullanılmıştır. Tablo 11’de kullanılan veriden örnek bir bölüm görülmektedir.

*Tablo 11. Bir E-Ticaret Sitesinin Alışveriş Kayıt Örneği*

Müşteri No	Alınan Ürünler
10001	HP 710C Inkjet Printer
10001	IBM Thinkpad Laptop
10002	Toshiba laptop
10003	Apollo p1200 Color Inkjet
10003	LCD Flat Panel HDTV
10004	Camcorder with HD Video
10004	LCD Flat Panel HDTV
10004	VCR
10004	Sony VAIO laptop

Verilerin yarısı (34929 adet) birliktelik modeli kurulması için kullanılmıştır. Yeniden değerlendirme ve minimum görülme sayısına ulaşamayan ürünlerin elenmesinden sonra 27264 adet hareket, 18172 adet müşteri ve 148 çeşit ürün analizde incelenmiştir. Analiz sonucunda 20 adet birliktelik kuralı oluşturulmuştur.

*Tablo 12. Analiz Sonucunda Elde Edilen Birliktelik Kuralları*

KURAL			Destek %	Güven %	Kaldırma Oranı
StarOffice 5.0 Personal Edition for Linux (Intel)	⇒	Red Hat Linux 6.0 for Intel Systems	0.53	49.11	20.39
VCR	⇒	LCD Flat Panel HDTV	0.17	43.37	92.08
Seagate 6.5 GB EIDE Hard Drives	⇒	Hard Drive Cable Pack	0.11	43.14	104.3
Sony VAIO laptop	⇒	LCD Flat Panel HDTV	0.12	42.11	89.39
Zobmondo!! Lite boardgame	⇒	Zobmondo!! Original boardgame	0.12	40.32	88.31

Tablo 12 incelendiğinde LCD Flat Panel HD televizyonların satışını artırmak için, içerisinde bu ürünün adı geçen 2. ve 4. kuralların kullanılabileceği görülecektir. 2. kural VCR satın alanların LCD Flat Panel HD televizyon da aldığını göstermektedir. Destek değeri (%0,17) iki ürünü de alanların oranını, güven değeri (%43,4) VCR satın alanların ne kadarının LCD Flat Panel HD televizyon aldığını göstermektedir. Kaldırma oranı değeri de VCR alan bir müşterinin, rastgele seçilen başka bir müşteriye göre 92,08 (ilişkinin güven değeri/ LCD alanların yüzdesi = %43,37/%0,47) kat daha fazla olasılıkla LCD Flat Panel HD televizyon alacağını



göstermektedir. 4. kural da Sony Vaio dizüstü bilgisayarların, sık sık LCD Flat Panel HD televizyon alan müşteriler tarafından alındığını göstermektedir. Müşterilerin %0,12'si bu iki ürünü beraber satın almıştır. Sony Vaio alanlar 89,39 kat daha fazla olasılıkla LCD Flat Panel HD televizyon almaktadır. Sony Vaio alanların %42,11'i LCD Flat Panel HD televizyon almaktadır.

Verilerin diğer yarısı da aynı şekilde birliktelik kuralları analizine tabi tutulmuştur. Yukarıda belirtilen kurallar bu analizde 3. ve 5. sıralarda elde edilmiştir. “VCR⇒ LCD Flat Panel HD televizyon” kuralı için destek, güven ve kaldırma oranı değerleri sırasıyla 0,17; 39,33 ve 91,78'dir. “Sony Vaio ⇒ LCD Flat Panel HD televizyon” kuralı için destek, güven ve kaldırma oranı değerleri sırasıyla 0,11; 34,92 ve 81,50'dir. Doğrulama sonucunda elde edilen kurallar model kurallarıyla paralel olduğundan, şirket yetkililerinin LCD Flat Panel HD televizyonun çapraz satışı için VCR veya Sony Vaio alan müşteri gruplarına göre uygulamalar yapmaları satışları arttıracaktır.

## 2. KÜMELEME ANALİZİ

Kümeleme, kayıtları, alanları, kişileri, nesnelere, gözlemleri, benzer eğilim ve özelliklerine göre gruplamak amacıyla kullanılan bir veri madenciliği yöntemidir. Özdamar'a (1999: 257) göre kümeleme analizi X veri matrisinde yer alan ve doğal gruplamaları kesin olarak bilinmeyen birimleri, değişkenleri ya da birim ve değişkenleri birbirleri ile benzer olan alt kümelere ayırmaya yardımcı olan yöntemdir.

Tatlıdil (2002: 329) kümeleme analizini, gruplanmamış verileri benzerliklerine göre sınıflandıran ve araştırmacıya uygun, işe yarar ve özetleyici bilgiler elde etmede yardımcı olan bir yöntem olarak tanımlamıştır. Koç (2001: 2) kümeleme analizinde temel amacın, dağınık bir halde bulunan verileri benzerliklerine göre bir araya getirip sınıflandırarak işlenebilir hale getirmek olduğunu belirtmiştir.

Altıntaş'a (2006:18) göre kümeleme analizi, nesnelere alt dizinlere gruplanmasını yapan bir işlemdir. Böylece nesnelere, örneklene kitle özelliklerini iyi

yansıtan etkili bir temsil gücüne sahip olmuş olur. Sınıflamanın aksine, yeniden tanımlanmış sınıflara dayalı değildir. Kümeleme, bir denetimsiz öğrenme (unsupervised learning) yöntemidir.

Kümeleme analizi; birimleri ve nesnelere düzenleyerek sınıflara ayıran çok değişkenli istatistik analizlerinden birisidir. Bu nesnelere kümelenmesindeki amaç, birimlerin kendi içlerinde birbirine benzeyecek şekilde gruplanmasını yapmaktır. Kümeleme sonucunda nesnelere küme içinde yüksek homojenlik gösterirken, kümeler arasında yüksek heterojenlik göstereceklerdir (Hair, Black, Barry, 2006: 559).. Kümeleme işlemi başarılı ise geometrik gösterimde küme içinde yer alan nesnelere birbirine yakın iken, farklı kümelerde yer alan nesnelere birbirinden uzakta yer alacaklardır (Selanik, 2007: 3).

Özdamar (1999: 257-258) kümeleme analizini temelde aşağıda belirtilen dört değişik amaca yönelik olarak uygulanan bir yöntem olarak belirtmektedir:

- n sayıda birimi, nesneyi, oluşumu; p sayıda değişkene göre saptanan özelliklerine göre olabildiğince kendi içinde homojen ve kendi aralarında heterojen alt gruplara (kümelere) ayırmak.
- p sayıda değişkeni, n sayıda birimde saptanan değerlere göre ortak özellikleri açıkladığı varsayılan alt kümelere ayırmak ve ortak faktör yapıları ortaya koymak.
- Hem birimleri hem de değişkenleri birlikte ele alarak n birimi p değişkene göre ortak özellikli alt kümelere ayırmak.
- Birimleri, p sayıda değişken baz alınarak saptanan değerlere göre, izledikleri biyolojik ve tipolojik sınıflamayı ortaya koymak (taksonomik sınıflandırma yapmak).

Çok değişkenli, çok sayıda gözlemi tanımlamak zor olabilir. Bu tür bir veri setinde istenilen bilgiyi çıkartmak ve özetlemek için, gözlemler gruplara ayrılmak zorundadır. Belirlenen ölçülerde gözlemler gruplara ayrıldığında, gruplar içinde homojenlik ve gruplar arasında heterojenlik sağlanmalıdır. Bu durum araştırmacıya hem zaman, hem de hesaplama yönünde kolaylık sağlar. Bu işlem bir miktar bilgi kaybına neden olacağından, kümeleme analizi bu bilgi kaybını asgariye indirger. Kümeleme analizinin, tahminden çok tanımlama ile, değişkenlerden çok bireyler ile,

kriter-tahmin deęişkenleri ilişkilerinden çok tüm veri setindeki ilişkiler ile ilgilendięi, bu analizde deęişkenler arasındaki ilişkinin doğrusal olmasının şart olmadığı ve analizin seçilen kümeleme yöntemine göre nominal, ordinal, aralıklı veya kombine bir ölçekle ölçülmüş verilere uygulanabileceęi söylenebilir (Selanik, 2007:5).

Doęan (2008:78) kümeleme analizinin işletme biliminde kullanımının pazarlama araştırmaları alanında yoğunlaştığını ve stratejik grupları tanımlamak, belirli bir mal veya hizmetin hitap ettięi tüketici gruplarını sınıflandırmak, bölgeler arası tüketim benzerliklerini/farklılıklarını belirlemek ve ürünleri tüketici karakteristiklerine göre sınıflandırmak gibi amaçlarla kullanıldığını görüldüğünü belirtmektedir.

## 2.1. UZAKLIK ÖLÇÜLERİ

Kümeleme analizinin ana teması, gözlenen birey veya nesnelere arasındaki benzerlikleri veya uzaklıkları/yakınlıkları tespit etmektir. Benzerlik, uzaklık kavramının tersi olup, büyük bir sayı çıktığında iki nesnenin birbirine yakın olduğunu, küçük bir sayı çıktığında ise iki nesnenin birbirinden uzak olduğunu göstermektedir (Nakip, 2003: 420). Kümeleme analizinde birimlerin p deęişkene göre birbirleri arasındaki uzaklıkları hesaplamak için çok çeşitli uzaklık ölçü birimleri ileri sürülmüştür (Özdamar, 1999: 261).

p sayıda deęişkene göre birimler ya da deęişkenler arasındaki uzaklıkları hesaplamak için kullanılan uzaklık ölçülerine genel olarak Minkowski uzaklık ölçüsü adı verilir (Tatlıdil, 2002: 332). Minkowski uzaklık ölçüsü aşağıdaki gibi hesaplanır:

$$d(x, y) = \left[ \sum_{i=1}^p |x_i - y_i|^m \right]^{\frac{1}{m}}$$

Uzaklık ölçüleri ya da benzerlik ölçüleri veri matrisinde yer alan deęişkenlerin ölçü birimlerine göre farklılık göstermektedir. Eđer deęişkenler oransal yada aralıklı ölçekle elde edilmiş deęerleri ise uzaklık ya da ilişki türü ölçülerden yararlanır. Ölçümler sayısal deęerler olarak yapılmış ise tercih edilen ölçüler ki-kare uzaklık ölçüsü ya da phi kare uzaklık ölçüsüdür. Eđer ikili (binary) gözlemlere

göre ölçümler yapılmış ise birimler arasındaki benzerlikleri belirlemede öklid, kare öklid, size difference, Lance and Willams difference gibi benzerlik ya da farklılık ölçülerinden yararlanılmaktadır.

Minkowski uzaklığının  $m=2$  değeri için özel hesaplanma formu Öklid uzaklığı ya da karesel öklid uzaklığı olarak bilinmektedir.  $n \times p$  boyutlu bir veri matrisinde  $i$ . ve  $j$ . birimler (gözlemler, nesnelere) arasındaki öklid uzaklığı aşağıdaki şekilde hesaplanır.

$$d(i, j) = \sqrt{\sum_{k=1}^p (X_{ik} - X_{jk})^2}$$

Öklid uzaklığının birimler ve değişkenler arasındaki uzaklıkları ve benzerlikleri hesaplamada kullanılan tutarlı bir ölçü olduğu kabul edilmektedir (Özdamar, 1999: 269).

## 2.2. KÜMELEME YÖNTEMLERİ

Kümeleme yöntemleri; uzaklık matrisi ya da benzerlik matrisinden yararlanarak birimler ya da değişkenleri kendi içinde homojen ve kendi aralarında heterojen gruplamalar oluşturmayı sağlayan yöntemlerdir. Kümeleme yöntemleri, grupları belirlemede izledikleri yaklaşımlara göre hiyerarşik kümeleme ve hiyerarşik olmayan kümeleme olmak üzere ikiye ayrılabilir.

### 2.2.1. Hiyerarşik Kümeleme

Bağlantı yöntemleri olarak da adlandırılan hiyerarşik kümeleme yöntemleri, birimleri birbirleri ile değişik aşamalarda bir araya getirerek; ardışık biçimde kümeler oluşturmaya ve bu kümelere girecek elemanların hangi uzaklık ya da benzerlik düzeyinde küme elemanı olduğunu belirlemeye yönelik yöntemlerdir (Doğan 2008: 88).

Hiyerarşik kümelemede, her kümede veri nesnelere içerecek bir bağlantı kurulur. Hangi yöntem olursa olsun kümeler birbirine benzer özellik gösteren nesnelere oluşturulur. Böylece kümeler kendi içinde aynı özelliği taşıyan nesnelere içermiş olur. Hiyerarşik kümeleme nesnelere yakınlık ilişkisine göre oluşturulan

kümelerden bir ağaç inşa eder. Hiyerarşik kümelemede kümeleme işlemi, temel olarak henüz aynı kümede olmayan iki en benzer değişkeni ve onların kümelerini belirlemektir. Bu kural, kümeleme analizinin başlangıcında her değişkenin kendisi tek başına bir küme olmak üzere bütün değişkenler tek bir küme oluncaya kadar her iki kümenin birleşmesi şeklinde tekrarlanmaktadır (Şimşek, 2006: 120).

Altıntaş (2006:26) hiyerarşik kümelemenin aşağıdaki özelliklere sahip olduğunu belirtmektedir:

- Bir veri tabanını bir kaç kümeye ayırıştırır.
- Bu ayırıştırma dendogram adı verilen bir ağaç sayesinde yapılır.
- Bu ağaç, yapraklardan gövdeye doğru veya gövdeden yapraklara doğru kurulabilir.

Selanik (2007:10) hiyerarşik kümeleme teknikleri uygulanırken, kaç küme oluşacağını önceden bilinmediğini belirtmektedir. Öyle ki, kümeleme sürecinin başlangıcında her birey bir kümedir, süreç sonunda ise tüm bireyler bir kümede toplanır.

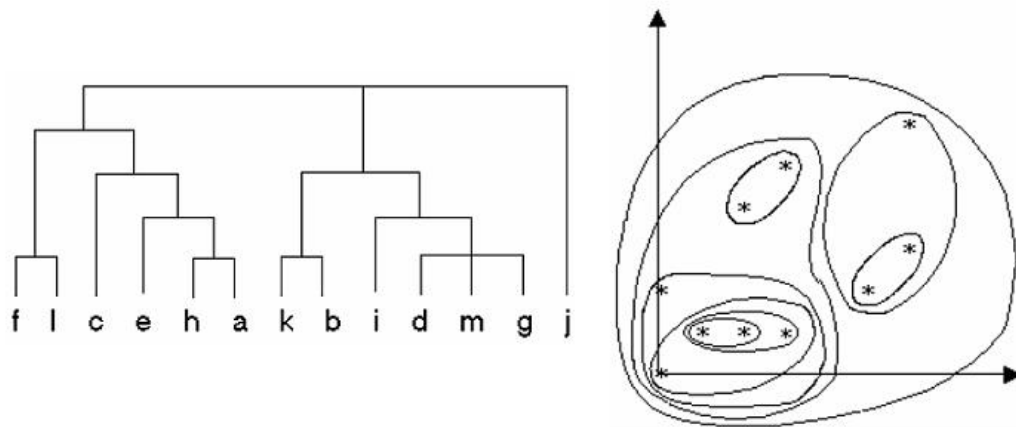
Doğan'ın (2008:89) Gordon'dan (1987) aktardığına göre, hiyerarşik toplama yönteminde kümelerin grafiksel olarak ifade edilebilmesi için ağaç diyagram (dendogram) örneğinden yararlanılır. Bu diyagramda hiyerarşik bir yapı gözlenmektedir. Kümeleme sürecinin başında her gözlem bir kümedir (Ağacın dalları). Süreç sonunda ise tüm gözlemler bir kümede toplanır. En düşük düzeyde ele alınan gözlemlerin tümü bağımsız iken, sonraki aşamalarda gözlemler birleşerek yeni kümeleri oluşturmaktadır.

Kümeleme sürecinin başlangıcında her birey bir kümedir. Süreç sonunda ise tüm bireyler bir kümede toplanır. Bu sürecin işleyişi aşağıda maddeler halinde sıralanmıştır (Tatlıdil, 1996: 334):

1. Öncelikle  $n$  adet birey,  $n$  adet küme olmak üzere işleme başlanır.
2. En yakın iki küme ( $d_{ij}$  değeri en küçük olan) birleştirilir.
3. Küme sayısı bir indirgenerek yinelenmiş uzaklıklar matrisi bulunur.
4. 2 ve 3 numaralı adımlar  $n-1$  kez tekrarlanır.

Hiyerarşik kümeleme yöntemleri de kendi içinde birleştirici ve ayrıştırıcı yöntemler olmak üzere ikiye ayrılmıştır. Birleştirici aşamalı kümeleme yöntemleri başlangıçta tüm birimlerin ayrı birer küme oluşturduğunu kabul ederek,  $n$  adet birimi sırasıyla  $n, n-1, n-2, \dots, n-r, \dots, 3, 2, 1$  kümeye yerleştirmeyi amaçlayan bir yaklaşımdır. Ayrıcı aşamalı kümeleme yöntemleri başlangıçta tüm birimlerin bir küme oluşturduğunu kabul ederek, birimleri aşamalı olarak sırasıyla  $1, 2, 3, \dots, n-r, \dots, n-2, n-1, n$  kümeye yerleştirmeyi amaçlayan bir yaklaşımdır (Özdamar, 1999: 272).

Birleştirici aşamalı kümeleme yöntemleri, birimlerin birbirleri ile hangi benzerlik düzeyinde ortak özelliklere sahip kümeler oluşturduklarını göstermeleri açısından yaygın olarak kullanılan aşamalı kümeleme yaklaşımıdır. Bu yöntemde, her birim başlangıçta tek başına birer küme olarak kabul edilir. Daha sonra birbirleri ile yüksek derecede benzerlik gösteren iki birim bir küme oluşturur. Daha sonra bu kümeye değişik benzerlik düzeylerinde diğer birimler eklenerek, birimlerini tümü bir kümede toplayacak biçimde birbirleri ile bağlanırlar. Bağlantılar, uzaklıkları ve birimlerin bağlanma düzeyleri dendogram adı verilen ağaç grafikleri ile gösterilirler. Şekil 13'de birleştirici aşamalı kümeleme yöntemine göre çizilmiş bir dendogram ve buna paralel olarak oluşturulmuş bir küme grafiği görülmektedir.



Şekil 13. Örnek bir dendogram ve küme grafiği

Birleştirici kümeleme algoritmalarında Merkezi Kümeleme (Centroid) Yöntemi, Tek Bağlantı (En Yakın Komşu- Single Linkage) Yöntemi, Tam Bağlantı

(En Uzak Komşu - Complete Linkage) Yöntemi, Ortalama Bağlantı (Average Linkage) Yöntemi ve Ward Yöntemi gibi yaklaşımlar uygulanmaktadır. Ayırıcı kümeleme algoritmalarında is Bölünmüş Ortalamalar (Splinter- Average Distance) Yöntemi, Otomatik Etkileşim Dedektörü (Automatic Interaction Detection-AID) Yöntemi gibi yaklaşımlar uygulanmaktadır. (Şimşek, 2006: 119).

### 2.2.2. Hiyerarşik Olmayan Kümeleme

Küme sayısı konusunda ön bilgi var ise ya da araştırmacı anlamlı olacak küme sayısına karar vermiş ise bu durumda, çok uzun zaman alan hiyerarşik yöntemler yerine, hiyerarşik olmayan kümeleme yöntemleri tercih edilmektedir. Ayrıca, bu yöntemlerin kurumsal dayanaklarının daha güçlü olması diğer bir tercih sebebidir (Selanik, 2007: 13). Hiyerarşik olmayan kümeleme yöntemleri prototip aracılığı ile alt populasyonların parametre tahminlerini yapmayı amaçlayan yöntemlerdir (Doğan, 2008:103).

Özdamar'a (1999: 303) göre hiyerarşik olmayan yöntemlerde birimlerin uygun oldukları kümelerde toplanmaları ve n birimin k sayıda kümeye parçalanması hedeflenmektedir. Birimlerin ayrılacakları küme sayısı belirlendikten sonra, kümeler için belirlenen küme belirleme kriterlerine göre birimlerin hangi kümelere girebileceklerine karar verilir ve atama işlemi gerçekleştirilir.

Hiyerarşik olmayan teknikler düğüm yöntemleri olarak da adlandırılır. Veri seti önceden belirlenmiş sayıda kümeye ayrılır. Bu kümelerin merkezleri yani düğüm noktaları hesaplanır. Bu adımlar her gözlem bir kümeye atanana kadar devam eder. Hiyerarşik kümelemede bir gözlem bir kümeye atandıktan sonra tekrar yer değiştirmez. Hiyerarşik olmayan kümeleme de ise küme sayısı başlangıçta belirlendiği için, gözlemler, kümeleme algoritması sonlanana kadar ilk atandıkları kümeden farklı bir kümeye de atanabilmektedirler (Doğan, 2008: 103).

Hiyerarşik olmayan kümelemede veri k adet gruba ayrılır ve her bir grup bir kümeyi belirtir. Dolayısıyla hiyerarşik kümelemenin tersine, küme sayısı önceden bilinmektedir. Hiyerarşik olmayan kümeleme analizinde aşağıdaki aşamalar izlenir (Sharma, 1996: 202):

1. k adet küme için k adet küme ortalaması analist tarafından belirlenir.
2. Gözlem birimleri hangi küme ortalamasına daha yakınsa, o kümeye dahil edilir.
3. Kümelere ait ortalama değerleri yeniden hesaplanır.
4. Küme elemanlarında herhangi bir değişiklik yoksa işlem durdurulur. Eğer değişiklik varsa ikinci adıma dönülür.

Hiyerarşik olmayan kümeleme yöntemleri arasında k-ortalamalar (k-means clustering) yöntemi, metoid kümeleme (metoid clustering), bulanık kümeleme (fuzzy clustering) ve yığma (hill climbing) kümeleme gibi pek çok yöntem bulunmaktadır. Ancak bunlardan en çok kullanılanı k-ortalamalar tekniğidir (Şimşek, 2006: 134).



## **DÖRDÜNCÜ BÖLÜM**

### **BİR SÜPERMARKETİN SATIŞ FİŞLERİ ÜZERİNDE ÜRÜNLERİN SATIŞINDA BİRLİKTELİK KURALLARI VE KÜMELEME ANALİZLERİ UYGULAMASI**

Bu bölümde öncelikle araştırmanın amaç, önem ve kapsamından bahsedilecek, daha sonra da önceki bölümlerde bahsedilen teorik konuların ve daha önce yapılmış çalışmaların ışığında bir araştırma modeli oluşturulacak ve bu modelin uygulanması sonucunda elde edilen bulgular belirtilecektir.

#### **1. ARAŞTIRMANIN AMACI, ÖNEMİ VE KAPSAMI**

##### **1.1. AMAÇ**

Bu araştırmanın amacı bir süpermarketin satış terminalleri aracılığıyla topladığı satışlara ait verileri (alışveriş detayları: fiş no, fiş tarihi, saat, alınan ürünler, vb.) veri madenciliği yöntemleri (birliktelik kuralları: GRI ve Apriori, hiyerarşik kümeleme analizi) ile incelemek ve kategoriler arası ilişkileri ortaya çıkarmaktır. Uygulanacak veri madenciliği yöntemleri ile birliktelik kuralları analizi yapılması, böylece ürün türleri hakkında birlikte satışla ilgili kuralların bulunması ve bu kurallar ışığında kategori yönetimi, tanzim-teşhir konularında çeşitli bilgilere ulaşılması amaçlanmaktadır.

##### **1.2. ÖNEM**

Bu çalışma ile işletme tarafından depolanan verilerden anlamlı bilgiler elde edilmesi ve bu bilgiler ışığında satış faaliyetlerinin düzenlenmesi amaçlanmaktadır. İşletme uzun bir süredir sürekli olarak müşteri verilerini depolamakta ama bu verilerden tam anlamıyla yararlanamamaktadır. Veritabanının içinde saklı olan birçok değerli bilginin keşfedilmesi ile işletme daha etkin satış yapabilecektir.

### 1.3. KAPSAM

Araştırmanın kapsamı Afyonkarahisar ilinde faaliyet gösteren bir süpermarketin 01.01.2007-31.12.2007 tarihleri arasındaki 1 yıllık market satış bilgileriyle sınırlıdır.

## 2. MODELİN OLUŞTURMASI VE UYGULANMASI

Bu çalışmada veri madenciliği uygulaması CRISP-DM referans modeli takip edilerek gerçekleştirilmiştir. CRISP-DM modeli ile ilgili ayrıntılı bilgi 2. Bölümde “Veri Madenciliği Süreci” başlığı altında verilmiştir. Bu bölümde CRISP-DM referans modelinin bu araştırmaya uygulanması yani problemin değerlendirilmesi ve amacın belirlenmesi, verinin incelenmesi, verinin hazırlanması ve model oluşturma aşamaları anlatılacaktır.

İşletmeden elde edilen verilerde Apriori ve GRI algoritmaları kullanılarak Pazar sepeti analizleri gerçekleştirilecektir. Aynı zamanda ürün sınıfları arasında birlikte satış sırasında oluşan kümeleri belirlemek amacıyla da hiyerarşik kümeleme analizleri uygulanacaktır.

Araştırma sırasında, veri madenciliği uygulamasında yol gösterici bir süreç olan ve çoğu veri madenciliği uygulamasında standart olarak kullanılan CRISP-DM modeli uygulanacaktır. Bu modele göre araştırmanın basamakları aşağıda belirtilmiştir.

### 2. 1. ARAŞTIRMANIN OLUŞTURULMASI

Verilerin alındığı işletme Afyonkarahisar’da faaliyet gösteren bir süpermarkettir. İşletmenin kolay ulaşılabilir bir yerde olması, fiyatlarının uygunluğu, ürün çeşitliliği, ürün kalitesi, kampanyaları, promosyonları sayesinde tüketicinin beğenisini kazanmakta ve alışveriş tercihi olarak ön plana çıkmaktadır.

İşletmenin amacı rekabetin hızla arttığı perakendecilik sektöründe müşterilere daha iyi hizmet verebilmek, satışlarını artırmak için politikalar geliştirmektir.

İşletmede çok sayıda ürün çeşidi ve markası bulunmaktadır. Çok miktardaki bu ürün çeşitlerinin satış istatistiklerinden elde edilen bilgilerle kategori yönetimi ve tanzim-teşhir işlemleri gerçekleştirilmektedir. Ama ürün kategorileri arasındaki birlikte satış bağıntıları bilinmediği için, bu faktörleri içeren bir yapılandırma uygulanamamaktadır.

Market yöneticileriyle yapılan görüşme sonucunda, marketin uzun bir süredir müşteri verilerini topladığı ve veritabanına kaydettiği ama bu bilgileri kullanmadığı ve yorumlayamadığı görülmüştür. Bunun üzerine veri madenciliği sistemi ve mağaza yönetimine ve işletmeye olabilecek katkıları anlatılmıştır. Bunun sonucunda da market yöneticileri market ismi saklı kalmak ve sonuçları paylaşmak kaydıyla verilerin kullanılmasına izin vermişlerdir.

## 2.2. VERİLERİN OLUŞTURULMASI

Ön görüşmelerden sonra işletmenin yazılımlarını hazırlayan ve uygulayan bilgi işlem birimiyle görüşülmüş ve veriler ham veri olarak sağlanmıştır.

2007 yılına ait alışveriş detayları sistemden okunarak txt formatında kaydedilmiştir. Alınan alışveriş verileri aylara göre alışveriş detaylarını yani tarih, kasa no, fiş no, barkod no, ürün ismi, adedi, birim fiyatı, KDV'si, toplam fiyatı gibi bilgilerden oluşmaktadır.

## 2.3. VERİLERİN HAZIRLANMASI

Elde edilen veriler bu aşamada modelde kullanılmak üzere hazır hale getirilmiştir. "txt" formatında olan veriler Microsoft Excel programına aktarılmıştır. Veriler bu aşamada işletmenin kullandığı kategori düzenleme sistemine göre gruplandırılmıştır. Kullanılan verilerde toplam 1.096.125 adet fiş içerisinde 6.009.317 alışveriş hareketi olduğu görülmüştür. Aylara göre fiş adedi ve hareket miktarı Tablo 13'de görülmektedir. Ortalama bir fişteki hareket adedi 5,48'dir.

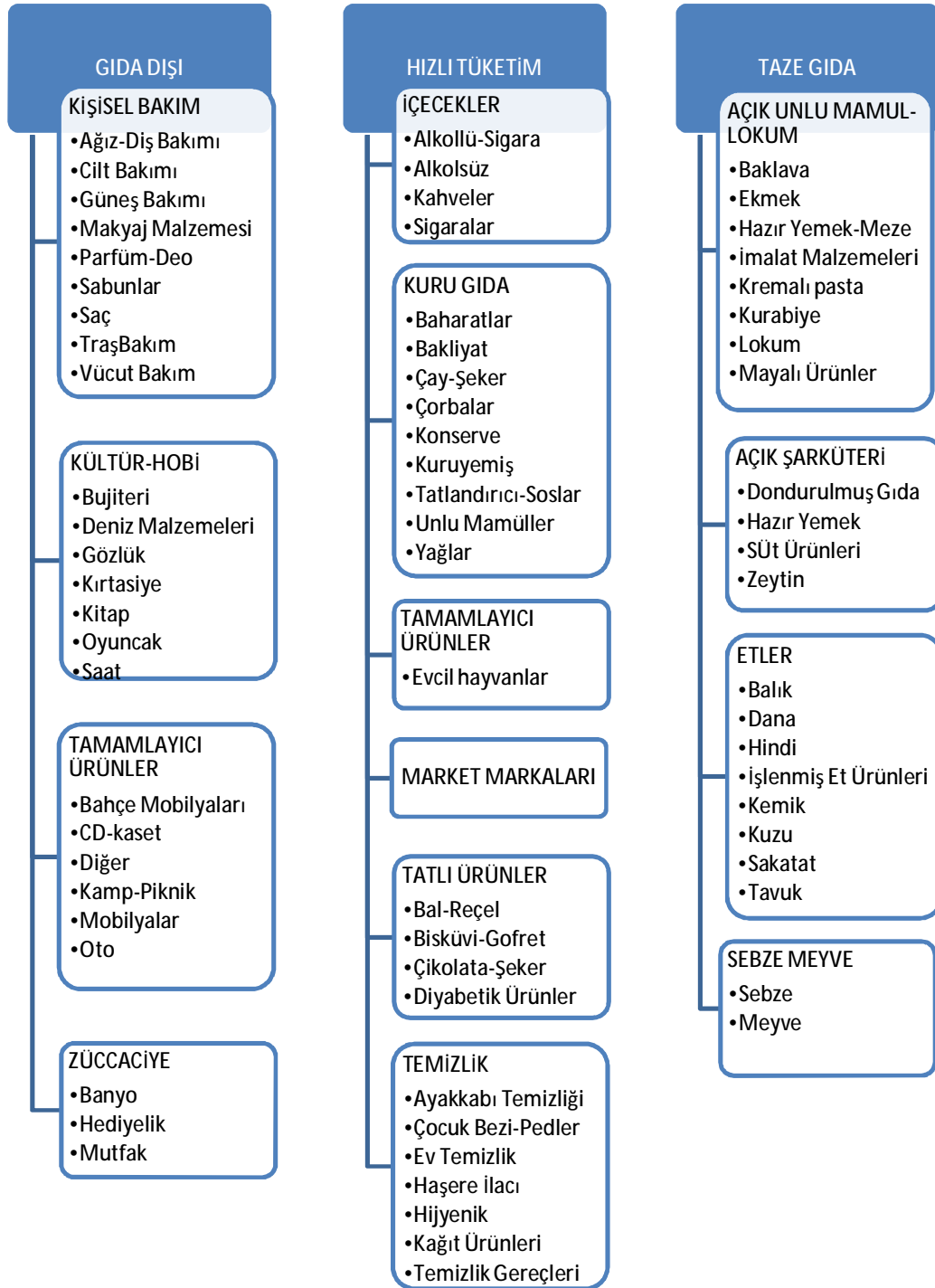
Şekil 14’de bu markette kullanılmakta olan kategori yönetimi çizelgesi görülmektedir. Ürünler ana olarak gıda dışı, hızlı tüketim ve taze gıda olmak üzere 3 kategoriye ayrılmıştır. Bu üç kategorinin altında ise alt kategoriler bulunmaktadır. İşletme her kategoriyle ilgili satış miktarları, karlar vb. gibi değerleri hesaplayabilmekte ama kategoriler arası veya kategori içi birlikte satış konusunda herhangi bir bilgiye ulaşamamaktadır.

*Tablo 13. Aylara Göre Fiş ve Hareket Miktarının Dağılımı*

<b>Aylar</b>	<b>Fiş sayısı</b>	<b>Hareket sayısı</b>
Ocak	75902	415468
Şubat	74368	440818
Mart	78848	438892
Nisan	82480	462999
Mayıs	80882	415776
Haziran	92252	472526
Temmuz	125128	640652
Ağustos	132491	668247
Eylül	94021	564886
Ekim	93593	528222
Kasım	74490	429756
Aralık	91670	531075

Bir fişte en çok görülen kategori sayısı 14, en az görülen kategori sayısı 1’dir. Tablo 14’de fişlerde görülen hareketlerin kategori ve aylara göre dağılımları görülmektedir. Bu tabloya göre en çok hareket “tatlı ürünler” ve “içecekler” kategori türlerinde görülmektedir. Tablo 15’de de fişlerde görülen kategori miktarlarının aylara göre dağılımı görülmektedir. Bu tabloya göre “içecekler” ve “tatlı ürünler” kategorilerin fişlerde en sık rastlanan kategoriler olduğu görülmektedir. En az görülen kategori ise “Evcil hayvanlar” kategorisidir.

Şekil 14. Alışveriş Merkezinde Kullanılan Ürün Kategorileri



Tablo 14. Kategorilerin Hareketlerdeki Görülme Sayılarının Aylara Göre Dağılımı

	Ocak	Şubat	Mart	Nisan	Mayıs	Haziran	Temmuz	Ağustos	Eylül	Ekim	Kasım	Aralık	
Kişisel bakım	22862	25595	24834	24181	22185	25425	44158	39390	31107	31764	24741	27565	343807
Kültür-hobi	9227	8543	9531	7749	7166	9079	13843	21773	34425	11916	8599	12904	154755
Tamamlayıcı ür.	8614	8136	6344	7709	6026	7768	10235	10491	7232	7096	5385	7583	92619
Züccaciye	7884	8217	8860	9949	9777	10596	14926	16334	11811	13754	8687	11244	132039
İçecekler	66872	69197	77727	83022	85108	95826	135187	149698	92522	88843	67268	85517	1096787
Kuru gıda	60789	65879	74269	72187	60283	63604	75466	75613	92739	77086	71360	82258	871533
Market mar.	9122	9052	13000	10142	8785	9711	12565	11504	11640	12113	10337	11160	129131
Evcil ürünler	175	168	240	252	199	253	452	394	372	396	397	278	3576
Tatlı ürünler	86563	94461	92125	95625	78666	89497	115411	123478	91028	107037	89878	112346	1176115
Temizlik	25450	29044	31818	32071	26217	31136	38942	36778	35946	36251	28515	34739	386907
Açık unlu mamul	20983	20404	21882	21856	20175	21603	29509	31008	27763	24473	15600	22324	277580
Açık şarküteri	34445	34270	32699	39482	37377	43332	53692	53545	48014	46667	38355	43137	505015
Etlere	8416	11066	3970	13393	12856	13529	19095	20535	19327	15349	13877	13195	164608
Sebze-meyve	14591	14194	15125	15557	13849	17091	25424	23084	15959	15491	16128	21123	207616

Tablo 15. Kategorilerin Fişlerde Görülme Sayılarının Aylara Göre Dağılımı

	Ocak	Şubat	Mart	Nisan	Mayıs	Haziran	Temmuz	Ağustos	Eylül	Ekim	Kasım	Aralık	
Kişisel bakım	11922	12670	12840	12622	11679	13378	20777	19749	15223	15588	12557	14558	173563
Kültür-hobi	6107	5515	6393	5317	5037	6418	9406	11284	10938	7675	5789	8513	88392
Tamamlayıcı ür.	5663	5421	4457	5237	4038	5212	6699	6903	4908	4783	3742	5159	62222
Züccaciye	5281	5576	6138	6716	6573	7144	10249	11565	8133	9036	5968	7721	90100
İçecekler	34413	34330	40415	42035	44323	49412	69068	75289	47838	45837	36261	44076	563297
Kuru gıda	17731	18519	23067	22770	19696	21231	25939	26121	24897	25430	21336	25773	272510
Market mar.	5359	5378	7180	6032	5095	5799	7658	6973	6789	7189	5881	6339	75672
Evcil ürünler	130	121	122	148	111	149	222	205	200	221	235	179	2043
Tatlı ürünler	28819	29292	30296	30506	26448	30588	41638	44282	31278	35989	28576	35952	393664
Temizlik	11357	12583	14344	14386	11926	14695	19311	18191	15989	16284	12501	16026	177593
Açık unlu mamul	13710	13183	14336	14434	13466	14529	19672	20457	18342	16355	10835	14738	184057
Açık şarküteri	14583	14470	14277	16486	15928	18654	24243	24133	20010	18937	15231	17827	214779
Etlere	5604	6905	3132	8170	7713	8247	11441	12063	11386	9657	8445	8387	101150
Sebze-meyve	6527	6063	6576	6687	6221	8234	12464	11731	7748	7541	7589	9272	96653

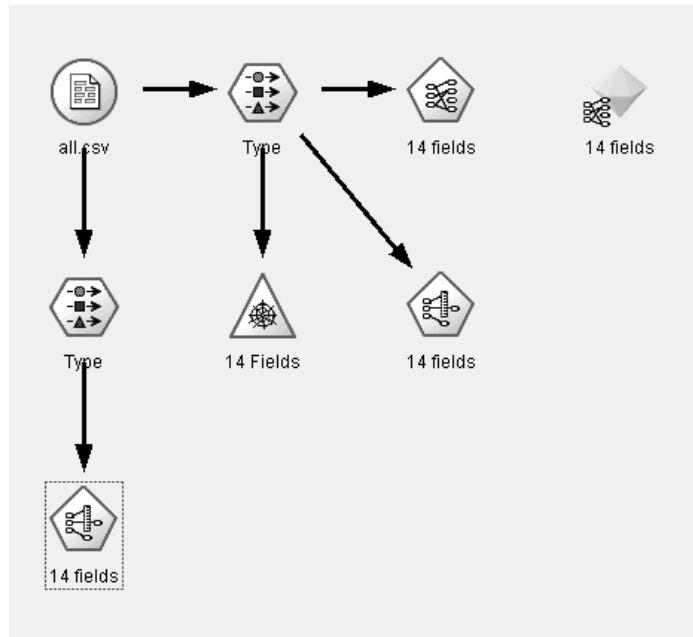
Verilerle ilgili bir sorun da verilerin veri madenciliği yazılımında kullanmaya tam olarak hazır olmamasıdır. Alışveriş detayları Microsoft Excel yazılımında satır satır görülmektedir. Ama veri madenciliği yazılımı açısından analizlerde, tek bir fişteki detayların tek bir satırda gözükmeleri gerekmektedir. Bu sorunun giderilmesi için de Visual Basic uygulaması (Visual Basic Application – VBA) kullanılarak bir kod yazılmıştır. Bu işlem sonucunda bir alışverişe ait fiş no, tarih, müşteri no ve o fişte hangi kategorideki ürünlerin alındığı bilgileri tek bir satırda toplanmıştır.

Veriler gerekli düzenlemelerden sonra “SPSS Clementine” yazılımında kullanılmak üzere “comma delimited-csv” formatına çevrilmiştir.

#### 2.4. MODELLEME

Modelleme aşamasında uygun modeller kurularak analizler gerçekleştirilmiştir. Birliktelik analizinde SPSS Clementine yazılımında bulunan Apriori, Web ve GRI modelleri uygulanmıştır. Kümeleme analizinde ise SPSS 13.0 yazılımında bulunan hiyerarşik kümeleme analizi kullanılmıştır.

Şekil 15. Clementine Programında Kullanılan Model Görüntüsü





Clementine yazılımında veri ilk önce “Sources” araç kutusundaki “Var.File” düğmesi ile yazılıma yüklenir. Daha sonra “Type” düğmesi ile veri tipleri ve veri özellikleri belirlenir. Bu kutuda aynı zamanda analize girecek ve analiz sonuçlarında yer alacak değişkenlerin belirlenmesi de mümkündür. Ürün kategorileri arasındaki satış ilişkilerini grafiksel olarak görebilmek için “Web” düğmesi kullanılmaktadır. “GRI” ve “Apriori” düğmeleri de birliktelik kurallarının tespit edilmesi için kullanılmıştır.

## 2.5. DEĞERLENDİRME

Bu aşamada oluşturulan modelin işletilmesi neticesinde elde edilen sonuçlar incelenecektir. Kurulan modeller kategori türlerinin birbirleri ve kendi grupları arasında birliktelik kuralları analizi sonuçlarını verecektir.

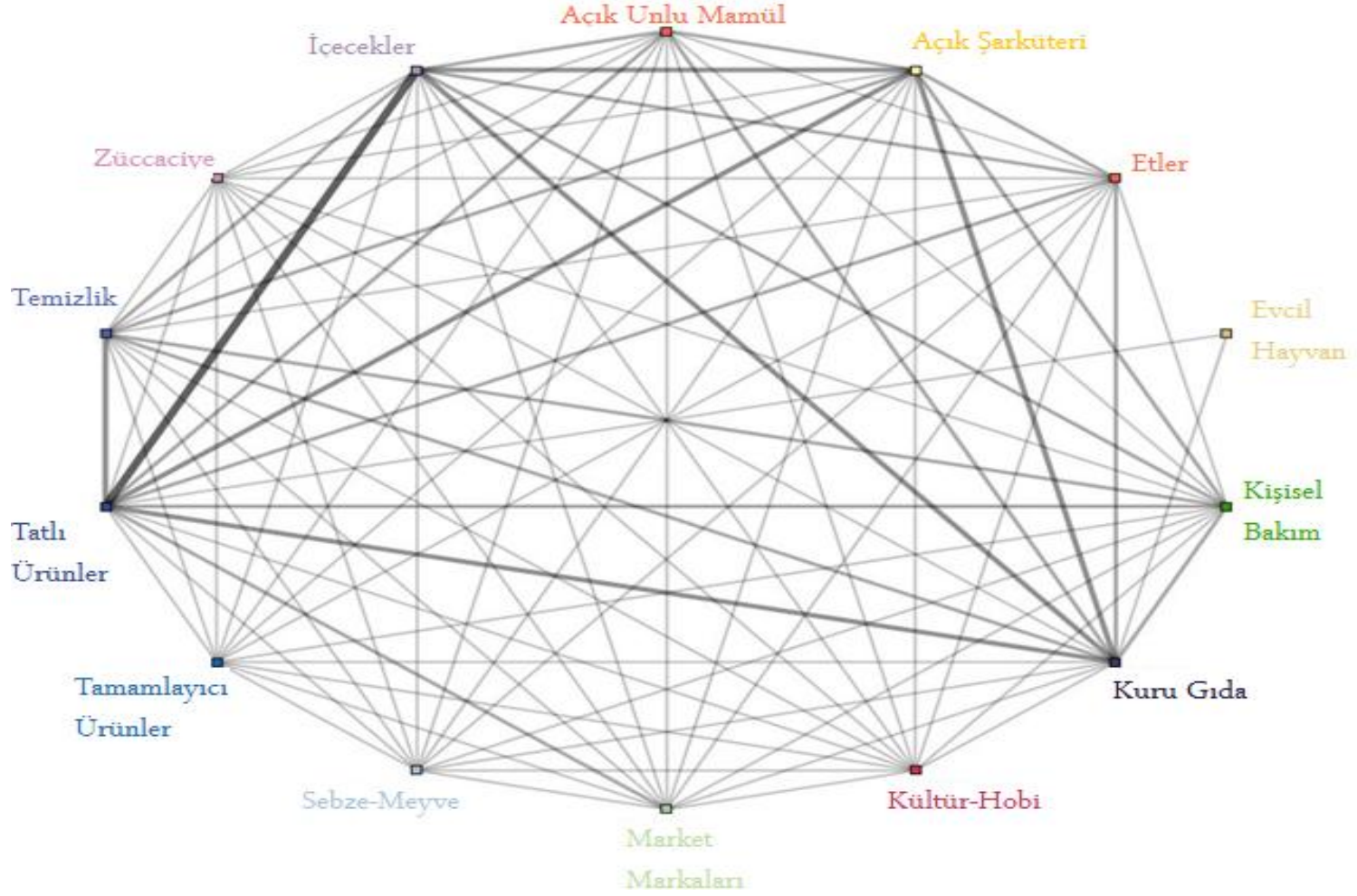
### 2.5.1. Kategoriler Arasındaki Birlikte Satış İlişkilerinin Grafiksel Olarak Gösterilmesi

Kategoriler arasındaki ilişkileri belirlemek amacıyla verilere ilk önce Web Analizi uygulanmıştır. Web Analizi iki veya daha fazla alan arasındaki ilişkinin gücünü göstermek için kullanılır. İki alan arasındaki ilişkinin gücüne göre çizginin şekli değişecektir. Güçlü ilişkiler kalın çizgilerle gösterilir. Kalın çizgi iki alanın kuvvetli bir biçimde ilişkili olduğunu ve incelenmesi gerektiğini gösterir. Orta derecedeki bağlantılar normal çizgiyle gösterilir. Zayıf bağlantılar ise kesikli çizgilerle gösterilir. Eğer iki alan arasında herhangi bir çizgi yoksa ya bu ikiliye aynı kayıta hiç rastlanmamış ya da rastlanma sayısı eşik değerinin altında kalmış demektir.

Şekil 16 incelenirse en kalın çizgilerin dolayısıyla da en kuvvetli bağlantıların “içecekler-tatlı ürünler”, “kuru gıda-tatlı ürünler”, “içecekler-kuru gıda”, “tatlı ürünler-açık şarküteri”, “içecekler-açık şarküteri”, “kuru gıda-açık şarküteri”, “tatlı ürünler-temizlik” kategori çiftleri arasında olduğu görülecektir.

Elde edilen bu grafik sonucunda kategoriler arasındaki ilişkiler görsel bir şekilde belirtilmiştir. Kategoriler arasındaki çizgilerin ince olması aralarındaki birlikte satış ilişkisinin az olduğunu göstermektedir. Ama bu ilişkilerin daha net, daha doğru belirlenebilmesi ve kategoriler arasındaki üçlü, dörtlü satış ilişkilerinin belirlenebilmesi için Apriori ve GRI modellerinden yararlanılması gerekmektedir.

Şekil 16. Kategoriler Arası Birlikte Satış Bağlantılarını Gösteren Web Grafiği



## 2.5.2. Kategoriler Arasındaki İkili Birlikte Satış İlişkilerinin İncelenmesi

Apriori ve GRI modelleri Clementine yazılımında birliktelik kurallarını belirlenmesi amacıyla kullanılmaktadır. Apriori modeli bir veri setinden en yüksek bilgi içeriğine sahip kuralları çıkarır. Bilgi içeriği bulunurken destek-genellik (support-generality) ve güven-doğruluk (confidence-accuracy) değerleri hesaplanır. Genel literatürde destek değeri ürünlerin birlikte bulunma olasılığını, güven değeri ise ilk ürünü içeren kayıtların ne kadarının ikinci ürünü de içereceğini gösterir. Ama SPSS Clementine programında elde edilen Destek (support) değeri öncül (andecent) türün veri seti içerisindeki toplam bulunma yüzdesini, güven (confidence) değeri öncül tür içeren kayıtların ne kadarının artçıl (consequent) türde kayıtlar içerdiğini, kural desteği (rule support) değeri de iki türün birlikte bulunma yüzdesini belirtir. Yani literatürdeki destek kavramı burada kural desteği olarak kullanılmaktadır. Tablolar değerlendirilirken SPSS Clementine programının tanımları dikkate alınmıştır. Analizlerde önce tüm kategori türleri bir arada incelenecek, genel bir görünüm elde edildikten sonra da kategoriler tek başına incelenecektir.

*Tablo 16. Kategorilere Uygulanmış Apriori Modelinin Sonuçları (Kural Desteğine Göre Sıralı)*

Öncül	Artçıl	Destek%	Güven %	Kural Desteği %
Tatlı ürünler	İçecekler	35,91	52,04	18,69
İçecekler	Tatlı ürünler	51,39	36,37	18,69
Kuru gıda	Tatlı ürünler	24,86	52,41	13,03
Tatlı ürünler	Kuru gıda	35,91	36,28	13,03
Kuru gıda	İçecekler	24,86	49,23	12,24
İçecekler	Kuru gıda	51,39	32,28	12,24
Tatlı ürünler	Açık şarküteri	35,91	32,70	11,74
Açık şarküteri	Tatlı ürünler	19,59	59,94	11,74
Açık şarküteri	İçecekler	19,59	53,32	10,44
Açık şarküteri	Kuru gıda	19,59	52,99	10,38
Kuru gıda	Açık şarküteri	24,86	41,76	10,38

Tablo 16 incelendiğinde en büyük kural desteğinin %18,69 ile içecekler⇒tatlı ürünler bağıntısında olduğu görülmektedir. Yani fişlerin yaklaşık %19'unda içecekler ve tatlı ürünler kategorilerindeki ürünler birlikte yer almaktadır. Birinci ve ikinci sıradaki bağıntılar arasındaki fark güven değeridir. Tatlı ürünler kategorisinden bir ürün alan müşterilerin yarısından fazlası (%52,04) içecekler

kategorisinden bir ürün de alırken, içecekler kategorisinden ürün alan müşterilerin üçte birinden biraz fazlası (%36,37) tatlı ürünler kategorisinden ürün almaktadır. Bu ilişkiyi %13,03 ile kuru gıda⇒tatlı ürünler bağıntısı izlemektedir. Yüzde 10'un üzerinde desteğe sahip diğer bağıntılar da kuru gıda⇒içecekler, tatlı ürünler⇒açık şarküteri, açık şarküteri⇒içecekler, açık şarküteri⇒kuru gıda ve kuru gıda⇒açık şarküteri bağıntılarıdır.

Birliktelik analizleri değerlendirilirken güven değeri de önemli sonuçlar verebilmektedir. Analiz sonuçlarının güven değerine göre sıralanmış hali Tablo 17'de görülmektedir. Güven değeri yüksek olmasına rağmen, kural desteği (toplam fişler arasında görülme adedi) düşük olduğu için Tablo 16'ya giremeyen bağıntılar bu şekilde incelenebilmektedir.

*Tablo 17. Kategorilere Uygulanmış Apriori Modelinin Sonuçları (Güven Değerine Göre Sıralı)*

Öncül	Artçıl	Destek%	Güven %	Kural Desteği %
Market markaları	Tatlı ürünler	6,90	69,46	4,79
Market markaları	Kuru gıda	6,90	62,50	4,31
Açık şarküteri	Tatlı ürünler	19,59	59,94	11,74
Market markaları	İçecekler	6,90	58,75	4,05
Temizlik	Tatlı ürünler	16,20	57,84	9,37
Etler	Açık şarküteri	9,22	55,75	5,15
Etler	Tatlı ürünler	9,22	55,59	5,13
Market markaları	Temizlik	6,90	55,20	3,81
Kişisel bakım	Tatlı ürünler	15,83	54,58	8,64
Etler	Kuru gıda	9,22	53,49	4,93

Tablo 17 incelendiğinde en yüksek güven değerinin %69,46 ile market markası⇒tatlı ürünler bağıntısında olduğu görülmektedir. Market markası kategorisinden ürün alan müşterilerin %70'i aynı zamanda tatlı ürünler kategorisinden de ürün almaktadır. Bu bağıntıya fişlerin yaklaşık %5'inde rastlanmaktadır. Market markası⇒kuru gıda, açık şarküteri⇒tatlı ürünler, market markası⇒içecekler, temizlik⇒tatlı ürünler, etler⇒açık şarküteri, market markası⇒temizlik, içecekler ⇒tatlı ürünler, etler⇒kuru gıda bağıntılarının hepsinde de %50'nin üzerinde güven değeri görülmektedir. Yani ilk kategoriden bir ürün alan müşterilerin yarıdan fazlası, ikinci kategoriden de ürün almaktadır. Açık

Şarküteri⇒tatlı ürünler bağıntısına fişlerin %10'undan fazlasında rastlanması da ilgi çekici bir başka bağıntıdır.

Yukarıda elde edilen bulguların istatistikî açıdan doğruluğunu test etmek amacıyla Bölüm 3'de "ki-kare yöntemiyle birliktelik kurallarının değerlendirilmesi" kısmında bahsedilen yöntem uygulanmıştır. Kurallara tek tek bu yöntem uygulandığında aşağıdaki sonuçlar elde edilmiştir.

*Tablo 18. Kategorilerdeki Birliktelik Kurallarının Ki-Kare Tablosu*

Öncül	Artçıl	Destek%	Güven %	Kural Desteği %	Ki-kare
Tatlı ürünler	İçecekler	35,91	52,04	18,69	90
İçecekler	Tatlı ürünler	51,39	36,37	18,69	90
Kuru gıda	Tatlı ürünler	24,86	52,41	13,03	42617
Tatlı ürünler	Kuru gıda	35,91	36,28	13,03	42617
Kuru gıda	İçecekler	24,86	49,23	12,24	710
Tatlı ürünler	Açık şarküteri	35,91	32,70	11,74	66384
Açık şarküteri	Tatlı ürünler	19,59	59,94	11,74	66384
Açık şarküteri	İçecekler	19,59	53,32	10,44	346
Açık şarküteri	Kuru gıda	19,59	52,99	10,38	112408
Kuru gıda	Açık şarküteri	24,86	41,76	10,38	112408

Ki-kare testi tüm k nesnelere k-şekilde bağımsız olup olmadığını belirlemektedir. Eğer değişkenler gerçekten bağımsızsa ki-kare değeri 0 olacaktır. Eğer kesim frekansından yüksekse (%5 güven aralığı için 3,841 ve %1 güven aralığı için 6,635) bağımsızlık varsayımı reddedilir. Tablo 18 incelendiğinde birliktelik kuralları bulunan tüm kategori türleri için ki-kare değerinin 6,635'den yüksek olduğu görülmektedir. Tüm kurallarda %1 anlamlılık derecesinde bir ilişki vardır.

Daha sonra yine Bölüm 3'de "ki-kare yöntemiyle birliktelik kurallarının değerlendirilmesi" kısmında bahsedilen ilgi değerleri hesaplanır.  $I(r)$  değeri 1'den uzaklaştıkça, r hücresindeki nesnelere bağımlılığı artar. 1'in üzerindeki değerler beklenenden fazla, 1'in altındaki değerler beklenenden az satış olduğunu gösterir. Daha önce hesaplanan ki-kare değeri anlamlı çıktığından bu ilişkilerin istatistiksel olarak anlamlı olduğunu söylenebilir.

Başka bir deyişle 1'in üzerindeki ilgi değerleri X ve Y arasında tamamlayıcı etkiyi, 1'in altındaki ilgi değerleri X ve Y arasında ikame edilebilir (yerine geçici)

etkiyi gösterir. İlgili deęerinin 1 olması veya 1'e çok yakın olması X ve Y arasında durumsal bağımsızlığa işaret eder

*Tablo 19. Tatlı ürünler-İçecekler İlgili Deęerleri Tablosu*

	“İçecekler”	“ <i>İçecekler</i> ”
Tatlı ürünler	1,011774	0,987543
<i>Tatlı ürünler</i>	0,993396	1,006987

Kural desteęi en yüksek olan tatlı ürünler⇒içecekler bağıntısının ilgili tablosu incelendiğinde ilginç bir durum ortaya çıkmıştır. Tablo 19'daki tüm hücreler 1'e yakın sayılardan oluşmaktadır. Dolayısıyla bu kategorideki ürünlerin satışının birbirinden bağımsız olduğu söylenebilir. Ki-kare deęeri istatistiksel olarak anlamlı olmasına rağmen, iki kategorinin bağımsız olarak belirlenmesi çelişki gibi gözükülebilir ama ki-kare hesaplanırken dört hücrenin de ki-kare deęerine katkı yaptığı unutulmamalıdır. Ayrıca Brin vd. (1997) göre ki-kare deęeri hesaplandıktan sonra, kategoriler arasındaki ilişki hakkında karar verebilmek için mutlaka ilgili deęerleri incelenmelidir.

*Tablo 20. Kuru gıda-tatlı ürünler İlgili Deęerleri Tablosu*

	“Kuru gıda”	“ <i>Kuru gıda</i> ”
Tatlı ürünler	1,457499	0,848478
<i>Tatlı ürünler</i>	0,743406	1,084983

Tablo 20 incelendiğinde kuru gıda kategorisi ile tatlı ürünler kategorileri arasında pozitif bir ilişki bulunduğu söylenebilir. Herhangi bir kategoriden ürün alıp, dięerinden almayanları belirten hücrelerdeki 0'dan küçük rakamlar da bu kategorilerin birlikte satıldığına işaret etmektedir.

Tablo 21. İçecekler- kuru gıda İlgili Değerleri Tablosu

	“İçecekler”	“İçecekler”
“kuru gıda”	0,95698	1,045514
“Kuru gıda ”	1,014248	0,984926

Tablo 21 incelendiğinde içecekler kategorisi ile kuru gıda kategorileri arasında bir bağıllık ilişkisi olmadığı görülecektir. Tüm hücrelerdeki bire yakın değerler bu kategoriler arasındaki satışların birbirinden bağımsız olduğunu göstermektedir.

Tablo 22. Açık Şarküteri-tatlı ürünler İlgili Değerleri Tablosu

	Açık şarküteri	“açık şarküteri”
Tatlı ürünler	1,665076	0,837641
Tatlı ürünler	0,626984	1,091061

Tablo 22 incelendiğinde açık şarküteri kategorisi ile tatlı ürünler kategorileri arasında güçlü bir pozitif ilişki bulunduğu söylenebilir. Herhangi bir kategoriden ürün alıp, diğerinden almayanları belirten hücrelerdeki 0’dan küçük rakamlar da bu kategorilerin birlikte satıldığına işaret etmektedir. Açık şarküteri kategorisinden ürün almış olmak ile tatlı ürünler kategorisinden ürün almamış olmak arasında büyük bir ters yönlü ilişki vardır.

Tablo 23. Açık Şarküteri-içecekler İlgili Değerleri Tablosu

	“İçecekler”	“İçecekler”
Açık şarküteri	1,034962	0,963011
“açık şarküteri”	0,991465	1,00903

Tablo 23 incelendiğinde içecekler kategorisi ile açık şarküteri kategorileri arasında bir bağıllık ilişkisi olmadığı görülecektir. Tüm hücrelerdeki bire yakın değerler bu kategoriler arasındaki satışların birbirinden bağımsız olduğunu göstermektedir.



Tablo 24. Açık Şarküteri-kuru gıda İlgili Değerleri Tablosu

	“Kuru gıda”	“Kuru gıda ”
Açık şarküteri	2,12622	0,626999
“açık şarküteri”	0,725066	1,091057

Tablo 24 incelendiğinde kuru gıda kategorisi ile açık şarküteri kategorileri arasında çok büyük bir pozitif ilişki bulunduğu söylenebilir. Herhangi bir kategoriden ürün alıp, diğerinden almayanları belirten hücrelerdeki 0’dan küçük rakamlar da bu kategorilerin birlikte satıldığına işaret etmektedir.

### 2.5.3. Kategoriler Arasındaki Çoklu Birlikte Satış İlişkilerinin İncelenmesi

Önceki kısımda bulunan ilişkiler sadece iki kategori arasındaki ilişkilereydi. Birden fazla kategorinin işin içine girdiği ilişkileri tespit etmek amacıyla GRI modelinin uygulama ayarları değiştirilmiş ve analiz yeniden gerçekleştirilmiştir. Elde edilen sonuçlar Tablo 25’de görülmektedir.

Tablo 25. Kategorilere Uygulanmış Çoklu GRI Modelinin Sonuçları (Kural Desteğine Göre Sıralı)

Öncül	Artçıl	Destek%	Güven %	Kural Desteği %
Kuru gıda&içecekler	Tatlı ürünler	12,24	66,44	8,13
Kuru gıda&tatlı ürünler	İçecekler	13,03	62,41	8,13
Tatlı ürünler&içecekler	Kuru gıda	18,69	43,51	8,13
Açık şarküteri&tatlı ürünler	Kuru gıda	11,74	64,08	7,52
Açık şarküteri&kuru gıda	Tatlı ürünler	10,38	72,49	7,52
Kuru gıda&tatlı ürünler	Açık şarküteri	13,03	57,76	7,52
Açık şarküteri&içecekler	Tatlı ürünler	10,44	70,30	7,34
Açık şarküteri&tatlı ürünler	İçecekler	11,74	62,53	7,34
Tatlı ürünler&içecekler	Açık şarküteri	18,69	39,29	7,34

Tablo 25 incelendiğinde en yüksek kural desteğinin (%8,13) içecekler, kuru gıda ve tatlı ürünler kategorileri arasında olduğu görülmektedir. Fişlerin %8’inde bu bağıntıya rastlanmaktadır. Özellikle “kuru gıda&içecekler ⇒Tatlı ürünler” ve “kuru gıda & tatlı ürünler ⇒ içecekler” bağıntılarındaki %60’ın üstündeki güven değeri

dikkat çekicidir. Yani ilk iki kategorideki ürünü alan müşteriler %60 olasılıkla üçüncü kategorideki ürünlerden de almaktadır.

Aynı analiz sonuçları güven değerine göre sıralandığında aşağıdaki tablo oluşmaktadır.

*Tablo 26. Kategorilere Uygulanmış Çoklu GRI Modelinin Sonuçları (Güven Değerine Göre Sıralı)*

Öncül	Artçıl	Destek%	Güven %	Kural Desteği %
Market markaları&açık şarküteri	Kuru gıda	3,54	80,45	2,85
Market markaları&açık şarküteri	Tatlı ürünler	3,54	79,67	2,82
Etler&temizlik	Kuru gıda	3,23	78,02	2,52
Market markaları&kuru gıda	Tatlı ürünler	4,31	77,85	3,35
Market markaları&içecekler	Tatlı ürünler	4,05	77,62	3,14
Temizlik&açık şarküteri	Kuru gıda	7,12	76,58	5,45
Market markaları&Kişisel Bakım	Tatlı ürünler	3,05	76,53	2,33
Market markaları&temizlik	Kuru gıda	3,81	76,02	2,89
Temizlik&açık şarküteri	Tatlı ürünler	7,12	75,63	5,38

Tablo 26 incelendiğinde “market markaları&açık şarküteri⇒kuru gıda” bağıntısındaki güven değerinin en yüksek olduğu görülmektedir. Yani market markaları ve açık şarküteri kategorilerinden ürün alan bir müşteri %80 ihtimalle kuru gıda kategorisinden bir ürün de almaktadır. Ama bu bağıntıya fişlerin sadece %2,85’inde rastlanmaktadır. “temizlik&açık şarküteri⇒kuru gıda” ve “temizlik&açık şarküteri⇒tatlı ürünler” bağlantılarında hem güven değeri %75’in üzerindedir hem de bu bağıntılara fişlerin %5’inden fazlasında rastlanmaktadır.

Yukarıda elde edilen bulguların istatistikî açıdan doğruluğunu test etmek amacıyla Bölüm 3’de “ki-kare yöntemiyle birliktelik kurallarının değerlendirilmesi” kısmında bahsedilen yöntem uygulanmıştır. Elde edilen kurallara tek bu yöntem uygulandığında aşağıdaki sonuçlar elde edilmiştir.

*Tablo 27. Kategorilerdeki Çoklu Birliktelik Kurallarının Ki-Kare Tablosu*

Öncül	Artçıl	Destek%	Güven %	Ki-kare
Kuru gıda&içecekler	Tatlı ürünler	12,24	66,44	61728
Kuru gıda&tatlı ürünler	İçecekler	13,03	62,41	7933
Tatlı ürünler&içecekler	Kuru gıda	18,69	43,51	46735
Açık şarküteri&tatlı ürünler	Kuru gıda	11,74	64,08	119720
Açık şarküteri&kuru gıda	Tatlı ürünler	10,38	72,49	73526
Kuru gıda&tatlı ürünler	Açık şarküteri	13,03	57,76	151074
Açık şarküteri&içecekler	Tatlı ürünler	10,44	70,30	65578
Açık şarküteri&tatlı ürünler	İçecekler	11,74	62,53	7207
Tatlı ürünler&içecekler	Açık şarküteri	18,69	39,29	61685

Tablo 27 incelendiğinde birliktelik kuralları bulunan tüm kategori türleri için ki-kare değerinin 6,63'den yüksektir ve %1 anlamlılık derecesinde anlamlı bir ilişkidir. Dolayısıyla kategorilerin istatistikî olarak birbirine bağımlı olduğu söylenebilir.

İlgi değerleri incelendiğinde aşağıdaki tablolar oluşmaktadır. Daha önce hesaplanan ki-kare değerleri anlamlı çıktığından, bu ilişkilerin istatistiksel olarak anlamlı olduğunu söylenebilir.

*Tablo 28. İçecekler&Kuru Gıda  $\bar{P}$  Tatlı Ürünler Bağıntısı İçin İlgili Tablosu*

	“İçecekler&Kuru Gıda”	$\overline{\text{“İçecekler&kuru gıda”}}$
Tatlı ürünler	1,848481	0,881662
$\overline{\text{Tatlı ürünler}}$	0,524119	1,066371

Tablo 28 incelendiğinde “içecekler ve kuru gıda” kategorisi ile tatlı ürünler kategorileri arasında pozitif bir ilişki bulunduğu söylenebilir. Herhangi bir kategoriden ürün alıp, diğerinden almayanları belirten hücrelerdeki 0’dan küçük rakamlar da bu kategorilerin birlikte satıldığına işaret etmektedir.

*Tablo 29. Kuru Gıda & Tatlı Ürünler  $\bar{P}$  İçecekler Bağıntısı İçin İlgili Tablosu*

	“Kuru gıda & Tatlı ürünler”	$\overline{\text{“Kuru gıda&Tatlı ürünler”}}$
İçecekler	1,213695	0,967984
$\overline{\text{İçecekler}}$	0,77392	1,033872

Tablo 29 incelendiğinde “kuru gıda ve tatlı ürünler” kategorisi ile içecekler kategorileri arasında pozitif bir ilişki bulunduğu söylenebilir. Ancak bu ilişki bir önceki kadar güçlü değildir.

*Tablo 30. İçecekler&Tatlı Ürünler  $\bar{P}$  Kuru Gıda Bağıntısı İçin İlgili Tablosu*

	“İçecekler&Tatlı Ürünler”	$\overline{\text{“İçecekler&Tatlı Ürünler”}}$
Kuru gıda	1,748373	0,827979
$\overline{\text{Kuru gıda}}$	0,752141	1,056973

Tablo 30 incelendiğinde “içecekler&tatlı ürünler” kategorisi ile kuru gıda kategorileri arasında güçlü pozitif bir ilişki bulunduğu söylenebilir. İçecekler ve tatlı ürünler kategorilerinden ürün alan müşterilerin, büyük ihtimalle kuru gıda kategorisinden ürün aldıkları söylenebilir.

*Tablo 31. Tatlı Ürünler&Açık Şarküteri P Kuru Gıda Bağıntısı İçin İlgili Tablosu*

	“Tatlı Ürünler&Açık Şarküteri”	“Tatlı Ürünler&Açık Şarküteri”
Kuru gıda	2,574558	0,790559
<i>Kuru Gıda</i>	0,478511	1,069366

Tablo 31 incelendiğinde “tatlı ürünler&açık şarküteri” kategorisi ile kuru gıda kategorileri arasında çok güçlü pozitif bir ilişki bulunduğu söylenebilir. Daha önce hesaplanan destek ve güven değerleri büyük ölçüde geçerlidir.

*Tablo 32. Kuru Gıda&Açık Şarküteri P Tatlı Ürünler Bağıntısı İçin İlgili Tablosu*

	“Kuru Gıda&Açık Şarküteri”	“Kuru Gıda&Açık Şarküteri”
Tatlı ürünler	2,016177	0,882305
<i>Tatlı Ürünler</i>	0,430065	1,066011

Tablo 32 incelendiğinde “kuru gıda&açık şarküteri” kategorisi ile tatlı ürünler kategorileri arasında çok güçlü pozitif bir ilişki bulunduğu söylenebilir. İlk iki kategoriden alışveriş yapanların, yapmayanlara göre üçüncü kategoriye alma olasılıkları artmaktadır.

*Tablo 33. Kuru gıda&Tatlı Ürünler P Açık Şarküteri Bağıntısı İçin İlgili Tablosu*

	“Kuru Gıda&Tatlı Ürünler”	“Kuru Gıda&Tatlı Ürünler”
Açık şarküteri	2,941227	0,709162
<i>Açık Şarküteri</i>	0,526106	1,071

Tablo 33 incelendiğinde “kuru gıda&tatlı ürünler” kategorisi ile açık şarküteri kategorileri arasında çok güçlü pozitif bir ilişki bulunduğu söylenebilir.

Tablo 34. *İçecekler&Açık Şarküteri P Tatlı Ürünler Bağıntısı İçin İlgili Tablosu*

	“İçecekler&Açık Şarküteri”	“İçecekler&Açık Şarküteri”
Tatlı ürünler	1,956602	0,88849
<i>Tatlı Ürünler</i>	0,463478	1,062542

Tablo 34 incelendiğinde “içecekler&açık şarküteri” kategorisi ile tatlı ürünler kategorileri arasında güçlü pozitif bir ilişki bulunduğu söylenebilir.

Tablo 35. *Açık Şarküteri&Tatlı Ürünler P İçecekler Bağıntısı İçin İlgili Tablosu*

	“Açık Şarküteri&Tatlı Ürünler”	“Açık Şarküteri&Tatlı Ürünler”
İçecekler	1,216166	0,971246
<i>İçecekler</i>	0,771305	1,03042

Tablo 35 incelendiğinde “açık şarküteri&tatlı ürünler” kategorisi ile içecekler kategorileri arasında pozitif bir ilişki bulunduğu söylenebilir. Ama bu ilişki daha önceki kategori çiftleri arındakiler kadar güçlü değildir.

Tablo 36. *Tatlı Ürünler&İçecekler P Açık Şarküteri Bağıntısı İçin İlgili Tablosu*

	“Tatlı Ürünler&İçecekler”	“Tatlı Ürünler&İçecekler”
Açık şarküteri	2,001443	0,769808
<i>Açık Şarküteri</i>	0,755527	1,056195

Tablo 36 incelendiğinde “tatlı ürünler&içecekler” kategorisi ile açık şarküteri kategorileri arasında güçlü pozitif bir ilişki bulunduğu söylenebilir.

#### 2.5.4. Kategorilerin Birlikte Satış İlişkilerinin Tek Tek İncelenmesi

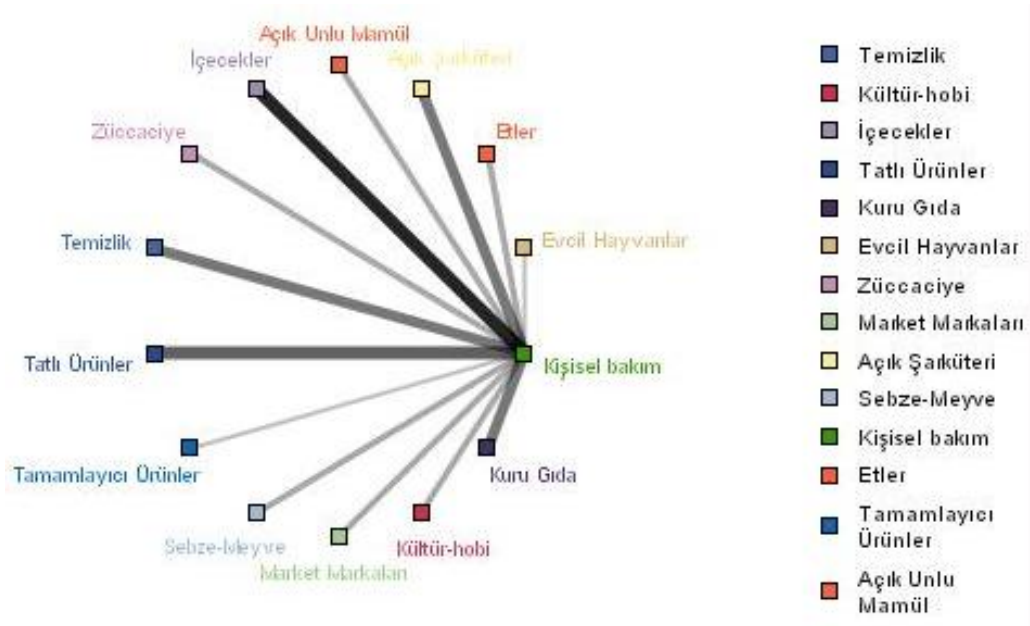
Tüm kategorilere ait genel bir değerlendirme yapıldıktan sonra, tek tek kategori bazında değerlendirmelere geçilmiştir. Bu amaçla model tekrar düzenlendikten sonra analizler tekrar uygulanmıştır. Bu sefer analizlerde girdi olarak

sadece tek bir kategori verilmiş, çıktı açısından tüm kategorilerin değerlendirilmesi istenmiştir.

İlk analiz olarak web grafiği uygulanmıştır. Burada tek bir kategorinin diğerleriyle olan ilişkisinin görülebilmesi için gerekli model ayarları yapıldıktan sonra analiz gerçekleştirilmiştir.

#### 2.5.4.1. Kişisel Bakım kategorisinin birlikte satış ilişkileri

Şekil 17’de kişisel bakım ile diğer kategoriler arasındaki ilişki grafiği görülmektedir. Şekil incelendiğinde en güçlü ilişkilerin “tatlı ürünler” ve “içecekler” kategorileriyle olduğu görülebilir.



Şekil 17. Kişisel Bakım Kategorisi İlişki Grafiği

Kişisel bakım kategorisi için Apriori algoritması sonucunda aşağıdaki bulgular elde edilmiştir. Tabloda güven sütununda “kişisel bakım” kategorisinden ürün alanların % kaçının diğer kategorilerden ürün aldığı görülmektedir. Aynı zamanda bulunan ilişkilerin geçerli olup olmadığının belirlenmesi amacıyla ki-kare değeri ve 1. hücredeki ilgi değeri de aynı tabloda gösterilmiştir.

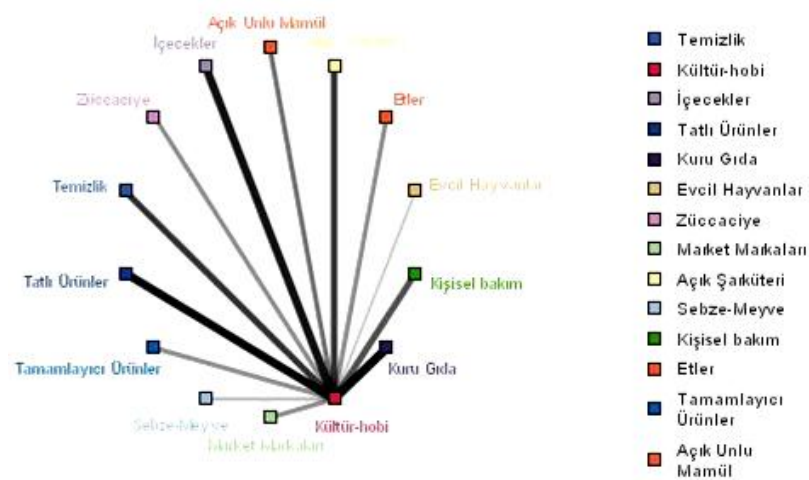
Tablo 37. “Kişisel Bakım” Kategorisi Birlikte Satış Kuralları Tablosu

Artçıl	Güven %	Kural Desteği %	Ki-kare değeri	İlgi değeri
Tatlı ürünler	54,59	8,64	31076	1,52
İçecekler	47,14	7,46	1525	0,91
Kuru gıda	45,81	7,25	48224	1,84
Temizlik	42,43	6,71	104087	2,61
Açık şarküteri	37,77	5,98	35358	1,92
Açık unlu mamul	20,98	3,32	2538	1,24
Market markaları	19,28	3,05	48779	2,79
Etler	17,39	2,75	16301	1,88
Züccaciye	14,71	2,32	11439	1,79
Sebze-meyve	13,85	2,19	6403	1,57
Kültür-hobi	13,17	2,08	7194	1,63
Tamamlayıcı ürünler	10,4	1,64	8534	1,83
Evcil hayvan	0,4	0,06	373	2,14

Tablo sonuçlarına göre “kişisel bakım” kategorisinden ürün alanların yarıdan fazlası (%54,5) aynı zamanda “tatlı ürünler” kategorisinden ürün de almaktadır. İlgili değeri geçerli bir bağıntı olduğunu göstermektedir. Ama aynı şey “kişisel bakım⇒içecekler” kategorisi için söylenemez. Çünkü ilgili değeri bu iki kategorinin neredeyse bağımsız olduğunu göstermektedir.

#### 2.5.4.2. Kültür-hobi kategorisinin birlikte satış ilişkileri

Şekil 18’de kültür-hobi kategorisi ile diğer kategoriler arasındaki ilişki grafiği görülmektedir. Şekil incelendiğinde en güçlü ilişkilerin “tatlı ürünler”, içecekler ve kuru gıda kategorileriyle olduğu görülebilir.



Şekil 18. Kültür-Hobi Kategorisi İlişki Grafiği

Kültür-hobi kategorisi için Apriori algoritması sonucunda aşağıdaki bulgular elde edilmiştir. Tabloda güven sütununda “kültür-hobi” kategorisinden ürün alanların % kaçının diğer kategorilerden ürün aldığı görülmektedir. Aynı zamanda bulunan ilişkilerin geçerli olup olmadığının belirlenmesi amacıyla ki-kare değeri ve 1. hücredeki ilgi değeri de aynı tabloda gösterilmiştir.

*Tablo 38. “Kültür-Hobi” Kategorisi Birliktelik Kuralları Tablosu*

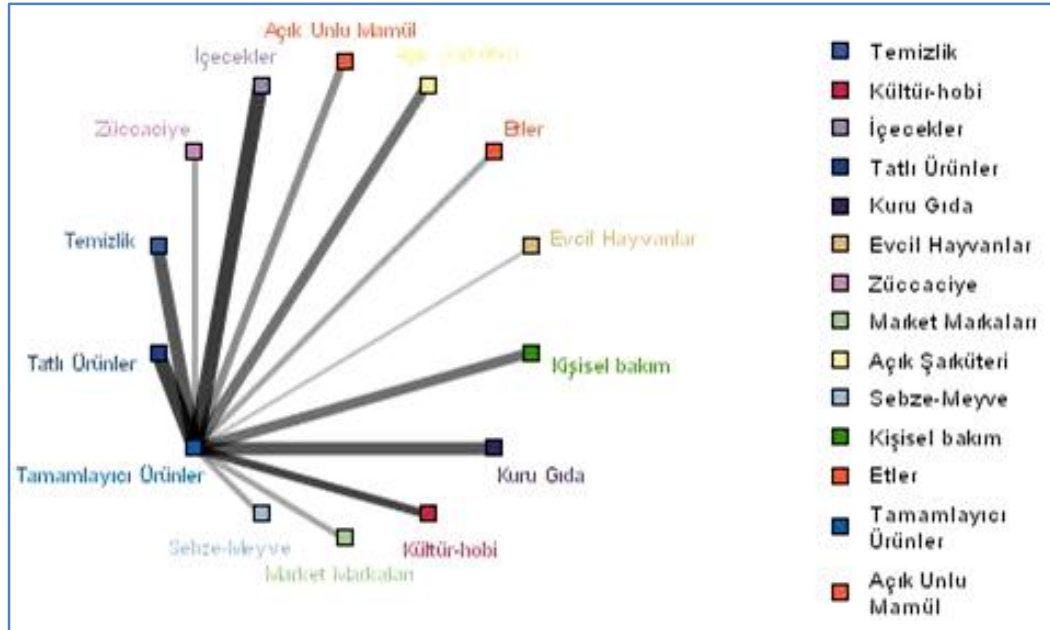
Artçıl	Güven %	Kural Desteği %	Ki-kare değeri	İlgi değeri
Tatlı ürünler	44.682	3.603	3161	1,24
İçecekler	36.976	2.982	8071	0,72
Kuru gıda	35.246	2.842	5487	1,41
Temizlik	29.832	2.406	13064	1,84
Açık şarküteri	28.888	2.330	5208	1,47
Kişisel bakım	25.856	2.085	7194	1,63
Açık unlu	17.379	1.401	19,5	1,03
Züccaciye	13.671	1.102	3744	1,66
Etler	12.964	1.045	1558	1,40
Market mar.	12.183	0.982	4102	1,76
Tamamlayıcı ür.	11.630	0.938	6316	2,04
Sebze-meyve	10.708	0.863	405	1,21
Evcil hayvanlar	0,335	0,027	77	1,79

Tablo sonuçlarına göre “kültür-hobi” kategorisinden ürün alanların yarıya yakınının (%44,6) aynı zamanda “Tatlı ürünler” kategorisinden ürün de almaktadır. İlgi değeri geçerli bir bağıntı olduğunu göstermektedir. Ama aynı şey “kültür-hobi ⇒içecekler” kategorisi için söylenemez. Çünkü ilgi değeri bu iki kategorinin bağımsız hatta ters yönlü olarak ilişkili olduğunu göstermektedir.

#### 2.5.4.3. Tamamlayıcı ürünler kategorisinin birlikte satış ilişkileri

Şekil 19’da tamamlayıcı ürünler kategorisi ile diğer kategoriler arasındaki ilişki grafiği görülmektedir. Şekil incelendiğinde en güçlü ilişkilerin “tatlı ürünler”, içecekler ve kuru gıda kategorileriyle olduğu görülebilir.





Şekil 19. Tamamlayıcı ürünler Kategorisi İlişki Grafiği

Tamamlayıcı ürünler kategorisi için Apriori algoritması sonucunda aşağıdaki bulgular elde edilmiştir. Tabloda güven sütununda “tamamlayıcı ürünler” kategorisinden ürün alanların % kaçının diğer kategorilerden ürün aldığı görülmektedir. Aynı zamanda bulunan ilişkilerin geçerli olup olmadığını belirlemek amacıyla ki-kare değeri ve 1. hücredeki ilgi değeri de aynı tabloda gösterilmiştir.

Tablo 39. “Tamamlayıcı Ürünler” Kategorisi Birliktelik Kuralları Tablosu

Artçıl	Güven %	Kural Desteği %	Ki-kare değeri	İlgi değeri
Tatlı ürünler	44.979	2.553	2308	1.252
İçecekler	40.630	2.306	3115	0.791
Kuru gıda	38.244	2.171	6529	1.538
Temizlik	35.222	1.999	17435	2.174
Açık şarküteri	30.693	1.742	5082	1.566
Kişisel bakım	29.019	1.647	8534	1.833
Açık unlu mamul	16.994	0.965	1,08*	1.012
Kültür-hobi	16.521	0.938	6316	2.049
Etler	14.135	0.802	1849	1.53
Market markaları	13.958	0.792	5035	2.022
Züccaciye	12.978	0.737	1959	1.579
Sebzeme-yuve	12.298	0.698	1025	1.395
Evcil hayvanlar	0.286	0.016	18	1.535



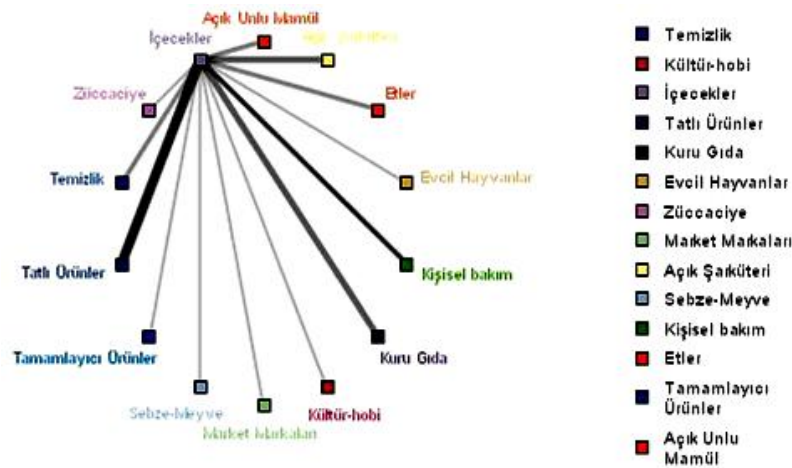
Tablo 40. “Züccaciye” Kategorisi Birliktelik Kuralları Tablosu

Artçıl	Güven %	Kural Desteği %	Ki-kare değeri	İlgi değeri
İçecekler	52.720	4.334	62	1.026
Tatlı ürünler	50.377	4.141	8836	1.403
Kuru gıda	38.266	3.145	9354	1.539
Açık şarküteri	34.554	2.840	13813	1.763
Temizlik	32.216	2.648	18415	1.988
Kişisel bakım	28.331	2.329	11439	1.789
Açık unlu mamul	24.420	2.007	4025	1.454
Etler	15.931	1.310	5195	1.726
Sebze-meyve	15.238	1.253	4562	1.728
Market markaları	14.941	1.228	9762	2.164
Kültür-hobi	13.411	1.102	3744	1.663
Tamamlayıcı ürünler	8.962	0.737	1959	1.579
Evcil hayvanlar	0.394	0.032	160	2.114

Tablo sonuçlarına göre “züccaciye” kategorisinden ürün alanların yarısından fazlası aynı zamanda içecekler kategorisinden ürün de almaktadır. “tatlı ürünler” kategorisi için de aynı şey geçerlidir. “tatlı ürünler” kategorisinin ilgi değeri geçerli bir bağıntı olduğunu göstermektedir. Ama aynı şey “züccaciye ⇒ içecekler” kategorisi için söylenemez. Çünkü ilgi değeri bu iki kategorinin nerdeyse tamamen bağımsız olduğunu göstermektedir.

#### 2.5.4.5. İçecekler kategorisinin birlikte satış ilişkileri

Şekil 21’de içecekler kategorisi ile diğer kategoriler arasındaki ilişki grafiği görülmektedir. Şekil incelendiğinde en güçlü ilişkinin “tatlı ürünler” kategorisiyle olduğu görülebilir.



Şekil 21. İçecekler Kategorisi İlişki Grafiği

İçecekler kategorisi için Apriori algoritması sonucunda aşağıdaki bulgular elde edilmiştir. Tabloda güven sütununda “içecekler” kategorisinden ürün alanların % kaçının diğer kategorilerden ürün aldığı görülmektedir. Aynı zamanda bulunan ilişkilerin geçerli olup olmadığının belirlenmesi amacıyla ki-kare değeri ve 1. hücredeki ilgi değeri de aynı tabloda gösterilmiştir.

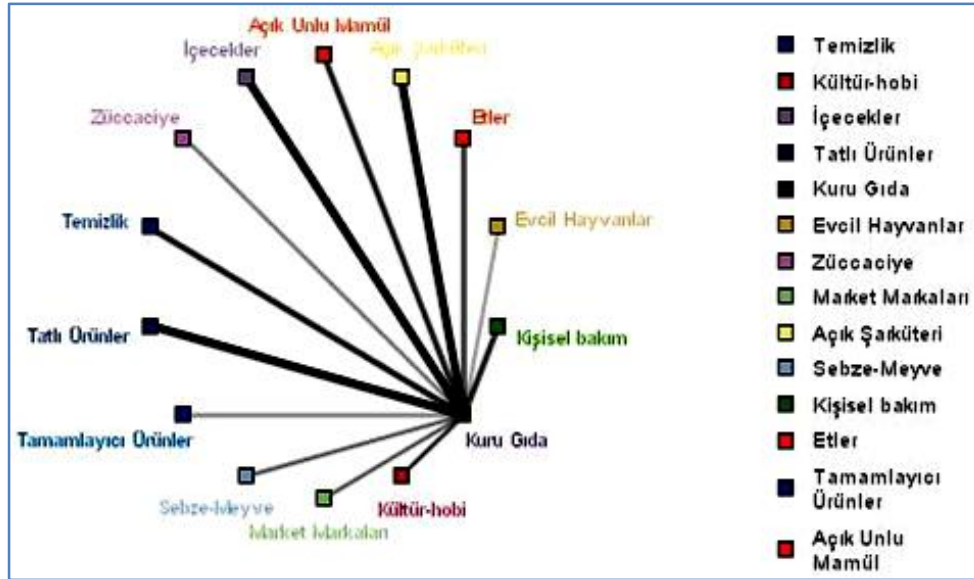
*Tablo 41. “İçecekler” Kategorisi Birliktelik Kuralları Tablosu*

Artçıl	Güven %	Kural Desteği %	Ki-kare değeri	İlgi değeri
Tatlı ürünler	36.374	18.693	90	1.013
Kuru gıda	23.820	12.241	708	0.958
Açık şarküteri	20.331	10.448	361	1.038
Temizlik	15.638	8.037	294	0.965
Kişisel bakım	14.524	7.464	1525	0.917
Açık unlu mamul	14.470	7.436	4595	0.862
Etlere	9.393	4.827	24	1.018
Sebze-meyve	8.723	4.483	22	0.989
Züccaciye	8.433	4.334	62	1.026
Market markaları	7.892	4.056	1684	1.143
Kültür-hobi	5.802	2.982	8071	0.720
Tamamlayıcı ürünler	4.488	2.306	3115	0.791
Evcil hayvanlar	0.155	0.080	139	0.831

Tablo sonuçlarına göre “içecekler” kategorisinden ürün alanların üçte biri (%36,37) aynı zamanda “tatlı ürünler” kategorisinden ürün de almaktadır. Üstelik bu bağıntıya fişlerin yaklaşık beşte birinde (%18,69) rastlanmaktadır. Ama hem bu bağıntı için hem de içecekler kategorisiyle ilgili diğer bağıntılar için ilgi değeri 1’e yakın olduğu için içecekler kategorisinin diğer kategorilerden bağımsız olduğu söylenebilir.

#### 2.5.4.6. Kuru gıda kategorisinin birlikte satış ilişkileri

Şekil 22’de kuru gıda kategorisi ile diğer kategoriler arasındaki ilişki grafiği görülmektedir. Şekil incelendiğinde en güçlü ilişkilerin “tatlı ürünler”, içecekler ve açık şarküteri kategorileriyle olduğu görülebilir.



Şekil 22. Kuru Gıda Kategorisi İlişki Grafiği

Kuru gıda kategorisi için Apriori algoritması sonucunda aşağıdaki bulgular elde edilmiştir. Tabloda güven sütununda “kuru gıda” kategorisinden ürün alanların % kaçının diğer kategorilerden ürün aldığı görülmektedir. Aynı zamanda bulunan ilişkilerin geçerli olup olmadığının belirlenmesi amacıyla ki-kare değeri ve 1. hücredeki ilgi değeri de aynı tabloda gösterilmiştir.

Tablo 42. “Kuru Gıda” Kategorisi Birliktelik Kuralları Tablosu

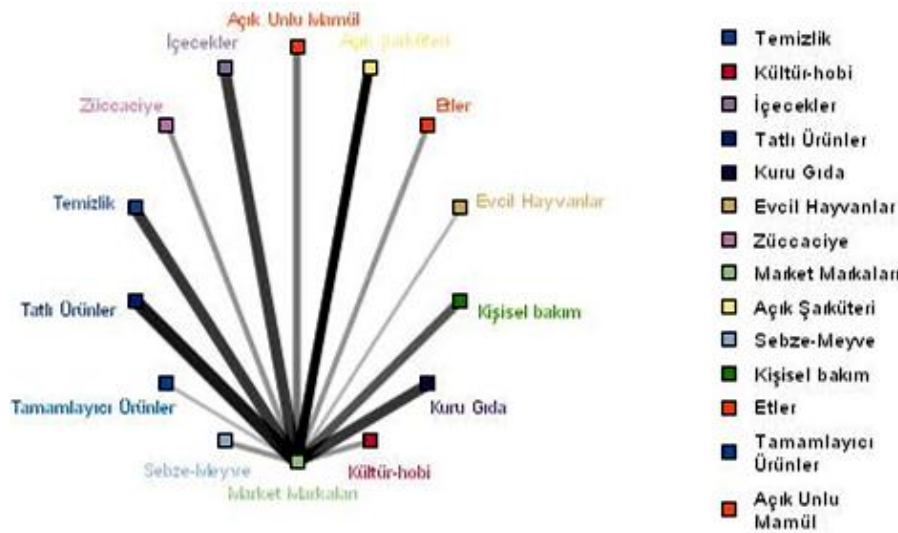
Artçıl	Güven %	Kural Desteği %	Ki-kare değeri	İlgi değeri
Tatlı ürünler	52.418	13.032	42658	1.460
İçecekler	49.238	12.241	708	0.958
Açık şarküteri	41.769	10.384	189562	2.132
Temizlik	34.410	8.555	88144	2.124
Kişisel bakım	29.175	7.253	48224	1.843
Açık unlu mamul	23.778	5.912	12508	1.416
Etler	19.857	4.937	48453	2.152
Market markaları	17.357	4.315	61191	2.514
Sebze-meyve	16.242	4.038	24552	1.842
Züccaciye	12.652	3.145	9354	1.539
Kültür-hobi	11.432	2.842	5487	1.418
Tamamlayıcı ürünler	8.732	2.171	6529	1.538
Evcil hayvanlar	0.338	0.084	290	1.813

Tablo incelendiğinde kuru gıda kategorisinin içecekler kategorisi dışındaki tüm kategorilerle aralarında pozitif bağımlılık olduğu söylenebilir. Kuru gıda

kategorisinden ürün alanların yarısından fazlası aynı zamanda “tatlı ürünler” kategorisinden ürün de almaktadır. Üstelik bu bağıntıya fişlerin %13’ünde rastlanmaktadır. Kuru gıda kategorisiyle açık şarküteri, temizlik, etler ve market markaları kategorileri arasında güçlü pozitif bağımlılık da dikkat çekmektedir.

#### 2.5.4.7. Market markaları kategorisinin birlikte satış ilişkileri

Şekil 23’de market markaları kategorisi ile diğer kategoriler arasındaki ilişki grafiği görülmektedir. Şekil incelendiğinde en güçlü ilişkilerin “tatlı ürünler”, içecekler ve açık şarküteri kategorileriyle olduğu görülebilir.



Şekil 23. Market Markaları Kategorisi İlişki Grafiği

Market markaları kategorisi için Apriori algoritması sonucunda aşağıdaki bulgular elde edilmiştir. Tabloda güven sütununda “market markaları” kategorisinden ürün alanların % kaçının diğer kategorilerden ürün aldığı görülmektedir. Aynı zamanda bulunan ilişkilerin geçerli olup olmadığının belirlenmesi amacıyla ki-kare değeri ve 1. hücredeki ilgi değeri de aynı tabloda gösterilmiştir.

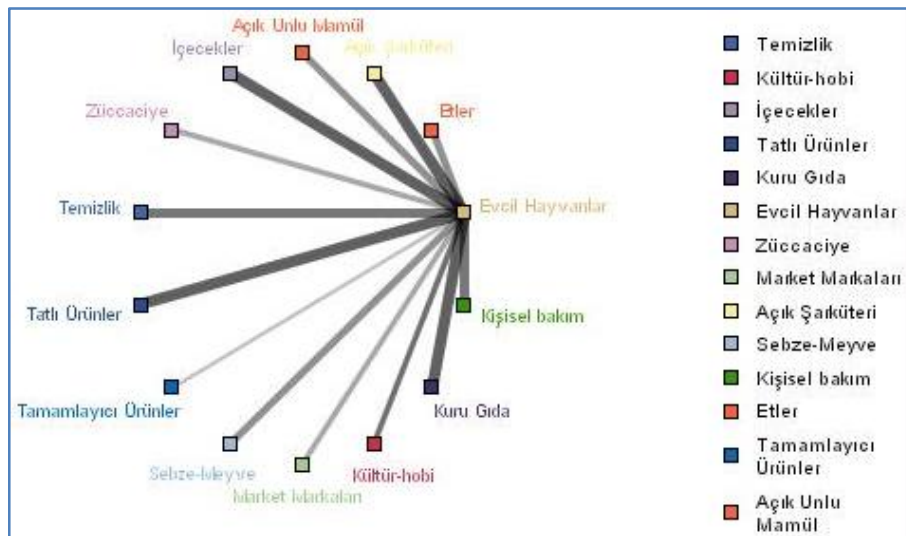
Tablo 43. Market Markaları Kategorisi Birliktelik Kuralları Tablosu

Artçıl	Güven %	Kural Desteği %	Ki-kare değeri	İlgi değeri
Tatlı ürünler	69.463	4.795	39354	1.934
Kuru gıda	62.505	4.315	61191	2.514
İçecekler	58.751	4.056	1684	1.143
Temizlik	55.105	3.811	90464	3.407
Açık şarküteri	51.319	3.543	51483	2.619
Kişisel bakım	44.216	3.052	48779	2.792
Açık unlu mamul	27.228	1.880	6229	1.622
Etler	23.898	1.650	20634	2.590
Züccaciye	17.790	1.228	9762	2.164
Sebze-meyve	16.431	1.134	5742	1.863
Kültür-hobi	14.231	0.982	4102	1.765
Tamamlayıcı ür.	11.477	0.792	5035	2.022
Evcil hayvanlar	0.367	0.025	97	1.971

Tablo incelendiğinde market markaları kategorisinin çoğu kategoriyle aralarında kuvvetli pozitif bağımlılık olduğu söylenebilir. Market markaları kategorisinden ürün alanların yarısından fazlası aynı zamanda “tatlı ürünler”, kuru gıda, içecekler ve temizlik kategorilerinden ürün de almaktadır.

#### 2.5.4.8. Evcil hayvanlar kategorisinin birlikte satış ilişkileri

Şekil 24’de evcil hayvanlar kategorisi ile diğer kategoriler arasındaki ilişki grafiği görülmektedir. Şekil incelendiğinde en güçlü ilişkilerin içecekler, tatlı ürünler, kuru gıda ve açık şarküteri kategorileriyle olduğu görülebilir.



Şekil 24. Evcil Hayvanlar Kategorisi İlişki Grafiği

Evcil hayvanlar kategorisi için Apriori algoritması sonucunda aşağıdaki bulgular elde edilmiştir. Tabloda güven sütununda evcil hayvanlar kategorisinden ürün alanların % kaçının diğer kategorilerden ürün aldığı görülmektedir. Aynı zamanda bulunan ilişkilerin geçerli olup olmadığının belirlenmesi amacıyla ki-kare değeri ve 1. hücredeki ilgi değeri de aynı tabloda gösterilmiştir.

*Tablo 44. Evcil Hayvanlar Kategorisi Birliktelik Kuralları Tablosu*

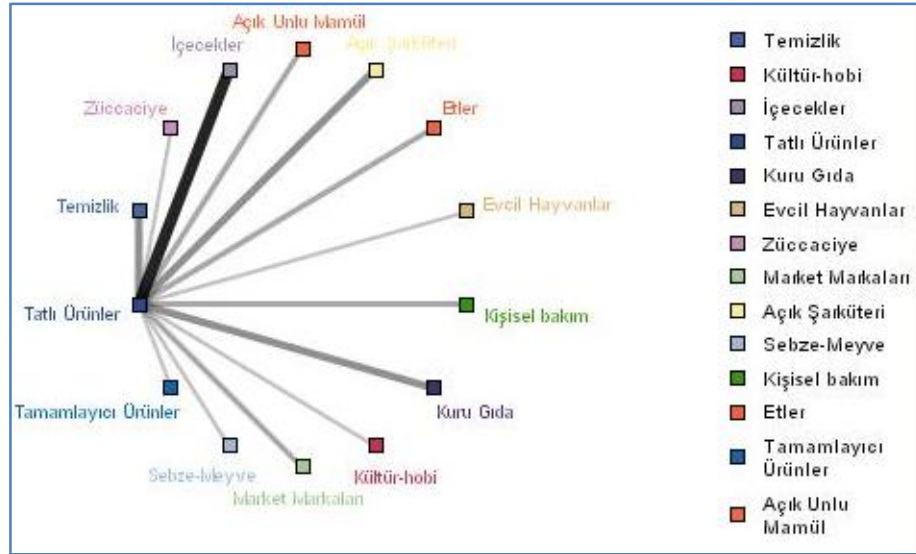
Artçıl	Güven %	Kural Desteği %	Ki-kare değeri	İlgi değeri
Tatlı ürünler	47.920	0.089	64	1.334
Kuru gıda	45.081	0.084	290	1.813
İçecekler	42.682	0.080	139	0.831
Açık şarküteri	40.137	0.075	410	2.048
Temizlik	36.711	0.068	475	2.266
Kişisel bakım	33.774	0.063	373	2.133
Açık unlu mamul	26.138	0.049	80	1.557
Etler	20.558	0.038	229	2.228
Sebze-meyve	20.362	0.038	216	2.309
Züccaciye	17.376	0.032	160	2.114
Kültür-hobi	14.488	0.027	77	1.797
Market markaları	13.607	0.025	97	1.971
Tamamlayıcı ür.	8.713	0.016	18	1.535

Tablo incelendiğinde evcil hayvanlar kategorisinin çoğu kategoriyle aralarında kuvvetli pozitif bağımlılık olduğu söylenebilir. Ama kural desteklerinin çok düşük olması bu bağıntıların uygulanabilirliğini düşürmektedir.

#### 2.5.4.9. Tatlı ürünler kategorisinin birlikte satış ilişkileri

Şekil 25'de tatlı ürünler kategorisi ile diğer kategoriler arasındaki ilişki grafiği görülmektedir. Şekil incelendiğinde tatlı ürünler kategorisi ile içecekler kategorisi arasındaki kuvvetli ilişki göze çarpmaktadır.





Şekil 25. Tatlı Ürünler Kategorisi İlişki Grafiği

Tatlı ürünler kategorisi için Apriori algoritması sonucunda aşağıdaki bulgular elde edilmiştir. Tabloda güven sütununda tatlı ürünler kategorisinden ürün alanların % kaçının diğer kategorilerden ürün aldığı görülmektedir. Aynı zamanda bulunan ilişkilerin geçerli olup olmadığını belirlemek amacıyla ki-kare değeri ve 1. hücredeki ilgi değeri de aynı tabloda gösterilmiştir.

Tablo 45. Tatlı ürünler Kategorisi Birliktelik Kuralları Tablosu

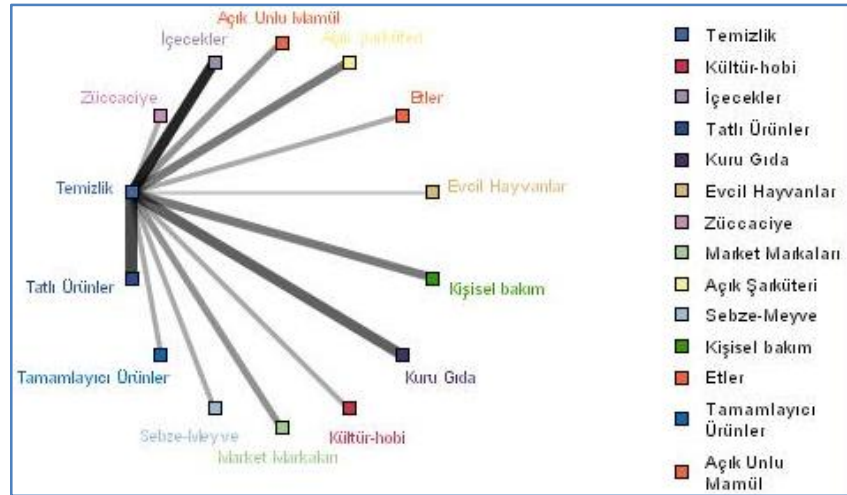
Artçıl	Güven %	Kural Desteği %	Ki-kare değeri	İlgi değeri
İçecekler	52.048	18.693	90	1.013
Kuru gıda	36.286	13.032	42658	1.460
Açık şarküteri	32.707	11.747	66580	1.669
Temizlik	26.094	9.372	43990	1.610
Kişisel bakım	24.066	8.643	31076	1.520
Açık unlu mamul	20.958	7.527	7488	1.248
Etler	14.284	5.130	18453	1.548
Market markaları	13.352	4.795	39354	1.934
Sebzeme meyve	12.616	4.531	10802	1.431
Züccaciye	11.530	4.141	8836	1.403
Kültür-hobi	10.033	3.603	3161	1.244
Tamamlayıcı ür.	7.109	2.553	2308	1.252
Evcil hayvanlar	0.249	0.089	64	1.334

Tablo incelendiğinde tatlı ürünler kategorisinin çoğu kategoriyle aralarında pozitif bağımlılık olduğu söylenebilir. Kural desteklerinin de yüksek olması bu bağıntıların değerini arttırmaktadır. Sadece içecekler kategorisiyle arasında yüksek

bir güven değeri ve kural desteği olmasında rağmen, ilgi değerinin 1 olması bu iki kategorinin istatistiksel olarak bağımsız olduğunu göstermektedir.

#### 2.5.4.10. Temizlik kategorisinin birlikte satış ilişkileri

Şekil 26'da temizlik kategorisi ile diğer kategoriler arasındaki ilişki grafiği görülmektedir. Şekil incelendiğinde temizlik kategorisi ile Tatlı ürünler, içecekler ve kuru gıda kategorileri arasındaki kuvvetli ilişki göze çarpmaktadır.



Şekil 26. Temizlik Kategorisi İlişki Grafiği

Temizlik kategorisi için Apriori algoritması sonucunda aşağıdaki bulgular elde edilmiştir. Tabloda güven sütununda temizlik kategorisinden ürün alanların % kaçının diğer kategorilerden ürün aldığı görülmektedir. Aynı zamanda bulunan ilişkilerin geçerli olup olmadığının belirlenmesi amacıyla ki-kare değeri ve 1. hücredeki ilgi değeri de aynı tabloda gösterilmiştir.

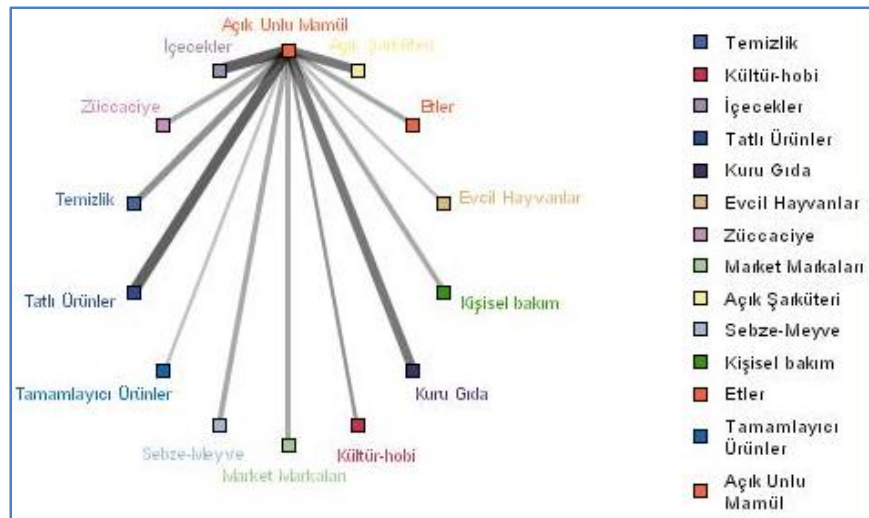
Tablo 46. Temizlik Kategorisi Birliktelik Kuralları Tablosu

Artçıl	Güven %	Kural Desteği %	Ki-kare değeri	İlgi değeri
Tatlı ürünler	57.840	9.372	43990	1.610
Kuru gıda	52.801	8.555	88144	2.124
İçecekler	49.601	8.037	294	0.965
Açık şarküteri	43.964	7.123	79394	2.244
Kişisel bakım	41.462	6.718	104087	2.619
Market markaları	23.521	3.811	90464	3.407
Açık unlu mamul	22.996	3.726	5739	1.369
Etler	19.971	3.236	28896	2.164
Züccaciye	16.344	2.648	18415	1.988
Sebze-meyve	15.709	2.545	12329	1.781
Kültür-hobi	14.848	2.406	13064	1.841
Tamamlayıcı ür.	12.340	1.999	17435	2.174
Evcil hayvanlar	0.422	0.068	475	2.266

Tablo incelendiğinde temizlik kategorisinin çoğu kategoriyle aralarında pozitif bağımlılık olduğu söylenebilir. Özellikle temizlik kategorisinden ürün alanların yarısından fazlasının tatlı ürünler veya kuru gıda kategorilerinden de ürün alıyor olmaları ilgi çekici bir bağıntıdır.

#### 2.5.4.11. Açık Unlu Mamul kategorisinin birlikte satış ilişkileri

Şekil 27’de açık unlu mamul kategorisi ile diğer kategoriler arasındaki ilişki grafiği görülmektedir. Şekil incelendiğinde açık unlu mamul kategorisi ile tatlı ürünler, içecekler ve kuru gıda kategorileri arasındaki kuvvetli ilişki göze çarpmaktadır.



Şekil 27. Açık Unlu Mamul Kategorisi İlişki Grafiği

Açık unlu mamul kategorisi için Apriori algoritması sonucunda aşağıdaki bulgular elde edilmiştir. Tabloda güven sütununda Açık unlu mamul kategorisinden ürün alanların % kaçının diğer kategorilerden ürün aldığı görülmektedir. Aynı zamanda bulunan ilişkilerin geçerli olup olmadığının belirlenmesi amacıyla ki-kare değeri ve 1. hücredeki ilgi değeri de aynı tabloda gösterilmiştir.

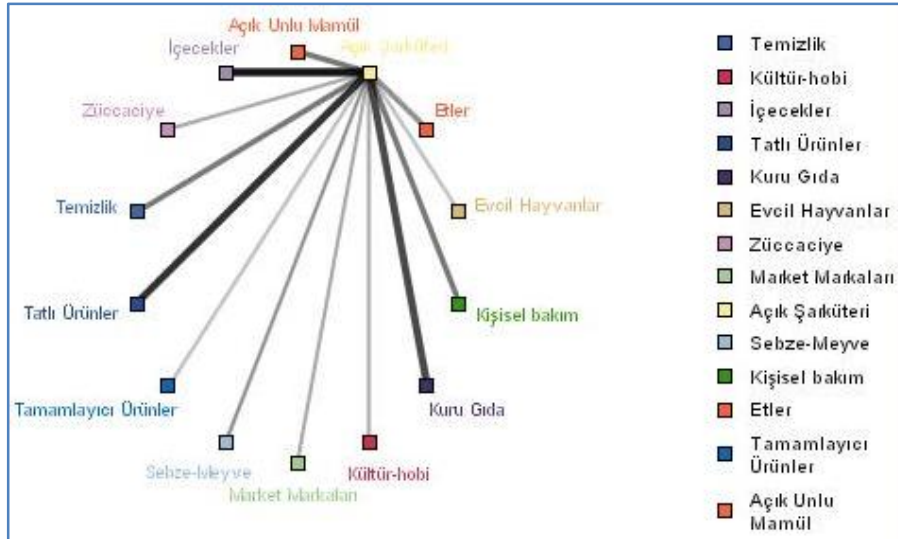
*Tablo 47. Açık Unlu Mamul Kategorisi Birliktelik Kuralları Tablosu*

<b>Artçıl</b>	<b>Güven %</b>	<b>Kural Desteği %</b>	<b>Ki-kare değeri</b>	<b>İlgi değeri</b>
Tatlı ürünler	44.826	7.527	7488	1.248
İçecekler	44.286	7.436	4595	0.862
Açık şarküteri	39.228	6.587	53663	2.002
Kuru gıda	35.205	5.912	12508	1.416
Temizlik	22.189	3.726	5739	1.369
Etler	21.390	3.592	38639	2.318
Sebze-meyve	20.001	3.358	33987	2.268
Kişisel bakım	19.782	3.322	2538	1.249
Züccaciye	11.954	2.007	4025	1.454
Market markaları	11.194	1.880	6229	1.622
Kültür-hobi	8.346	1.401	19,5	1.035
Tamamlayıcı ür.	5.745	0.965	1,08	1.012
Evcil hayvanlar	0.290	0.049	80	1.557

Tablo incelendiğinde açık unlu mamul kategorisinin çoğu kategoriyle aralarında pozitif bağımlılık olduğu söylenebilir. İçecekler kategorisiyle arasında ters yönlü ilişki söz edilebilir. Kültür-hobi ve tamamlayıcı ürünler kategorilerinden de bağımsız olduğu ilgi değerinden anlaşılmaktadır.

#### 2.5.4.12. Açık Şarküteri kategorisinin birlikte satış ilişkileri

Şekil 28'de açık şarküteri kategorisi ile diğer kategoriler arasındaki ilişki grafiği görülmektedir. Şekil incelendiğinde açık şarküteri kategorisi ile Tatlı ürünler, içecekler ve kuru gıda kategorileri arasındaki kuvvetli ilişki göze çarpmaktadır.



Şekil 28. Açık Şarküteri Kategorisi İlişki Grafiği

Açık şarküteri kategorisi için Apriori algoritması sonucunda aşağıdaki bulgular elde edilmiştir. Tabloda güven sütununda açık şarküteri kategorisinden ürün alanların % kaçının diğer kategorilerden ürün aldığı görülmektedir. Aynı zamanda bulunan ilişkilerin geçerli olup olmadığının belirlenmesi amacıyla ki-kare değeri ve 1. hücredeki ilgi değeri de aynı tabloda gösterilmiştir.

Tablo 48. Açık Şarküteri Kategorisi Birliktelik Kuralları Tablosu

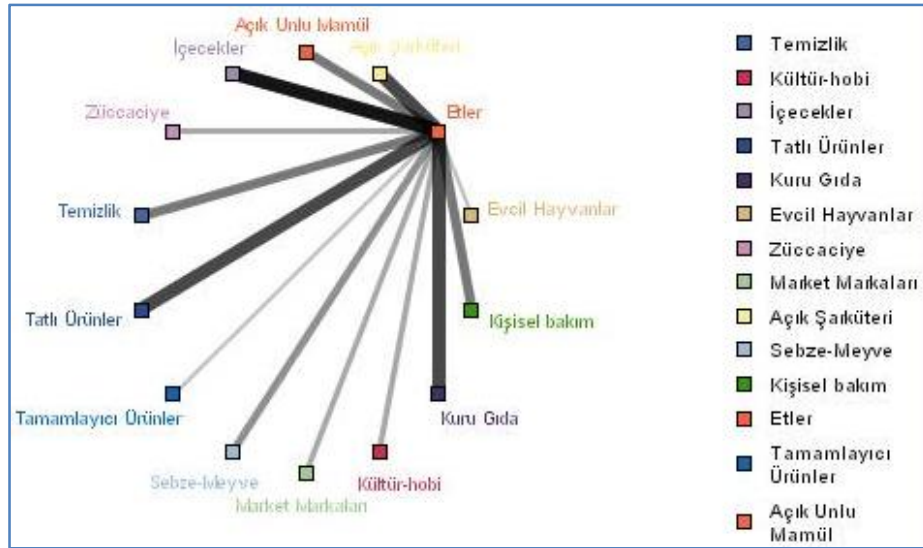
Artçıl	Güven %	Kural Desteği %	Ki-kare değeri	İlgi değeri
Tatlı ürünler	59.948	11.747	66580	1.669
İçecekler	53.322	10.448	361	1.038
Kuru gıda	52.996	10.384	189562	2.132
Temizlik	36.353	7.123	79394	2.244
Açık unlu mamul	33.616	6.587	53663	2.002
Kişisel bakım	30.525	5.981	35358	1.928
Etiler	26.259	5.145	91692	2.846
Sebzeye-meyve	21.952	4.301	56724	2.489
Market markaları	18.081	3.543	51483	2.619
Züccaciye	14.496	2.840	13813	1.763
Kültür-hobi	11.889	2.330	5208	1.474
Tamamlayıcı ür.	8.892	1.742	5082	1.566
Evcil hayvanlar	0.382	0.075	410	2.048

Tablo incelendiğinde açık şarküteri kategorisinin hemen hemen tüm kategorilerle arasında pozitif bağımlılık olduğu söylenebilir. İçecekler kategorisinden bağımsız olduğu ilgi değerine bakarak söylenebilir. Açık şarküteri kategorisinden

ürün alanları yarısından fazlasının tatlı ürünler veya kuru gıda kategorilerinde de ürün alması ilgi çekici bir bağıntıdır.

#### 2.5.4.13. Etler kategorisinin birlikte satış ilişkileri

Şekil 29'da etler kategorisi ile diğer kategoriler arasındaki ilişki grafiği görülmektedir. Şekil incelendiğinde etler kategorisi ile Tatlı ürünler, içecekler ve kuru gıda kategorileri arasındaki kuvvetli ilişki göze çarpmaktadır.



Şekil 29. Etler Kategorisi İlişki Grafiği

Etler kategorisi için Apriori algoritması sonucunda aşağıdaki bulgular elde edilmiştir. Tabloda güven sütununda etler kategorisinden ürün alanların % kaçının diğer kategorilerden ürün aldığı görülmektedir. Aynı zamanda bulunan ilişkilerin geçerli olup olmadığının belirlenmesi amacıyla ki-kare değeri ve 1. hücredeki ilgi değeri de aynı tabloda gösterilmiştir.

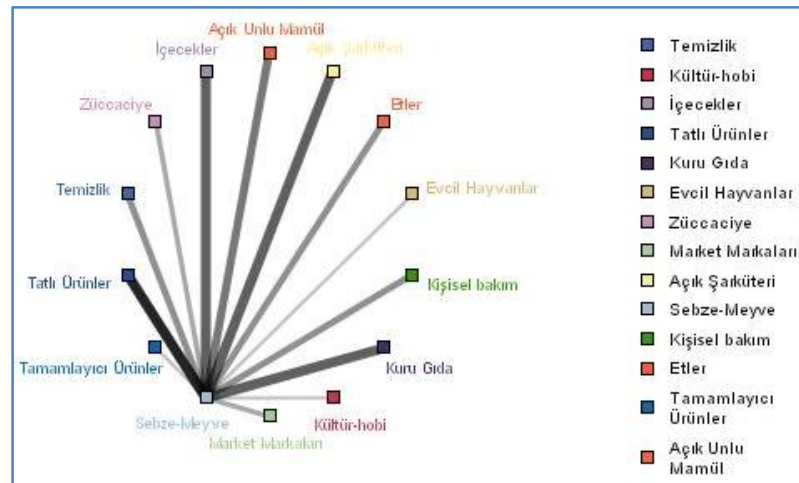
Tablo 49. Etler Kategorisi Birliktelik Kuralları Tablosu

Artçıl	Güven %	Kural Desteği %	Ki-kare değeri	İlgi değeri
Açık şarküteri	55.759	5.145	91962	2.846
Tatlı ürünler	55.594	5.130	18453	1.548
Kuru gıda	53.497	4.937	48453	2.152
İçecekler	52.312	4.827	24	1.018
Açık unlu mamul	38.921	3.592	38639	2.318
Temizlik	35.065	3.236	28896	2.164
Kişisel bakım	29.836	2.753	16201	1.884
Sebze-meyve	25.955	2.395	40216	2.943
Market markaları	17.878	1.650	20634	2.590
Züccaciye	14.191	1.310	5195	1.726
Kültür-hobi	11.329	1.045	1558	1.405
Tamamlayıcı ür...	8.695	0.802	1849	1.532
Evcil hayvanlar	0.415	0.038	229	2.228

Tablo incelendiğinde etler kategorisinin hemen hemen tüm kategorilerle arasında pozitif bağımlılık olduğu söylenebilir. Sadece içecekler kategorisinden bağımsız olduğu ilgi değerine bakarak söylenebilir. Etler kategorisinden ürün alanları yarından fazlasının tatlı ürünler, açık şarküteri veya kuru gıda kategorilerinde de ürün alması ilgi çekici bir bağıntıdır.

#### 2.5.4.14. Sebze-meyve kategorisinin birlikte satış ilişkileri

Şekil 30'da sebze-meyve kategorisi ile diğer kategoriler arasındaki ilişki grafiği görülmektedir. Şekil incelendiğinde sebze-meyve kategorisi ile tatlı ürünler ve kuru gıda kategorileri arasındaki kuvvetli ilişki göze çarpmaktadır.



Şekil 30. Sebze-meyve Kategorisi İlişki Grafiği

Sebze-meyve kategorisi için Apriori algoritması sonucunda aşağıdaki bulgular elde edilmiştir. Tabloda güven sütununda sebze-meyve kategorisinden ürün alanların % kaçının diğer kategorilerden ürün aldığı görülmektedir. Aynı zamanda bulunan ilişkilerin geçerli olup olmadığının belirlenmesi amacıyla ki-kare değeri ve 1. hücredeki ilgi değeri de aynı tabloda gösterilmiştir.

*Tablo 50. Sebze-Meyve Kategorisi Birliktelik Kuralları Tablosu*

<b>Artçıl</b>	<b>Güven %</b>	<b>Kural Desteği %</b>	<b>Ki-kare değeri</b>	<b>İlgi değeri</b>
Tatlı ürünler	51.385	4.531	10802	1.431
İçecekler	50.840	4.483	22	0.989
Açık şarküteri	48.780	4.301	56724	2.489
Kuru gıda	45.793	4.038	24552	1.842
Açık unlu mamul	38.087	3.358	33987	2.268
Temizlik	28.864	2.545	12329	1.781
Etlere	27.162	2.395	40216	2.943
Kişisel bakım	24.879	2.194	6403	1.571
Züccaciye	14.205	1.253	4562	1.728
Market markaları	12.864	1.134	5742	1.863
Kültür-hobi	9.793	0.863	405	1.214
Tamamlayıcı ür.	7.917	0.698	1025	1.395
Evcil hayvanlar	0.430	0.038	216	2.309

Tablo incelendiğinde sebze-meyve kategorisinin hemen hemen tüm kategorilerle arasında pozitif bağımlılık olduğu söylenebilir. Sadece içecekler kategorisinden bağımsız olduğu ilgi değerine bakarak söylenebilir. Sebze-meyve kategorisi ile açık şarküteri, açık unlu mamul ve etler kategorileri arasındaki kuvvetli ilişki de dikkat çekicidir.



Tablo 51. Kategoriler Arasındaki İlgili Değerleri

	Kişisel bakım	Kültür-hobi	Tamamlayıcı ürün	Züccaciye	İçecekler	Kuru gıda	Market mark.	Evcil hayvan	Tatlı ürünler	Temizlik	Açık unlu mamul	Açık şarküteri	Etler
Kültür-hobi	1,63	-											
Tamamlayıcı ürün	1,83	2,04	-										
Züccaciye	1,79	1,66	1,57	-									
İçecekler	0,91	0,72	0,79	1,02	-								
Kuru gıda	1,84	1,41	1,53	1,53	0,95	-							
Market mark.	2,79	1,76	2,02	2,16	1,14	2,51	-						
Evcil hayvan	2,14	1,79	1,53	2,11	0,83	1,81	1,97	-					
Tatlı ürünler	1,52	1,24	1,25	1,40	1,01	1,46	1,93	1,33	-				
Temizlik	2,61	1,84	2,17	1,98	0,96	2,12	3,40	2,26	1,61	-			
Açık unlu mam.	1,24	1,03	1,01*	1,45	0,86	1,41	1,62	1,55	1,24	1,36	-		
Açık şarküteri	1,92	1,47	1,56	1,76	1,03	2,13	2,61	2,04	1,66	2,24	2,00	-	
Etler	1,88	1,40	1,53	1,72	1,01	2,15	2,59	2,28	1,54	2,16	2,31	2,84	-
Sebze-meyve	1,57	1,21	1,39	1,72	0,98	1,84	1,86	2,30	1,43	1,78	2,26	2,48	2,94

Tablo 51’de kategoriler arasındaki ilk hücrelerdeki ilgi değerleri tek bir tabloda gösterilmiştir. Genel kategori olarak gıda dışı, hızlı tüketim ve taze gıda kategorileri arasındaki ilişkiler gölgeli olarak gösterilmiştir.

### 2.5.5. Ürün Sınıfları Arasındaki Kümelenmenin Belirlenmesi

Kategoriler arasındaki ilişkiler belirlendikten sonra, alışverişler sırasında müşterilerin aldığı ürünlerin çoklu ilişkilerinin tespit edilebilmesi amacıyla hiyerarşik kümeleme analizi gerçekleştirilmiştir. Öncelikle ürünler toplam 83 sınıfa ayrılmış ve bu sınıfların her fişle dolayısıyla her alışveriş kaydıyla ilişkilendirilmesi sağlanmıştır.

Toplam 1.096.125 adet fiş içerisinde 6.009.317 alışveriş hareketi bulunmaktadır. Ürün sınıfları ve bu sınıfların fişlerde görülme frekans ve yüzdeleri Tablo 52’de verilmiştir.

Tablo 52. Ürün Sınıflarının Fişlerde Görülme Frekansları Ve Yüzdeleri

Gıda Maddeleri			Temizlik ve Diğer		
Sınıf Adı	Frekans	Yüzde	Sınıf Adı	Frekans	Yüzde
Çikolata, Şeker	207227	19,3	Sigara	129544	12
Unlu Mamul-hazır yemek	192272	17,9	Diğer	78727	7,3
Bisküvi, kek	172757	16,1	Bebek malzemeleri	65263	6,1
Kola, Gazoz	163927	15,2	züccaciye	61052	5,7
Su, Maden suyu	148474	13,8	Tuv.kâğıdı-peçete-havlu	55462	5,2
Meyve suyu	102990	9,6	Genel temizlik mal.	54509	5,1
Sebze Meyve	96666	9,0	Sabun -Duş jeli	53020	4,9
Peynir	90050	8,4	Oyuncak	49724	4,6
Sakız	83634	7,8	Şampuan	44812	4,2
Süt, Katkılı süt	82220	7,6	Plastik -ahşap malz.	44794	4,2
Hazır çorba	79732	7,4	Diş macunu -fırça	42430	3,9
Lokum	67047	6,2	Deodorant - krem, Kolonya	42294	3,9
Kuruyemiş	66385	6,2	Gazete	40579	3,8
Cips	60886	5,7	Çamaşır deterjan (makine)	40386	3,8
Sucuk-sosis	58020	5,4	İçki	35499	3,3
Makarna	45847	4,3	Kırtasiye	31991	3
Tavuk	42169	3,9	Kitap	30106	2,8
Çay, Bitki çayı	40893	3,8	Kadın pedi	24520	2,3
Yoğurt	40498	3,8	Makyaj-tüy dökücü	24325	2,3
Salca-konserve-sirke	39995	3,7	Erkek tıraş ürün.	23719	2,2
Zeytin	35891	3,3	Müzik cd-kaset	21942	2
Baharat - harçlar	35540	3,3	Oto-Araba malz.	20510	1,9
Çocuk gel. Yoğurt	34494	3,2	Oda parfümü	20121	1,9
Pekmez-tahin-helva	33506	3,1	Saç boyası-jöle	16233	1,5
Yumurta	33142	3,1	Çamaşır suyu	16147	1,5
Margarin	31728	2,9	Bulaşık det. (el)	15722	1,5
Bakliyat	29642	2,8	Dergi	14582	1,4
Şeker	29131	2,7	Bulaşık det. (makine)	13924	1,3
Kahve	29085	2,7	Yumuşatıcı	13288	1,2
Hazır Tatlı	27826	2,6	DVD-vcd	10243	1
Un, nişasta, Pasta Mal.	26267	2,4	Bebek temizlik mal.	7330	0,7
Dişabetik-diyet ürünler	25468	2,4	Pil	7185	0,7
Kıyma-et	23283	2,2	Prezervatif	4916	0,5
Ayran	22674	2,1	Güneş sütü- kremi	4901	0,5
Sürme çikolata	21726	2,0	Sinek öldürücü	3903	0,3
Mısır gevreği	20683	1,9	Evcil hayvan	2063	0,2
Ayçiçek yağı	19432	1,8	Çamaşır deterjan (el)	2042	0,2
Dondurma	18558	1,7			
Ekmek	16891	1,6			
Bebek maması-yiyeceği	15449	1,4			
Reçel-bal	14388	1,3			
Ketçap-mayonez	14372	1,3			
Tuz	13571	1,3			
Dondurulmuş ürünler	9550	0,9			
Zeytinyağı	6583	0,6			
Enerji içeceği	6389	0,6			

Tablo 52 incelendiğinde fişlerde en sık rastlanan gıda türü ürün sınıflarının çikolata-şeker (%19,3), Unlu mamul-hazır yemek (%17,9), Bisküvi-kek (%16,1), Kola-gazoz (%15,2) ve Su-maden suyu (%13,8) olduğu görülmektedir. Gıda dışı ürünlerde ise en sık rastlanan ürün sınıfları sigara (%12), bebek malzemeleri (%6,1), züccaciye (%5,7), kağıt temizlik ürünleri (%5,2) ve genel temizlik maddeleridir (%5,1).

Ürün sınıfları arasında gerçekleştirilen hiyerarşik kümeleme analizi sonucunda EK 1’de görülen yakınlık matrisi (proximity matrix), Tablo 53’de görülen toplu tablo (agglomeration schedule) ve EK 2’de görülen dendogram diyagramı elde edilmiştir.

Yakınlık matrisi ürünler arasındaki uzaklıkları veya benzerlikleri gösteren bir tablodur. Başka bir deyişle ürün sınıfları arasındaki ikili ilişkileri göstermektedir. İki sınıf arasındaki ilişki “-1 ile 1” aralığında değişebilmektedir. Bu değer 1 olması ürün sınıflarının sürekli bir arada görüldüğü, -1 olması ise bu sınıfların hiçbir zaman birlikte görülmediği anlamına gelmektedir. EK 1’de görülen yakınlık matrisi incelendiğinde çoğu ürün sınıfı arasında, küçük değerlerde de olsa, pozitif yönde ilişkiler olduğu söylenebilir. Ters yönlü ilişkiler genellikle içki, sigara, gazete, dergi gibi marketin ayrı bir bölümünde satışı yapılan ürünlerle diğer ürünler arasında söz konusudur.

En kuvvetli ilişkiler, toplu tabloda da kümeleme işlemleri sırasında ilk olarak aynı kümelere yerleştirilen ürün çiftleri arasında görülmektedir. Makarna-bakliyat (0,403), zeytin-peynir (0,354), çay-şeker (0,337), pasta malzemeleri (un, kabartma tozu, vb.)-hazır tatlı (puding, vb.) (0,308) bu ürün çiftleri arasında gösterilebilir.

Tablo 53’de verilen “toplu tablo” aslında kümeleme işleminin aşamalarını göstermektedir.

Tablo 53. Ürün Sınıfları Arasındaki Kümeleme Aşamalarını Gösteren Toplu Tablo

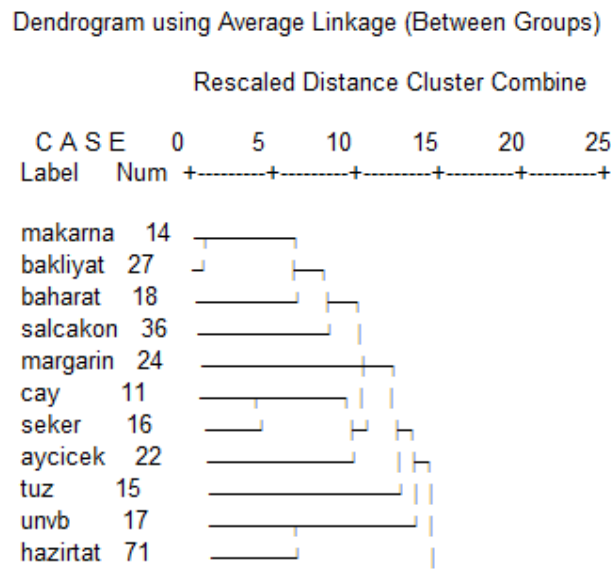
Aşama	Birleştirilen Kümeler		Katsayı	Kümenin ilk görülüş sırası		Sonraki aşama
	Küme 1	Küme 2		Küme 1	Küme 2	
1	Makarna	Bakliyat	,403	0	0	6
2	Zeytin	Peynir	,354	0	0	16
3	Çay	Şeker (Toz-Küp)	,337	0	0	12
4	Un-Pas.Mal.	Hazır Tatlı	,308	0	0	24
5	Çikolata	Bisküvi	,287	0	0	25
6	Makarna	Baharat	,284	1	0	9
7	Gen. Tem.	Kâğıt Tem.	,282	0	0	8
8	Gen. Tem.	Çamaşır Det.	,250	7	0	13
9	Makarna	Salça-Konserve	,249	6	0	14
10	Et-Kıyma	Tavuk	,240	0	0	32
11	Süt Ürünleri	Yoğurt	,238	0	0	17
12	Çay	Ayçiçek Yağı	,230	3	0	15
13	Gen. Tem.	Sabun	,230	8	0	18
14	Makarna	Margarin	,225	9	0	15
15	Çay	Makarna	,212	12	14	22
16	Zeytin	Sucuk-Sosis	,211	2	0	23
17	Süt Ürünleri	Yumurta	,206	11	0	23
18	Gen. Tem.	Çam. Suyu	,206	13	0	19
19	Gen. Tem.	Bulaşık Det.	,193	18	0	21
20	Ayran	Sebze-meyve	,184	0	0	30
21	Gen. Tem.	Yumuşatıcı	,180	19	0	27
22	Çay	Tuz	,180	15	0	24
23	Süt Ürünleri	Zeytin	,175	17	16	30
24	Çay	Un-Pas.Mal.	,173	22	4	29
25	Çikolata	Meyve Suyu	,172	5	0	31
26	Şampuan	Ağız Bakım Ür.	,162	0	0	36
27	Gen. Tem.	Bebek Mal.	,161	21	0	29
28	Pekmez-Helv	Lokum	,158	0	0	73
29	Çay	Gen. Tem.	,154	24	27	34
30	Süt Ürünleri	Unlu Mamul	,150	23	20	32
31	Çikolata	Kola	,142	25	0	37
32	Süt Ürünleri	Et-Kıyma	,137	30	10	44
33	Bebek Mama	Çocuk Yoğ.	,134	0	0	50
34	Çay	Bulaşık Mak.	,133	29	0	35
35	Çay	Ketçap-May.	,129	34	0	38
36	Şampuan	Erkek Traş.ür.	,127	26	0	43
37	Çikolata	Cips	,122	31	0	46
38	Çay	Hazır Çorba	,119	35	0	39
39	Çay	Krem Çik.	,116	38	0	42
40	Makyaj Mal.	Deodorant	,116	0	0	43
41	zuccaciye	Diğer	,114	0	0	48
42	Kahve	Çay	,106	0	39	44
43	Makyaj	Şampuan	,105	40	36	49

Tablo 53.'ün devamı

Aşama	Birleştirilen Kümeler		Katsay 1	Kümenin ilk görüldüğü sırası		Sonraki aşama
	Küme 1	Küme 2		Küme 1	Küme 2	
44	Süt Ürünleri	Kahve	,103	32	42	46
45	Kitap	Kırtasiye	,087	0	0	65
46	Çikolata	Süt	,087	37	44	51
47	Diyet Ür.	Ekmek	,084	0	0	56
48	Plastik-ağşap	züccaciye	,084	0	41	63
49	Makyaj	Kadın Pedi	,084	43	0	54
50	Bebek Mama	Mısır Gevreği	,083	33	0	55
51	Çikolata	Oda Parf.	,081	46	0	52
52	Çikolata	Zeytinyağı	,079	51	0	53
53	Çikolata	Reçel Bal	,079	52	0	55
54	Makyaj	Saç Boyası	,077	49	0	61
55	Çikolata	Bebek Maması	,074	53	50	56
56	Çikolata	Diyet Ürünler	,067	55	47	57
57	Çikolata	Dondurulmuş	,066	56	0	61
58	Kuruyemiş	İçki	,062	0	0	75
59	DVD-VCD	Müzik kaset	,061	0	0	70
60	Su-Maden su.	Sakız	,060	0	0	71
61	Çikolata	Makyaj	,060	57	54	63
62	Pil	Oyuncak	,059	0	0	65
63	Çikolata	Plastik-ağşap	,058	61	48	66
64	Dergi	Gazete	,041	0	0	81
65	Kitap	Pil	,034	45	62	70
66	Çikolata	Unlu Mamul	,032	63	0	68
67	Güneş Krem	Sinek Öl.	,031	0	0	78
68	Çikolata	Çamaşır El	,028	66	0	69
69	Çikolata	Bebek Tem.	,027	68	0	71
70	Kitap	DVD-VCD	,027	65	59	74
71	Çikolata	Su-Maden su.	,026	69	60	72
72	Çikolata	Dondurma	,026	71	0	73
73	Çikolata	Pekmez-Helva	,024	72	28	74
74	Çikolata	Kitap	,017	73	70	76
75	Enerji İç.	Kuruyemiş	,013	0	58	80
76	Çikolata	Prezervatif	,013	74	0	77
77	Çikolata	Araba Bakım	,012	76	0	78
78	Çikolata	Güneş Kremi	,011	77	67	79
79	Çikolata	evcil hayvan	,009	78	0	80
80	Çikolata	Enerji iç.	,008	79	75	82
81	Sigara	Dergi	-,010	0	64	82
82	Çikolata	Sigara	-,037	80	81	0

Tablo 53 incelendiğinde kümeleme işleminin 82 aşama sürdüğü görülmektedir. İlk aşamada aralarındaki ilişki en kuvvetli olan “makarna-bakliyat” çiftinin ilk kümeyi oluşturduğu görülmektedir. Sonraki aşama sütunundan görüleceği gibi, daha sonra bu kümeye 6. aşamada baharat eklenmektedir. 6. aşamadaki 0,284 değeri baharat sınıfının makarna veya bakliyat sınıfıyla değil, makarna-bakliyat kümesiyle arasındaki ilişki kuvvetini göstermektedir. Daha sonra ise 9. aşamada salça-konserve ürün sınıfı bu kümeye eklenmektedir.

Kümeleme işlemi sırasında oluşturulan toplu tablo, dendogram grafiği üzerinde daha rahat incelenebilir. Dendogram, toplu tablodaki ilişkilerin kuvvetini çizgilerle gösteren bir grafikdir. Analiz sonucunda oluşan EK 2’deki dendogramın ilk bölümü aşağıda açıklanmıştır.



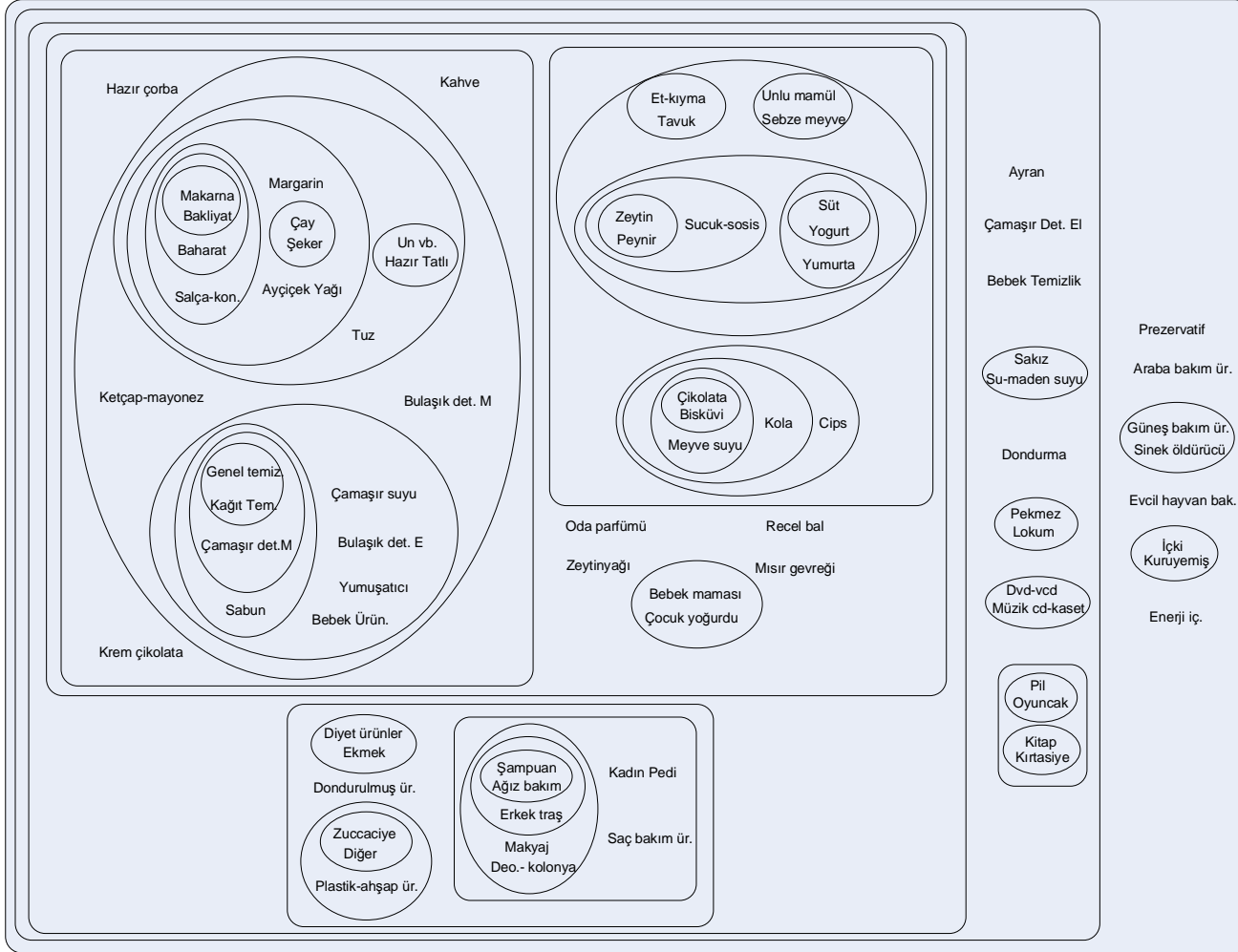
Şekil 31. Ürün Sınıfları İlişkilerini Gösteren Dendogram

Şekil 31 incelendiğinde ürün sınıfları aralarındaki ilişkiler ve oluşturdukları kümeler görülebilmektedir. Makarna ve bakliyat sınıfları arasındaki kuvvetli ilişki bu iki ürün sınıfının ilk önce birleşmesini ve ilk kümeyi oluşturmasını sağlamıştır. Birleşme noktası sağa kaydıçça küme içindeki ilişkinin kuvvetinin azaldığı söylenebilir. Yukarıdaki şekilde 1. kümeyi “makarna, bakliyat, baharat ve salça-konserve” ürün sınıflarının oluşturduğu, 2. kümeyi de “çay, şeker, margarin ve ayçicek yağı”nın

oluşturduğu söylenebilir. Veya daha genel bir küme oluşturularak “makarna, bakliyat, baharat, salça-konserve, çay, şeker, margarin, ayçiçek yağı, tuz, pasta malzemeleri ve hazır tatlı”nın bir küme oluşturduğu yorumu yapılabilir. Ürün sınıfları arasındaki ikili ilişki kuvvetleri EK 1’de verilen yakınlık matrisinde, kümeler arasındaki ilişkiler de tablo 53’deki “toplu tablo”da görülebilir.

Dendogramın diğer kısımları incelendiğinde tüm ürün sınıflarının en sonda tek bir kümede toplandığı görülebilir. Başka bir deyişle dendogram soldan sağa doğru incelendiğinde elde edilen küme sayısı arttırılabilir. Dendogramdaki ilişkiler küme grafiği halinde çizildiğinde Şekil 32 elde edilmektedir. Bu grafikte ürün sınıfları arasındaki sayısal ilişkiler, grafik çok karışacağı için gösterilmemiştir. Sayısal ilişkiler tablo 53’de verilen toplu tablodan ve Ek 1’deki yakınlık matrisinden elde edilebilir.

Şekil 32. Analiz sonucunda ürün sınıflarının oluşturduğu küme grafiği





## 2.5.6. Ürün Sınıfları Arasındaki Birliktelik Kurallarının Belirlenmesi

Ürün sınıfları arasında kümeleme analizi yapıldıktan sonra, ürünlerin birbirleri ile ikili ilişkilerini inceleyebilmek amacıyla birliktelik kuralları analizi gerçekleştirilmiştir. Analiz sonucunda elde edilen kuralların, kural desteği sıralamasına göre ilk 10'u Tablo 54'de görülmektedir.

Tablo 54. Ürün Sınıfları Arasındaki Birliktelik Kuralları (Kural Desteğine Göre Sıralı)

Ürün Sınıfı (Öncül)	Ürün Sınıfları (Artçıl)	Kural Desteği %	Güven %	Kaldırma Oranı (İlgi)
Bisküvi, kek	Çikolata, şeker	7,24	45,12	2,34
Çikolata, şeker	Bisküvi, kek	7,24	37,61	2,34
Çikolata, şeker	Kola, Gazoz	4,99	25,90	1,70
Bisküvi, kek	Kola, Gazoz	4,60	28,67	1,88
Peynir	Unlu Mamul-hazır yemek	3,77	45,08	2,52
Çikolata, şeker	Meyve Suyu	3,62	18,83	1,96
Bisküvi, kek	Meyve Suyu	6,58	22,34	2,33
Çikolata, şeker	Süt	3,42	17,77	2,32
Bisküvi, kek	Süt	3,07	19,16	2,50
Peynir	Sebze meyve	2,46	29,43	3,27

Tablo 54 incelendiğinde en yüksek kural desteğinin “Bisküvi, kek⇒Çikolata, şeker” kuralında olduğu görülmektedir. Alışveriş yapanların %7,24’ü hem bisküvi, kek sınıfından, hem de çikolata, şeker sınıfından ürün almışlardır.

Analiz sonucunda elde edilen kuralların, güven değerine göre sıralanmış ilk 10'u Tablo 55'de görülmektedir.

Tablo 55. Ürün Sınıfları Arasındaki Birliktelik Kuralları (Güven Değerine Göre Sıralı)

Ürün Sınıfı (Öncül)	Ürün Sınıfları (Artçıl)	Kural Desteği %	Güven %	Kaldırma Oranı (İlgi)
Zeytin	Peynir	2,04	61,16	7,30
Bakliyat	Makarna	1,45	52,64	12,35
Bisküvi, kek	Çikolata, şeker	7,24	45,12	2,34
Peynir	Unlu Mamul-hazır yemek	3,77	45,08	2,52
Şeker	Çay	1,15	42,48	11,17
Baharat	Makarna	1,24	37,63	8,82
Çikolata, şeker	Bisküvi, kek	7,24	37,61	3,34
Yumurta	Sebze meyve	1,12	36,65	4,07
Et-Kıyma	Tavuk	0,76	35,20	8,97
Makarna	Süt	1,48	34,89	4,56

Tablo 55 incelendiğinde en yüksek güven değerinin “Zeytin⇒ Peynir” kuralında olduğu görülmektedir. Zeytin alanların %61’i peynir sınıfından bir ürün de almaktadır. İkinci sıradaki kurala göre de bakliyat sınıfında ürün alanların yarısı makarna sınıfından bir ürün de almaktadır.

Analiz sonucunda elde edilen kuralların, kaldırma oranı (ilgi) değerine göre sıralanmış ilk 10’u Tablo 56’da görülmektedir.

*Tablo 56. Ürün Sınıfları Arasındaki Birliktelik Kuralları (Kaldırma Oranı Değerine Göre Sıralı)*

Ürün Sınıfı (Öncül)	Ürün Sınıfları (Artçıl)	Kural Desteği %	Güven %	Kaldırma Oranı (İlgi)
Hazır tatlı malz.	Un, kab. toz.	0,81	31,63	12,95
Un, kab. toz.	Hazır tatlı malz.	0,81	33,50	12,95
Bakliyat	Makarna	1,45	52,64	12,35
Makarna	Bakliyat	1,45	34,03	12,34
Çam. Deterjan (mak.)	Yumuşatıcı	0,52	13,87	11,22
Şeker	Çay	1,15	42,48	11,17
Çay	Şeker	1,15	30,26	11,17
Makarna	Tuz	0,55	12,95	10,26
Genel Temizlik Mad.	Çamaşır Suyu	0,76	15,07	10,03
Makarna	Ketçap- Mayonez	0,55	12,98	9,71

Tablo 56 incelendiğinde un, kabartma tozu sınıfından ürün almak, hazır tatlı malzemeleri sınıfından ürün alma olasılığını yaklaşık 13 kat arttırmaktadır. Aynı durum ürün sınıfları yer değiştirdiğinde de geçerlidir. Bu kurallara, kural desteği değerinin düşük olmasına nedeniyle fişlerde çok rastlanmamaktadır. Ama gene de bu kurallardan çıkartılacak sonuçlar özellikle raf dizaynı ve promosyon kampanyalarında yararlı olabilir.

Tablo 57’de ise tek tek tüm ürün sınıflarının en çok ilişkide olduğu sınıflar listelenmiştir. 81 ürün sınıfı bulunmaktadır ve her sınıfın diğer 80 sınıfla ilişkisi hesaplanmıştır. Ancak yer darlığından dolayı, her ürün sınıfının en yüksek kural desteğine sahip 3 ürün sınıfıyla ilişkisi verilmiştir.

Tablo 57. Ürün Sınıfları Arasındaki Birliktelik Kuralları Tablosu

Ürün Sınıfı (Öncül)	Ürün Sınıfları (Artçıl)	Kural Destegi %	Güven %	Kaldırma Oranı (İlgi)
Çikolata, Şeker	Bisküvi, kek	7,24	37,61	2,34
	Kola, Gazoz	4,99	25,90	1,70
	Unlu Mamul-hazır yemek	4,58	23,82	1,33
Unlu Mamul-hazır yemek	Çikolata, Şeker	4,58	25,67	1,33
	Bisküvi, kek	3,93	22,01	1,37
	Kola, Gazoz	3,89	21,78	1,42
Bisküvi, kek	Çikolata, Şeker	7,24	45,12	2,34
	Kola, Gazoz	4,60	28,67	1,88
	Unlu Mamul-hazır yemek	3,93	24,50	1,37
Kola, Gazoz	Çikolata, Şeker	4,99	32,75	1,70
	Bisküvi, kek	4,60	30,21	1,88
	Unlu Mamul-hazır yemek	3,89	25,54	1,42
Su, Maden suyu	Çikolata, Şeker	3,10	22,50	1,16
	Kola, Gazoz	2,76	20,2	1,31
	Bisküvi, kek	2,61	18,94	1,17
Meyve Suyu	Çikolata, Şeker	3,62	37,89	1,96
	Bisküvi, kek	3,58	37,47	2,33
	Kola, Gazoz	2,71	28,37	1,86
Sebze Meyve	Unlu Mamul-hazır yemek	3,61	40,26	2,25
	Çikolata, Şeker	2,67	29,72	1,54
	Peynir	2,46	27,42	3,25
Peynir	Unlu Mamul-hazır yemek	3,77	45,08	2,52
	Çikolata, Şeker	3,17	37,87	1,96
	Bisküvi, kek	2,70	32,36	2,01
Sakız	Çikolata, Şeker	2,72	35,08	1,82
	Bisküvi, kek	2,05	26,47	1,64
	Kola, Gazoz	1,65	21,24	1,39
Süt, Katkılı süt	Çikolata, Şeker	3,42	44,79	2,32
	Bisküvi, kek	3,07	40,27	2,50
	Unlu Mamul-hazır yemek	2,53	33,17	1,85
Hazır çorba	Çikolata, Şeker	2,58	34,90	1,81
	Bisküvi, kek	2,30	31,10	1,93
	Kola, Gazoz	2,21	29,92	1,96
Lokum	Unlu Mamul-hazır yemek	1,83	29,47	1,64
	Çikolata, Şeker	1,50	24,11	1,25
	Bisküvi, kek	1,02	16,39	1,02
Kuru yemiş	Çikolata, Şeker	1,48	23,99	1,24
	Bisküvi, kek	1,11	18,02	1,12
	Kola, Gazoz	1,08	17,50	1,14

Tablo 57'nin devamı

Cips	Çikolata, Şeker	2,33	41,30	2,14
	Kola, Gazoz	1,93	34,11	2,23
	Bisküvi, kek	1,89	33,40	2,08
Sucuk-sosis	Unlu Mamul-hazır yemek	2,12	39,35	2,20
	Peynir	2,03	37,78	4,51
	Çikolata, Şeker	1,89	35,12	1,82
Makarna	Çikolata, Şeker	2,15	50,64	2,62
	Bisküvi, kek	2,00	46,97	2,92
	Peynir	1,56	36,63	4,37
Tavuk	Unlu Mamul-hazır yemek	1,68	43,02	2,40
	Çikolata, Şeker	1,45	37,04	1,92
	Peynir	1,28	32,82	3,92
Çay, Bitki çayı	Çikolata, Şeker	1,66	43,90	2,27
	Bisküvi, kek	1,49	39,30	2,44
	Peynir	1,24	32,67	3,90
Yoğurt	Unlu Mamul-hazır yemek	1,55	41,38	2,31
	Peynir	1,52	40,45	4,83
	Çikolata, Şeker	1,52	40,37	2,09
Salça-konserve-sirke	Çikolata, Şeker	1,63	43,96	2,28
	Unlu Mamul-hazır yemek	1,45	39,23	2,19
	Bisküvi, kek	1,44	38,90	2,42
Zeytin	Peynir	2,04	61,16	7,30
	Unlu Mamul-hazır yemek	1,45	43,69	2,44
	Çikolata, Şeker	1,28	38,43	1,99
Baharat – harçlar	Çikolata, Şeker	1,59	48,21	2,50
	Bisküvi, kek	1,42	43,06	2,68
	Makarna	1,24	37,63	8,82
Çocuk Gel. Yoğurt	Çikolata, Şeker	1,68	52,45	2,72
	Bisküvi, kek	1,44	45,12	2,80
	Süt	1,18	36,90	4,82
Pekmez-tahin-helva	Çikolata, Şeker	1,15	37,19	1,93
	Unlu Mamul-hazır yemek	1,13	36,56	2,04
	Bisküvi, kek	0,91	29,46	1,83
Yumurta	Peynir	1,35	44,12	5,27
	Unlu Mamul-hazır yemek	1,35	44,05	2,46
	Çikolata, Şeker	1,27	41,42	2,15
Margarin	Çikolata, Şeker	1,47	50,05	2,59
	Bisküvi, kek	1,35	45,80	2,85
	Peynir	1,32	44,79	5,35
Bakliyat	Makarna	1,45	52,64	12,35
	Çikolata, Şeker	1,36	49,64	2,57
	Bisküvi, kek	1,24	45,11	2,80

Tablo 57'nin devamı

Şeker	Çikolata, Şeker	1,26	46,67	2,42
	Çay	1,15	42,48	11,17
	Bisküvi, kek	1,12	41,54	2,58
Kahve	Çikolata, Şeker	1,27	47,12	2,44
	Bisküvi, kek	1,07	39,83	2,48
	Kola, Gazoz	0,81	30,26	1,98
Hazır Tatlı	Çikolata, Şeker	1,35	52,55	2,72
	Bisküvi, kek	1,25	48,38	3,01
	Süt	0,91	35,27	4,61
Un, nişasta, pasta mal.	Çikolata, Şeker	1,28	52,47	2,72
	Bisküvi, kek	1,19	48,88	3,04
	Süt	0,87	35,99	4,70
Diyabetik-diyet ürünler	Bisküvi, kek	1,15	48,64	3,02
	Çikolata, Şeker	1,00	42,60	2,21
	Unlu Mamul-hazır yemek	0,72	30,57	1,71
Kıyma-et	Unlu Mamul-hazır yemek	1,06	49,28	2,75
	Sebze meyve	0,84	39,01	4,34
	Tavuk	0,76	35,20	8,97
Ayran	Unlu Mamul-hazır yemek	0,89	42,35	2,36
	Çikolata, Şeker	0,64	30,45	1,58
	Bisküvi, kek	0,55	26,22	1,63
Krem çikolata	Çikolata, Şeker	1,14	56,79	2,94
	Bisküvi, kek	1,02	50,71	3,15
	Peynir	0,71	35,53	4,24
Mısır gevreği	Çikolata, Şeker	1,02	53,30	2,76
	Bisküvi, kek	0,86	44,71	2,78
	Süt	0,74	38,82	5,07
Ayçiçek Yağı	Çikolata, Şeker	0,85	47,09	2,44
	Bisküvi, kek	0,76	42,15	2,62
	Makarna	0,74	41,30	9,69
Dondurma	Çikolata, Şeker	0,52	30,31	1,57
	Bisküvi, kek	0,44	25,93	1,61
	Kola	0,43	24,98	1,63
Ekmek	Unlu Mamul-hazır yemek	0,73	47,16	2,63
	Peynir	0,63	40,24	4,80
	Çikolata, Şeker	0,62	39,90	2,07
Bebek maması-yiyeceği	Bisküvi, kek	0,51	35,84	1,23
	Çikolata, Şeker	0,50	35,28	1,83
	Süt	0,42	29,56	3,86
Reçel-bal	Çikolata, Şeker	0,60	45,23	2,34
	Peynir	0,57	42,92	5,12
	Bisküvi, kek	0,51	38,21	2,37
Ketçap-mayonez	Çikolata, Şeker	0,70	53,04	2,75
	Bisküvi, kek	0,64	48,11	2,99
	Makarna	0,55	41,42	9,71

Tablo 57'nin devamı

Tuz	Çikolata, Şeker	0,63	50,52	2,62
	Bisküvi, kek	0,58	46,08	2,86
	Makarna	0,55	43,76	10,26
Dondurulmuş ürünler	Çikolata, Şeker	0,43	48,76	2,53
	Bisküvi, kek	0,37	42,12	2,62
	Unlu Mamul-hazır yemek	0,36	41,41	2,31
Zeytinyağı	Çikolata, Şeker	0,28	46,53	2,41
	Peynir	0,26	43,11	5,15
	Bisküvi, kek	0,24	39,86	2,48
Enerji içeceği	Çikolata, Şeker	0,19	32,38	1,68
	Kola, Gazoz	0,16	27,72	1,81
	Su, Maden suyu	0,13	23,38	1,69
Sigara	Kuruyemiş	0,45	3,81	0,61
	Gazete	0,42	3,49	0,92
	Bisküvi, kek	0,41	3,46	0,21
Diğer	Genel temizlik mal.	1,32	18,11	3,57
	Tuv. kağıdı-peçete-havlu	1,14	15,60	3,02
	Sabun	1,02	13,94	2,82
Bebek malzemeleri	Çikolata, Şeker	2,43	40,17	2,08
	Bisküvi, kek	2,05	33,82	2,10
	Tuv.kağıdı-peçete-havlu	1,55	25,61	4,96
Züccaciye	Çikolata, Şeker	1,85	32,75	1,70
	Unlu Mamul-hazır yemek	1,47	25,90	1,44
	Bisküvi, kek	1,46	25,86	1,61
Tuv.kağıdı-peçete-havlu	Çikolata, Şeker	2,30	44,62	2,31
	Bisküvi, kek	2,01	39,14	2,43
	Genel temizlik mal.	1,63	31,61	6,23
Genel temizlik mal.	Çikolata, Şeker	2,24	44,22	2,29
	Bisküvi, kek	1,92	37,92	2,36
	Tuv.kağıdı-peçete-havlu	1,63	32,16	6,23
Sabun -Duş jeli	Çikolata, Şeker	2,01	40,95	2,12
	Bisküvi, kek	1,74	35,41	2,20
	Tuv.kağıdı-peçete-havlu	1,42	28,98	5,62
Oyuncak	Çikolata, Şeker	1,15	25,00	1,29
	Bisküvi, kek	0,83	17,95	1,11
	Unlu Mamul-hazır yemek	0,76	16,64	0,93
Şampuan	Çikolata, Şeker	1,57	37,74	1,95
	Bisküvi, kek	1,27	30,59	1,90
	Sabun -Duş jeli	1,16	28,01	5,68
Plastik -ahşap malz.	Kola, gazoz	1,45	34,90	2,29
	Unlu Mamul-hazır yemek	1,16	27,94	1,56
	Çikolata, Şeker	1,13	27,20	1,41
Diş macunu -fırça	Çikolata, Şeker	1,38	35,21	1,82
	Bisküvi, kek	1,09	27,67	1,72
	Unlu Mamul-hazır yemek	0,87	22,24	1,24

Tablo 57'nin devamı

Deodorant – krem, Kolonya	Çikolata, Şeker	1,31	33,51	1,73
	Bisküvi, kek	1,00	25,46	1,58
	Kola, gazoz	0,83	21,12	1,38
Gazete	Sigara	0,42	11,14	0,92
	Kuruyemiş	0,14	3,90	0,63
	Dergi	0,14	3,76	2,77
Çamaşır deterjan (makine)	Çikolata, Şeker	1,63	43,48	2,25
	Bisküvi, kek	1,40	37,28	2,32
	Genel temizlik mal.	1,26	33,57	6,62
İçki	Cips	0,59	18,08	3,19
	Kola, gazoz	0,58	17,81	1,16
	Çikolata	0,55	16,77	0,87
Kırtasiye	Çikolata, Şeker	0,93	31,40	1,63
	Bisküvi, kek	0,72	24,24	1,50
	Unlu Mamul-hazır yemek	0,63	21,27	1,19
Kitap	Çikolata, Şeker	0,78	27,89	1,44
	Bisküvi, kek	0,56	20,07	1,25
	Unlu Mamul-hazır yemek	0,51	18,25	1,02
Kadın pedi	Çikolata, Şeker	0,92	40,36	2,09
	Bisküvi, kek	0,76	33,66	2,09
	Kola, gazoz	0,57	25,07	1,64
Makyaj-tüy dökücü	Çikolata, Şeker	0,75	33,48	1,73
	Bisküvi, kek	0,59	26,39	1,64
	Unlu Mamul-hazır yemek	0,50	22,38	1,25
Erkek tıraş ürün.	Çikolata, Şeker	0,72	32,84	1,70
	Bisküvi, kek	0,56	25,47	1,58
	Diş macunu -firça	0,49	22,37	5,67
Müzik cd-kaset	Çikolata, Şeker	0,35	17,16	0,89
	Sakız	0,26	13,03	1,67
	Su, maden suyu	0,26	12,91	0,93
Oto-Araba malz.	Çikolata, Şeker	0,43	22,57	1,17
	Kola, gazoz	0,31	16,37	1,07
	Bisküvi, kek	0,30	16,11	1,00
Oda parfümü	Çikolata, Şeker	0,73	39,42	2,04
	Genel temizlik mal.	0,62	33,30	6,57
	Bisküvi, kek	0,59	31,86	1,98
Saç boyası-jöle	Çikolata, Şeker	0,50	33,10	1,71
	Bisküvi, kek	0,38	25,34	1,57
	Kola, gazoz	0,32	21,78	1,42
Çamaşır suyu	Genel temizlik mal.	0,76	50,89	10,03
	Çikolata, Şeker	0,72	48,32	2,50
	Bisküvi, kek	0,64	43,20	2,69
Bulaşık det. (el)	Genel temizlik mal.	0,69	47,37	9,34
	Çikolata, Şeker	0,66	45,45	2,35
	Tuv.kağıdı-peçete-havlu	0,59	40,80	7,91

Tablo 57'nin devamı

Dergi	Gazete	0,14	10,46	2,77
	Sigara	0,10	8,07	0,67
	Çikolata, Şeker	0,07	5,14	0,26
Bulaşık det. (makine)	Çikolata, Şeker	0,62	48,47	2,51
	Genel temizlik mal.	0,58	45,22	8,92
	Bisküvi, kek	0,51	40,12	2,49
Yumuşatıcı	Çikolata, Şeker	0,57	46,76	2,42
	Genel temizlik mal.	0,56	45,62	9,00
	Çamaşır deterjan (makine)	0,52	42,17	11,22
DVD-vcd	Çikolata, Şeker	0,26	27,28	1,41
	Bisküvi, kek	0,18	19,57	1,21
	Unlu Mamul-hazır yemek	0,17	18,72	1,04
Bebek temizlik mal.	Çikolata, Şeker	0,24	36,17	1,87
	Bisküvi, kek	0,22	32,76	2,04
	Bebek malzemeleri	0,22	32,46	5,35
Pil	Çikolata, Şeker	0,21	31,44	1,63
	Bisküvi, kek	0,16	24,44	1,52
	Unlu Mamul-hazır yemek	0,14	21,39	1,19
Prezervatif	Çikolata, Şeker	0,13	28,66	1,48
	Bisküvi, kek	0,10	22,29	1,38
	Unlu Mamul-hazır yemek	0,08	18,41	1,03
Güneş sütü- kremi	Sakız	0,09	24,83	2,80
	Su-maden suyu	0,08	17,75	1,28
	Oyuncak	0,07	15,75	3,40
Sinek öldürücü	Çikolata, Şeker	0,11	30,64	1,59
	Bisküvi, kek	0,09	26,01	1,62
	Kola, gazoz	0,09	24,75	1,62
Evcil hayvan	Çikolata, Şeker	0,05	26,66	1,38
	Unlu Mamul-hazır yemek	0,04	24,24	1,35
	Peynir	0,04	20,79	2,48
Çamaşır deterjan (el)	Çikolata, Şeker	0,07	41,67	2,16
	Genel temizlik mal.	0,07	38,00	7,49
	Bisküvi, kek	0,07	37,56	2,33

Tablo 57'de kural desteği öncül ve artçıl ürün sınıflarının ikisinin birden aynı alışveriş sırasında alınma oranını, güven öncül sınıftan ürün alan müşterilerin yüzde kaçının artçıl sınıftan da bir ürün aldığını, kaldırma oranı (ilgi) ise öncül sınıftan bir ürün almanın, artçıl sınıftan bir ürün alma oranını ne kadar arttırdığını göstermektedir. Tablo 57'de çikolata-şeker, bisküvi-kek ve unlu mamul sınıflarının çoğu sınıfla kuvvetli bir şekilde ilişkide olduğu görülmektedir. Temizlik ve kişisel bakım ürünlerinin kendi aralarındaki yüksek kaldırma oranları da ilgi çekicidir.



## SONUÇ

Çalışmada perakende sektöründe ürün kategorileri ve ürün sınıfları arasındaki birlikte satış ilişkilerinin ve kümelerin belirlenmesi amaçlamıştır. Aynı zamanda bu araştırma perakende sektöründe faaliyet gösteren bir işletme tarafından depolanan alışveriş detaylarıyla ilgili verilerden anlamlı bilgiler elde edilmesi ve bu bilgiler ışığında pazarlama faaliyetlerinin düzenlenmesine ilişkin önerilerde bulunulması konusunda bir örnek oluşturacaktır. İşletme satış terminalleri aracılığıyla topladığı satışlara ait verileri (alışveriş detayları: fiş no, fiş tarihi, saat, alınan ürünler, vb.) bir süredir depolamamdadır, ama kullandıkları yazılım hala geliştirilmekte olduğu ve özellikleri sınırlı olduğu için bu verilerden etkili bir şekilde yararlanamamaktadır. Bu araştırma sonucunda veritabanının içinde saklı olan birçok değerli bilginin keşfedilmesi ile işletme hem şimdiye kadar yaptığı yatırımların geri dönüşünü sağlayacak, hem de günümüzün rekabetçi ortamında kendi sektöründe diğer işletmelerin bir adım önüne geçecektir.

Bu araştırmanın amacı ürün kategorileri ve ürün sınıfları arasındaki ilişkileri ortaya çıkarmaktır. Bu amaçla bir perakendeci işletmenin, alışveriş detaylarını içeren verileri veri madenciliği yöntemleri (birliktelik kuralları: GRI ve Apriori, hiyerarşik kümeleme analizi) ile incelenmiştir. Bulunan ilişki kuralları ve kümeler perakendeci işletmeler açısından kategori yönetimi, tanzim-teşhir konularında yeni düzenlemelerin bilimsel sonuçlar ışığında yapılmasını sağlayacaktır.

Araştırma kapsamında işletmeden alınan veriler aylara göre alışveriş detaylarını yani tarih, kasa no, fiş no, barkod no, ürün ismi, adedi, birim fiyatı, KDV'si, toplam fiyatı gibi bilgilerden oluşmaktadır ve öncelikle veri madenciliği yazılımında kullanılabilecek hale dönüştürülmüştür. Düzenlenen verilerde toplam 1.096.125 adet fiş içerisinde, 6.009.317 alışveriş hareketi olduğu görülmüştür.

Özellikle Temmuz-Ağustos aylarında kesilen fiş adedinin arttığı görülmektedir. Ortalama bir fişteki hareket adedi 5,48'dir.

Dönüştürülen veriler SPSS Clementine yazılımında veri madenciliği yöntemleriyle analiz edilmiştir. Araştırmada esas olarak birliktelik kuralları analizi ve kümeleme analizi kullanılmıştır. Araştırma sonunda elde edilen bulgulardan dikkat çekici olanlar şunlardır:

İlk olarak gerçekleştirilen birliktelik kuralları analizi sonucunda en büyük kural desteğinin %18,69 ile içecekler⇒tatlı ürünler bağıntısında olduğu belirlenmiştir. Yani fişlerin yaklaşık %19'unda içecekler ve tatlı ürünler kategorilerindeki ürünler birlikte yer almaktadır. Tatlı ürünler kategorisindeki bir ürünü alan müşterilerin yarıdan fazlası (%52,04) içecekler kategorisindeki bir ürünü de alırken, içecekler kategorisinden ürün alan müşterilerin üçte birinden biraz fazlası (%36,37) tatlı ürünler kategorisinden bir ürün almaktadır.

Analiz sonuçları güven değerine göre sıralandığında ise aşağıdaki sonuçlara ulaşılmıştır. Market markası kategorisinden ürün alan müşterilerin yaklaşık %70'i aynı zamanda tatlı ürünler kategorisinden de ürün almaktadır. Bu bağıntıya fişlerin yaklaşık %5'inde rastlanmaktadır. Market markası⇒kuru gıda, açık şarküteri⇒tatlı ürünler, market markası⇒içecekler, temizlik⇒tatlı ürünler, etler⇒açık şarküteri, market markası⇒temizlik, içecekler ⇒tatlı ürünler, etler⇒kuru gıda bağıntılarının hepsinde de %50'nin üzerinde güven değeri tespit edilmiştir. Yani ilk kategoriden bir ürün alan müşterilerin yarıdan fazlası, ikinci kategoriden de ürün almaktadır.

Bulunan bağıntıların istatistiksel olarak anlamlı olup olmadığının belirlenmesi amacıyla ki-kare testi uygulanmıştır. Birliktelik kuralları bulunan tüm kategori türleri için ki-kare değerinin 6,635'den yüksek olduğu görülmektedir. Tüm kurallarda %1 anlamlılık derecesinde bir ilişki vardır. Dolayısıyla kategorilerin birbirine bağımlı olduğu söylenebilir.

Daha sonra bulunan bağıntılar için ilgi değerleri tabloları hazırlanmıştır. İçecekler kategorisini içeren bağıntılar haricindeki tüm bağıntılarda ilgi değerlerinin 1'in üstünde olduğu belirlenmiştir. Yani bulunan bağıntılarda kategorilerden birinin satılması diğerinin satılması ihtimalini arttırmaktadır.

Tespit edilen bağıntıların hepsi iki kategori arasında olduğu için, analiz ayarları değiştirilerek daha çok kategorinin işin içine girdiği bağıntılar tespit edilmiştir. Bu analiz sonucunda en yüksek kural desteğinin (%8,13) içecekler, kuru gıda ve tatlı ürünler kategorileri arasında olduğu belirlenmiştir. Fişlerin %8’inde bu bağıntıya rastlanmaktadır. Analiz sonuçları güven değerine göre sıralandığında “Market markaları& Açık şarküteri⇒Kuru gıda” bağıntısındaki güven değerinin en yüksek olduğu görülmektedir. Market markaları ve açık şarküteri kategorilerinden ürün alan bir müşteri %80 ihtimalle kuru gıda kategorisinden bir ürün de almaktadır. Ama bu bağıntıya fişlerin sadece %2,85’inde rastlanmaktadır.

Tüm kategorilere ait genel bir değerlendirme yapıldıktan sonra, tek tek kategori bazında değerlendirmeler yapılmıştır. 14 kategorinin teker teker diğer kategorilerle ilişkisi incelenmiş ve kategoriler arasındaki ilişkiler de tespit edilmiştir.

Kategoriler arasındaki ilişkiler belirlendikten sonra, alışverişler sırasında müşterilerin aldığı ürünlerin çoklu ilişkilerinin tespit edilebilmesi amacıyla hiyerarşik kümeleme analizi gerçekleştirilmiştir. Öncelikle ürünler toplam 83 sınıfa ayrılmış ve bu sınıfların her fişle dolayısıyla her alışveriş kaydıyla ilişkilendirilmesi sağlanmıştır. Fişlerde en sık görülen ürün sınıfları çikolata-şeker (%19,3), Unlu Mamul-hazır yemek (%17,9), Bisküvi-kek (%16,1), Kola-Gazoz (%15,2), Su-Maden suyu (%13,8), Meyve suyu (%9,6), sebze-meyve (%9,0)’dir. Gıda dışı ürünlerde ise en sık rastlanan ürün sınıfları sigara (%12), Bebek malzemeleri (%6,1), züccaciye (%5,7), Kâğıt temizlik ürünleri (%5,2) ve genel temizlik maddeleridir (%5,1).

Araştırmanın ikinci kısmında gerçekleştirilen kümeleme analizi sonucunda en kuvvetli ilişkiler, toplu tabloda da kümeleme işlemleri sırasında ilk olarak aynı kümelere yerleştirilen ürün çiftleri arasında görülmektedir. Makarna-bakliyat (0,403), zeytin-peynir (0,354), çay-şeker (0,337), pasta malzemeleri (un, kabartma tozu, vb.)-hazır tatlı (puding, vb.) (0,308) bu ürün çiftleri arasında gösterilebilir.

Dendogram sonuçlarında ilk göze çarpan, makarna ve bakliyat sınıfları arasındaki kuvvetli ilişki sebebiyle bu iki ürün sınıfının ilk önce birleşmesi ve ilk kümeyi oluşturmasıdır. Daha genel bir bakışla 1. kümeyi “makarna, bakliyat, baharat ve salça-konserve” ürün sınıflarının oluşturduğu, 2. kümenin de “çay, şeker, margarin ve ayçiçek yağı”ndan oluştuğu söylenebilir. Veya daha genel bir küme

oluşturularak “makarna, bakliyat, baharat, salça-konserve, çay, şeker, margarin, ayçiçek yağı, tuz, pasta malzemeleri ve hazır tatlı”nın bir küme oluşturduğu yorumu yapılabilir. Diğer küme grupları dendogramda görülebilir.

Günümüzün artan rekabet şartları sonucu işletmeler kârlılıklarını arttırabilmek için, maliyetleri düşürmek, yeni pazarlar yaratmak ve stratejik kararlar almak zorundadır. Doğru kararlar alabilmek için öncelikle karar alınmasını sağlayacak verilerin tutarlı ve doğru olması gerekmektedir. İşletmeler, rekabette üstünlük sağlayabilecek, işletme ile ilgili faaliyetler sonucu ortaya çıkan verileri, teknolojik gelişmelerin sayesinde, birçok teknoloji yatırım gerçekleştirerek biriktirmektedirler. Ama bunun sonucunda ortaya çıkan veri fazlalığı, karar alırken yarardan çok zarar getirebilmekte ve yapılan teknoloji yatırımlarının geri dönmemesine sebep olmaktadır. Karar verici kişiler veri fazlalığı ve verilerin işlenip bilgiye dönüştürülememesi sebebiyle karar alırken verilerden yararlanamamaktadır.

Günümüzde rekabetin arttığı perakende sektöründe, artık perakendeci işletme kendine müşteri hazırlamak yerine, kendini müşteriye hazırlayabilmelidir. Müşteri odaklı olma günümüzde her işletmenin ana hedeflerinden biri olmaktadır. Müşterinin alışveriş alışkanlıklarının, örüntülerinin bilinmesi market yönetimi sırasında alınacak kararlara yön verebilir.

Bu araştırmanın sonuçları perakendeci işletmeler tarafından çeşitli promosyon kampanyalarında değerlendirilebilir. Birbiriyle kuvvetli satış ilişkisi bulunan ya da aynı küme içerisinde bulunan ürün sınıfları aynı promosyonda birleştirilebilir. Aynı zamanda market yerleşimi ve raf dizaynlarında da bu sonuçlardan yararlanılabilecektir. Mağaza içi planlama yapılırken birbiriyle ilişkili sınıflar uygun şekillerde yerleştirilebilir. İşletmenin kategori yönetiminde kullandığı ürün kategorileri ve sınıfları kümeleme analizi sonucunda elde edilen kümelere göre tekrar oluşturularak daha verimli bir şekilde kullanılabilir. Aynı zamanda bulunan örüntüler, kararsız müşteriler için ikna edici, yön gösterici çeşitli faaliyetler için de kullanılabilir.

Veri madenciliği analizi ile eldeki büyük miktardaki veri analiz edilip, daha önce bilinmeyen bilgiler elde edilebilir. Ama bu çalışmada da olduğu gibi, bu konudaki en büyük sorun işletmelerin veri depolama sistemine entegre bir veri

madenciliği yazılımı olmamasıdır. Analistler öncelikle verinin dönüşümünü ve temizlenmesi yapmalı, daha sonra analizleri gerçekleştirmelidir. Bu sebeple yazılım şirketlerinin ya mevcut veri madenciliği yazılımlarına kolayca aktarabilecek veri sistemi tasarımları ya da kendi yazılımlarına temel veri madenciliği modüllerini entegre etmeleri uygun olacaktır. Başarılı bir veri madenciliği sisteminde ham veriden, raporlama sürecine kadar kesintisiz, sürekli bir akış sağlanmalı, veri madenciliği sistemi düzenli aralıklarla, yeni gelen verileri de hesaba katarak raporlar üretebilmelidir.

İleriki araştırmalarda zaman aralığının daha geniş tutularak daha çok veri kullanılması, dönemlere göre (aylar, mevsimler, tatil günleri gibi) incelemeler yapılması, değişik modeller uygulanarak daha çeşitli bilgilerin elde edilmesi düşünülmektedir.

Daha değişik veri madenciliği analizleri yapabilmek için daha detaylı müşteri verilerine (gelir, meslek, tahsil, medeni hal, çocuk sayısı, vb.) ihtiyaç duyulmaktadır. Gelecekte işletme müşteri kartı- sadakat kartı gibi uygulamalar sayesinde elde edilecek demografik bilgiler ve alışveriş hareketleri ilişkilendirilip farklı veri madenciliği uygulamaları geliştirilebilir.

## KAYNAKÇA

- ACNielsen. (1992). *Category Management- Positioning Your Organization to Win*. Chicago, NTC Business Books, USA.
- ACNielsen, Karolefski, J. ve Heller A. (2006). *Consumer-Centric Category Management: How to increase Profits by Managing Categories on Consumer Needs*. New Jersey, USA: John Wiley & Sons Inc.
- Adriaans, P. ve Zantinge D. (1996). *Data Mining*, USA: Addison-Wesley Professional.
- Agrawal, R., Imelinski, T. ve Swami, A. (1993). Mining Association Rules Between Sets Of Items In Large Databases, *Proceedings of the ACM SIGMOD*, 207-216.
- Agrawal, R. ve Srikant, R. (1994). Fast Algorithms For Mining Association Rules In Large Databases, *Proceedings of the 20th International Conference Very Large Data Bases*, 478-499.
- Ahn, J.S. ve Sohn S.Y. (2008). Customer Pattern Search For After-Sales Service In Manufacturing, *Expert Systems with Applications*, doi:10.1016/j.eswa.2008.06.061.
- Akan, A. (2007). *Türk Perakende Sektörünün Değişimi ve Ekonomi Üzerindeki Etkileri*, İstanbul: Pricewaterhouse.
- Akbulut, S. (2006). *Veri Madenciliği Teknikleri İle Bir Kozmetik Markanın Ayrılan Müşteri Analizi ve Müşteri Segmentasyonu*. (Yayınlanmamış Yüksek Lisans Tezi). Gazi Üniversitesi Fen Bilimleri Enstitüsü, Ankara.
- Akpınar, H. (2000). Veri Tabanlarında Bilgi Keşfi ve Veri Madenciliği. *İstanbul Üniversitesi İşletme Fakültesi Dergisi*, c: 29, 1, 1-22.
- Aksulu, İ. (2002). *Dünyada ve Türkiye’de Perakendecilik ve Özel Markalar*, İzmir: İlkem Matbaası.
- Aldana, W.A. (2000). *Data Mining Industry: Emerging Trends and New Opportunities*. (Yayınlanmamış Yüksek Lisans Tezi). Massachusetts Institute of Technology, USA.
- Alışveriş Merkezleri ve Perakendeciler Derneği (AMPD) Web Sitesi. (Ocak 2008). Türkiye Perakende Sektörü. 10.04.2008, <http://www.ampd.org/arastirmalar/default.aspx?SectionId=97> ()

- Alpaydın, E. (2000). *Zeki Veri Madenciliği: Ham Veriden Altın Bilgiye Ulaşma Yöntemleri*. Bilişim 2000 Eğitim Semineri. 15.01.2007, [www.cmpe.boun.edu.tr/~ethem/files/papers/veri-maden\\_2knotlar.doc](http://www.cmpe.boun.edu.tr/~ethem/files/papers/veri-maden_2knotlar.doc).
- Altıntaş, T. (2006). *Veri Madenciliği Metotlarından Olan Kümeleme Algoritmalarının Uygulamalı Etkinlik Analizi*. (Yayınlanmamış Yüksek Lisans Tezi). Sakarya Üniversitesi, Fen Bilimleri Enstitüsü, Sakarya.
- Arslan, F.M. ve Bayçu, S. (2006). *Mağaza Atmosferi*. Eskişehir: Anadolu Üniversitesi Yayınları.
- Aydoğan, F. (2003). *E-Ticarette Veri Madenciliği Yaklaşımlarıyla Müşteriye Hizmet Sunan Akıllı Modüllerin Tasarımı ve Gerçekleştirilmesi*. (Yayınlanmamış Yüksek Lisans Tezi). Hacettepe Üniversitesi Fen Bilimleri Enstitüsü, Ankara.
- Aydın, K. (2005). *Perakende Yönetiminin Temelleri*. Birinci Basım, Ankara.
- Aydın, K. (12 Şubat 2007). *Kategori Yönetimi*. 10.05.2008, <http://perakende.org/haber.php?hid=1197544339>.
- Aydın, K. ve Candan, B. (2007). *Kategori Yönetimi*, İstanbul: Beta Basım Yayım.
- Barutçugil, İ. (25 Temmuz 2006). *Perakende Sektöründe Rekabet (I)*. Erişim: 25.03.2008, <http://perakende.org/haber.php?hid=1197528292>
- Baykal, A. (2006). Veri Madenciliği Uygulama Alanları, *Dicle Üniv. Ziya Gökalp Eğitim Fakültesi Dergisi*, Sayı 7.
- Baykal, N. (2003). *Veri Tabanı ve Veri Madenciliği*. 10.10.2006, [http://www.turkmia.org/eski/file/231verimadenciligi\\_baykal.ppt](http://www.turkmia.org/eski/file/231verimadenciligi_baykal.ppt).
- Bayram, E. (2001). *Customer Segmentation And Churn Modeling In Wireless Communication*. (Yayınlanmamış Yüksek Lisans Tezi). Boğaziçi Üniversitesi Fen Bilimleri Enstitüsü, İstanbul.
- Beksırgakızı, N. (2001). *Perakendecilikte Kategori Yönetimi ve Gıda Sektöründe Bir Uygulama*. (Yayınlanmamış Yüksek Lisans Tezi). Marmara Üniversitesi Sosyal Bilimler Enstitüsü, İstanbul.
- Bellazi, R., Larizza, C. ve Magni P. (2005). Temporal Data Mining For The Quality Assessment Of Hemodialysis Services, *Artificial Intelligence in Medicine*, 34 (1): 25-39.
- Bennett, P.D. (1995). *Dictionary of Marketing Terms*, Second Edition, Chicago, USA: American Marketing Association.
- Berman, B. (1996). *Marketing Channels*. Canada: John Wiley & Sons Inc.

- Berman, B. ve Evans, J.R. (2001). *Retail management: A Strategic Approach*. USA: Prentice Hall.
- Berry, M. ve Linoff, G. S. (2004). *Data Mining Techniques : For Marketing, Sales, And Customer Relationship Management*. Indiana, USA: Wiley Publishing.
- Boginski, V., Butenko, S. ve Pardalos, P. (2006). Mining Market Data: A Network Approach, *Computers & Operations Research*, 33 (11): 3171-3184.
- Bramer, M. (2007). *Principles of Data Mining*. Springer-Verlag, London, UK.
- Bransten, L. (1999). Technology– Power Tools – Looking For Patterns: Data Mining Enables Companies To Better Manage The Ream Of Statistics They Collect; The Goal: Spot The Unexpected, *Wall Street Journal*, 27 (12): 16- 20.
- Brin, S., Motwani, R., Silverstein, C. (1997). Beyond Market Baskets: Generalizing Association Rules to Correlations, *SIGMOD 1997, Proceedings ACM SIGMOD International Conference on Management of Data*, Mayıs 13-15, 1997, Tucson, Arizona, USA
- Brijs, T.. (2002). *Retail Market Basket Analysis: A Quantitative Modelling Approach*, (Yayınlanmamış Doktora Tezi). Limburg University Center, Faculty of Applied Economics, Belgium.
- Brijs, T., Swinnen, G., Vanhoof, K., Wets, G. (1999). Using Association Rules for Product Assortment Decisions: A Case Study, *KDD-99*, San Diego, USA.
- Canitez, E. C. (2007). *Veri Madenciliği ve Veri Ambarlarının Perakendecilik Sektöründe Uygulamaları*. (Yayınlanmamış Yüksek Lisans Tezi). Dokuz Eylül Üniversitesi Sosyal Bilimler Enstitüsü, İzmir.
- Caskey, K. (2001). A Manufacturing Problem Solving Environment Combining Evaluation, Search, And Generalisation Methods, *Computers in Industry*, 44 (2): 175-187.
- Cavique, L. (2007). A scalable Algorithm For The Market Basket Analysis, *Journal of Retailing and Consumer Services*, 14, 400-407.
- Ceran, G. (2006). *Esnek Akis Tipi Çizelgeleme Problemlerinin Veri Madenciligi Ve Genetik Algoritma Kullanılarak Çözülmesi*. (Yayınlanmamış Yüksek Lisans Tezi). Selçuk Üniversitesi Fen Bilimleri Enstitüsü, Konya.
- Cervone, G., Kafatos, M., Napoletani, D. ve Singh R. (2006). An early Warning System For Coastal Earthquakes, *Advances in Space Research*, 37 (4): 636-642.



- Chae, Y. M., Kima, H. S., Tarkb, K. C., Parkb, H. J. ve Ho, S. H. (2003). Analysis Of Healthcare Quality Indicator Using Data Mining And Decision Support System. *Expert Systems With Applications*, 24, 167–172.
- Chang, H.J., Hung, L.P. ve Ho C.L. (2008). An Anticipation Model Of Potential Customers' Purchasing Behavior Based On Clustering Analysis And Association Rules Analysis, *Expert Systems with Applications*, 32, 753–764.
- Chen, M., Huang, C., Chen, K. ve Wu H. (2005). Aggregation Of Orders In Distribution Centers Using Data Mining, *Expert Systems with Applications*, 28 (3): 453-460.
- Chen, Y.L., Tang, K., Shen, R.J. ve Hu Y.H. (2005). Market Basket Analysis In A Multiple Store Environment. *Decision Support Systems*, 40, 339– 354.
- Chiang, D., Keh, H., Huang, H. ve Chyr D. (2008). The Chinese text Categorization System With Association Rule And Category Priority, *Expert Systems with Applications*, 35, 102–110.
- Chien, C.F. ve Chen, L.F. (2008). Data Mining To Improve Personnel Selection And Enhance Human Capital: A Case Study In High-Technology Industry, *Expert Systems with Applications*, 34, 280–290.
- Chiu, S. ve Tavella, D. (2008). *Data Mining and Market Intelligence for Optimal Marketing Returns*, USA: Butterworth-Heinemann, Elsevier.
- Cox, I., Lewis, R., Ransing, R., Laszczewski, H. ve Berni G. (2002). Application Of Neural Computing In Basic Oxygen Steelmaking, *Journal of Materials Processing Technology*, 120 (1): 310-315.
- Crespo, F. ve Weber R. (2005). A Methodology For Dynamic Data Mining Based On Fuzzy Clustering, *Fuzzy Sets and Systems*, 150 (2): 267-284.
- Çerkez, H.S. (2003). *Müşteri İlişkileri Yönetiminde İş Zekası ve Veri Madenciliği Yöntemleri*. (Yayınlanmamış Yüksek Lisans Tezi). İstanbul Teknik Üniversitesi Fen Bilimleri Enstitüsü, İstanbul.
- Dal, B. (03.05.2006). *Geleceğin Devleri: Bilimselleşen Perakendeciler*, 10.10.2007, <http://turk.internet.com/haber/yazigoster.php3?yaziid=15422>.
- Davis, B. (1999). Data mining transformed, *Information Week*, 751: 86 .
- Demir, H. (2007). *Perakende Kuruluşlarda Yeni Teknolojiler Ve Gıda Perakendeciliğinde Kullanımına İlişkin Bir Araştırma*, (Yayınlanmamış Yüksek Lisans Tezi). Anadolu Üniversitesi Sosyal Bilimler Enstitüsü, Eskişehir.

- Doğan, B. (2008). *Bankaların Gözetiminde Bir Araç Olarak Kümeleme Analizi: Türk Bankacılık Sektörü İçin Bir Uygulama*, (Yayınlanmamış Doktora Tezi). Kadir Has Üniversitesi Sosyal Bilimleri Enstitüsü, İstanbul.
- Dolgun, M. Ö. (2006). *Büyük Alışveriş Merkezleri İçin Veri Madenciliği Uygulamaları*, (Yayınlanmamış Yüksek Lisans Tezi). Hacettepe Üniversitesi Fen Bilimleri Enstitüsü, Ankara.
- DuMouchel, W. (1999). Bayesian Data Mining In Large Frequency Tables, With An Application To The Fda Spontaneous, *American Statistician*, 53 (3): 177.
- Eker, H. (2005). *Veri Madenciliği Veya Bilgi Keşfi*. 10.05.2008, [http://www.bilgiyonetimi.org/cm/pages/mkl\\_gos.php?nt=538](http://www.bilgiyonetimi.org/cm/pages/mkl_gos.php?nt=538).
- Facca, F. ve Lanzi, P. (2005). Mining interesting Knowledge From Weblogs: A Survey, *Data & Knowledge Engineering*, 53 (3): 225-241.
- Fernandez, I. B., Zanakis, S.H. ve Walczak, S. (2002). Knowledge Discovery Techniques for Predicting Country Investment Risk, *Computers&Industrial Engineering*, No: 43, 787-800.
- Ganti, V., Gehrke, J. ve Ramakrishnan, R. (1999). Mining Very Large Databases. *IEEE Computer*, 32: 8, 38-45.
- Godin, S. (1999). Permission Marketing: Turning Strangers Into Friends and Friends Into Customers. *Simon and Shuster*, New York, 40-45.
- Goebel, M. ve Gruenwald, L. (1999). A Survey of Data Mining and Knowledge Discovery Software Tools, *ACM SIGKDD Explorations Newsletter*, 1, 1, 20-33.
- Guha, S., Rastogi, R. ve Shim, K. (2000). Rock: A robust Clustering Algorithm For Categorical Attributes, *Information Systems*, 25 (5): 345-366.
- Güllü, K. (2005). *Türk Gıda Perakendecilerinin Uluslararasılaşması Ve Bir Uygulama*, (Yayınlanmamış Doktora Tezi). Erciyes Üniversitesi Sosyal Bilimler Enstitüsü, Kayseri.
- Gürman, A.A. (2006). *Dünya Perakendeciliğinde Globalleşme Eğilimleri ve Türk Perakendecilik Sektörüne Etkileri*. (Yayınlanmamış Yüksek Lisans Tezi). Dokuz Eylül Üniversitesi Sosyal Bilimler Enstitüsü, İzmir.
- Hair, J.F., Black, W.C. ve Barry, B.J. *Multivariate Data Analysis*. Pearson, Prentice Hall, New Jersey, USA.
- Han, J., ve Kamber, M. (2001). *Data Mining: Concepts and Techniques*. USA: Morgan Kaufmann.

- Hand, D.J. (1998). Data mining: statistics and more?, *The American Statistician*, 52: 112-118.
- Hatipođlu, Z. (1986). *Pazarlama Yönetimi ve Stratejisi*, Temel Araştırma AŞ.Yayınları, No: 8. İstanbul.
- Hong, G., Park, S., Jang, D. ve Rho H. (2005). An Effective Supplier Selection Method For Constructing A Competitive Supply Relationship, *Expert Systems with Applications*, 28 (4): 629-639.
- Houtsma, M.A.W. ve Swami, A. (1995). Set-Oriented Mining for Association Rules in Relational Databases, *Eleventh International Conference on Data Engineering*, Taipei.
- Hsieh, N.-C. (2004). An Integrated Data Mining And Behavioral Scoring Model For Analyzing Bank Customers, *Expert Systems with Applications*, 27, 623-633.
- Hsu, C. ve Chen Y. (2007). Mining Of Mixed Data With Application To Catalog Marketing, *Expert System with Applications*, 32(1), s. 12-23.
- Hudairy, H. (2004). *Data mining And Decision Making Support In The Governmental Sector*. (Yayınlanmamış Yüksek Lisans Tezi). Faculty of Graduate School of The University of Louisville, Kentucky.
- Hung, S.-Y., Yen, D. C., ve Wang, H.-Y. (2006). Applying Data Mining To Telecom Churn Management. *Expert Systems with Applications*, 31, 515-524.
- 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 System with Applications*, 26: 181-188.
- İnal, E. (2008). Veri Madenciliđi-OLAP, 15.09.2008, <http://www.bilmuh.gyte.edu.tr/~htakci/vm/olap.doc>
- Jacobs, P. (1999). Data Mining: What general managers need to know, *Harvard Management Update*, 4 (10): .
- Jeng, B., Chen, J. ve Liang, T. (2006). Applying Data Mining To Learn System Dynamics In A Biological Model, *Expert Systems with Applications*, 30 (1): 50-58.
- Jiao, J., Zhang, Y. ve Helander, M. (2006). A Kansei Mining System For Affective Design, *Expert Systems with Applications*, 30 (4): 658-673.
- Karabatak, M. ve İnce, M.C. (2008). An Expert System For Detection Of Breast Cancer Based On Association Rules And Neural Network, *Expert Systems with Applications*, doi:10.1016/j.eswa.2008.02.064

- Karakaş, M. (2005). *Veri Ambarları Genel Yapısı*. 08.10.2006, [http://bilgiyonetimi.org/cm/pages/mkl\\_list.php?id=11](http://bilgiyonetimi.org/cm/pages/mkl_list.php?id=11).
- Kaya, Y. (2003). *Pazar Sepet Analizinde Ürün İlişkilerinin Bulunması*. (Yayınlanmamış Yüksek Lisans Tezi). Gebze Yüksek Teknoloji Enstitüsü, Mühendislik ve Fen Bilimleri Enstitüsü, Gebze.
- Kianmehr, K. ve Alhadj, R. (2008). CAR SVM: A Class Association Rule-Based Classification Framework And Its Application To Gene Expression Data, *Artificial Intelligence in Medicine*, 44, 7-25.
- Kim, E. ve Lee, Y. (2003). Combination Of Multiple Classifiers For The Customer's Purchase Behavior Prediction, *Decision Support Systems*, 34 (2): 167-175.
- Kitler, R. ve Wang, W., 1998, The Emerging Role Of Data Mining. *Solid State Technology*, 42 (11): 45.
- Koç, S. (2001). *Türkiye'de İllerin Sosyo-Ekonomik Özelliklere Göre Sınıflandırılması*, V. Ulusal Ekonometri ve İstatistik Sempozyumu, Çukurova Üniversitesi, Adana.
- Kolay, G. (2006). *İşletmelerde Bilgi Sistemleri Verimliliğini Arttırmada Veri Madenciliği Yöntemi: Bir Simülasyon Çalışması*. (Yayınlanmamış Yüksek Lisans Tezi). Zonguldak Karaelmas Üniversitesi Sosyal Bilimler Enstitüsü, Zonguldak.
- Koyuncu, F. (2004). *Veri Madenciliğinde Apriori Temelli İlişkilendirme Kuralı Algoritmalarının Uygulama ve Karşılaştırılması*. (Yayınlanmamış Yüksek Lisans Tezi). Yıldız Teknik Üniversitesi, Fen Bilimleri Enstitüsü, İstanbul.
- Koyuncugil, A. S. (2006). *Bulanık Veri Madenciliği Ve Sermaye Piyasalarına Uygulanması*. (Yayınlanmamış Doktora Tezi). Ankara Üniversitesi Fen Bilimleri Üniversitesi, Ankara.
- Kumar, A.V.S. ve Wahidabanu R.S.D. (2008). "Data Mining Association Rules for Making Knowledgeable Decisions". H. Rahman (ed.), *Data Mining Applications for Empowering Knowledge Societies*. New York, USA: IGI Global.
- Kuşaksızoğlu, B. (2006). *Fraud Detection In Mobile Communication Networks Using Data Mining*. (Yayınlanmamış Yüksek Lisans Tezi). The University Of Bahcesehir, İstanbul.
- Last, M. ve Kandel, A. (2004). Discovering Useful And Understandable Patterns In Manufacturing Data, *Robotics and Autonomous Systems*, 49 (3): 137-152.

- Larose, D. T. (2005). *Discovering knowledge In Data : An Introduction To Data Mining*. New Jersey, USA: John Wiley & Sons,.
- Lazcorreta, E., Botella, F. ve Fernandez-Caballero, A. (2008). Towards Personalized Recommendation By Two-Step Modified Apriori Data Mining Algorithm, *Expert Systems with Applications*, 35, 1422–1429.
- Lee, T., Chiu, C., Chou, Y. ve Lu C. (2006). Mining The Customer Credit Using Classification And Regression Tree And Multivariate Adaptive Regression Splines, *Computational Statistics & Data Analysis*, 50 (4): 1113-1130.
- Lian, J., Lai, M., Lin, Q. ve Yao, F. (2002). Application Of Data Mining And Process Knowledge Discovery In Sheet Metal Assembly Dimensional Variation Diagnosis, *Journal of Materials Processing Technology*, 129 (1): 315-320.
- Liao, C.W. ve Perng, Y.H. (2008). Data Mining For Occupational Injuries In The Taiwan Construction Industry, *Safety Science*, 46, 1091–1102.
- Liao, S.H., Chang, W.H. ve Lee C.C. (2008a). Mining Marketing Maps For Business Alliances, *Expert Systems with Applications*, 35, 1338–1350.
- Liao, S.H., Ho, H.H. ve Lin H.W. (2008b). Mining Stock Category Association And Cluster On Taiwan Stock Market, *Expert Systems with Applications*, 35, 19–29.
- Liao, S.H., Hsieh,C.L. ve Huang, S.P. (2008c). Mining Product Maps For New Product Development, *Expert Systems with Applications*, 34, 50–62.
- Liao, S.H., Chen C.M. ve Wu, C.H. (2008d). Mining Customer Knowledge For Product Line And Brand Extension In Retailing, *Expert Systems with Applications*, 34, 1763–1776.
- Limitedbrands (2004). Achieving Greater Efficiencies with Market Basket Analysis, *Microstrategy World 2004 Conference*, Miami.
- Lin, F. ve McClean, S. (2001). A Data Mining Approach To The Prediction Of Corporate Failure, *Knowledge-Based Systems*, 14 (3): 189-195.
- Lin, R.H. ve Chuang, C.L. (2008). An Integrated Method For Finding Key Suppliers In SCM, *Expert Systems With Applications*, doi:10.1016/j.eswa.2008.07.078.
- Lin, C. ve Hong, C. (2008). Using Customer Knowledge In Designing Electronic Catalog, *Expert Systems with Applications*, 34, 119–127.
- Liu, D. ve Shih, Y. (2005). Integrating AHP and Data Mining For Product Recommendation Based On Customer Lifetime Value, *Information & Management*, 42 (3): 387-400 .

- Massey, P. ve Mitzi, M. (2001). Reengineering CRM: Leveraging Knowledge Assets In IBM, *Decision Support Systems*, 32 (2): 155-170.
- Meriç, O.A. (2004). *Veri Madenciliği Aracı Olarak Genetik Algoritmalar ile Yapay Sinir Ağları ve Genetik Algoritma-Yapay Sinir Ağı Melez Modelinin Müşteri Değerlendirilmesinde Kullanılması*. (Yayınlanmamış Doktora Tezi). İstanbul Üniversitesi, Sosyal Bilimler Enstitüsü, İstanbul.
- Mucuk, İ. (1994). *Pazarlama İlkeleri*, Altıncı Basım, Der Yayınları, İstanbul.
- Nakip, M. (2003). *Pazarlama Araştırmaları*, Seçkin Yayınevi, Ankara.
- Nath, V. (2003). *Data Warehousing And Mining In The Wireless Industry*. (Yayınlanmamış Yüksek Lisans Tezi). Faculty of the College of Business of Florida Atlantic University, Florida.
- Nielsen (2007). *Perakende Endeksi 2007 yılı Değerlendirmesi*, 10.04.2008: [http://www.ampd.org/images/tr/Arastirmalar/AMPD\\_Perakende\\_Endeksi/2007\\_Yilli\\_Degerlendirmesi.pdf](http://www.ampd.org/images/tr/Arastirmalar/AMPD_Perakende_Endeksi/2007_Yilli_Degerlendirmesi.pdf).
- Öğüdücü, Ş.G. (2002). *Veri Madenciliğine Giriş*, 10.04.2008: <http://www3.itu.edu.tr/~sgunduz/courses/verimaden/slides/d1.pdf>.
- Özakar, B. (2002). Veri Tabanı Yaklaşımlı Genel Web Kullanım Madenciliği Sistemi, *TBD 19. Bilişim Kurultayı*, İstanbul.
- Özçınar, H. (2006). *Kpss Sonuçlarının Veri Madenciliği Yöntemleriyle Tahmin Edilmesi*. (Yayınlanmamış Yüksek Lisans Tezi). Pamukkale Üniversitesi Fen Bilimleri Enstitüsü, Denizli.
- Özdamar, K. (1999). *Paket Programlar İle İstatistiksel Veri Analizi*, İkinci Baskı, Kaan Kitabevi, Eskişehir.
- Özdemir, A. (2004). *Veritabanlarında Bilgi Keşfi ve Veri Madenciliği-Sağlık Sektöründe Bir Uygulama*. (Yayınlanmamış Doktora Tezi). Atatürk Üniversitesi, Sosyal Bilimler Enstitüsü, Erzurum.
- Özdemir, Ş. (2006). *Müşteri Odaklı Mağaza Yönetimi*, Nobel Yayın Dağıtım, Ankara.
- Özmen, Ş. (2002). *İş Hayatı Veri Madenciliği ile İstatistik Uygulamalarını Yeniden Keşfediyor*, 12.05.2007, [www.mimoza.marmara.edu.tr/~sozmen/teblig.php](http://www.mimoza.marmara.edu.tr/~sozmen/teblig.php).
- Pande, A. ve Abdel-Aty, M. (2008). Market Basket Analysis Of Crash Data From Large Jurisdictions And Its Potential As A Decision Support Tool, *Safety Science*, doi:10.1016/j.ssci.2007.12.001

- Phillips-Wren, G., Sharkey, P. ve Dy S.M. (2008). Mining Lung Cancer Patient Data To Assess Healthcare Resource Utilization, *Expert Systems with Applications*, 35(4), s. 1611-1619.
- Piatetsky-Shapiro, G. (1991). Discovery, analysis, and presentation of strong rules. G. Piatetsky-Shapiro and W. Frawley (Editörler). *Knowledge discovery in Databases*, AAAI Press/MIT Press, 1991: 229-248.
- PricewaterhouseCoopers (2007). *Türk Perakende Sektörünün Değişimi ve Ekonomi Üzerinde Etkileri*, PricewaterhouseCoopers, 10.04.08, [http://www.ampd.org/images/tr/Arastirmalar/Raporlar/etki\\_arastirmasi.pdf](http://www.ampd.org/images/tr/Arastirmalar/Raporlar/etki_arastirmasi.pdf).
- Romero, C., Ventura, S. ve Garcia, E. (2008). Data Mining In Course Management Systems: Moodle Case Study And Tutorial, *Computers & Education*, 51, 368–384.
- Russell, G. ve Petersen A. (2000). Analysis Of Cross-Category Dependence In Market Basket Selection, *Journal of Retailing*, 76-3, 367-392.
- Russel, S. (2001). A framework for CRM, *California Management Review*, 43 (4): 89-108.
- Sanchez, D., Vila, M.A., Cerda, L. ve Serrano, J.M. (2008). Association Rules Applied To Credit Card Fraud Detection, *Expert Systems with Applications*, doi:10.1016/j.eswa.2008.02.001.
- Sangün, L. (2007). *Temel Bileşenler Analizi, Ayrırma Analizi, Kümeleme Analizleri Ve Ekolojik Verilere Uygulanması Üzerine Bir Araştırma*. (Yayınlanmamış Doktora Tezi). Çukurova Üniversitesi, Fen Bilimleri Enstitüsü, Adana.
- Selanik, M. (2007). *Türk Tarımının Avrupa Birliği İçindeki Yerinin Kümeleme Analizi İle Belirlenmesi*. (Yayınlanmamış Yüksek Lisans Tezi). Gazi Üniversitesi, Sosyal Bilimler Enstitüsü, Ankara.
- Seyrek (Birinkulu), Ö.S. (2006). *Müşteri İlişkileri Yönetiminde Veri Madenciliği ve Bir Uygulama*. (Yayınlanmamış Yüksek Lisans Tezi). İstanbul Üniversitesi Sosyal Bilimler Enstitüsü, İstanbul.
- Sforna, M. (2000). Data Mining In A Power Company Customer Database, *Electric Power Systems Research*, 55 (3): 201-209.
- Sharma, S. (1996). *Applied Multivariate Techniques*, John Wiley & Sons, Inc, ABD.
- Shih, C.C., Chiang, D.A., Lai, S.W. ve Hu Y.W. (2008). Applying Hybrid Data Mining Techniques To Web-Based Self-Assessment System Of Study and Learning Strategies Inventory, *Expert Systems with Applications*, doi:10.1016/j.eswa.2008.06.089.

- Silahtaroglu, G. (2008). *Kavram ve Algoritmalarıyla Temel Veri Madenciliği*, Papatya Yayıncılık, İstanbul.
- Sohn, S.Y. ve Kim, Y. (2008). Searching Customer Patterns Of Mobile Service Using Clustering And Quantitative Association Rule, *Expert Systems with Applications*, 34, 1070–1077.
- Srikant, R. ve Agrawal, R. (1996). Mining Quantitative Association Rules in Large Relational Tables, *ACM SIGMOD Int'l Conference on Management of Data*.
- Statsoft Internet sitesi (2007). *Data Mining Techniques*. 20.12.2007, <http://www.statsoft.com/textbook/stathome.html?stdatmin.html&1>.
- Suh, E., Lim, S., Hwang, H. ve Kim, S. (2004). A Prediction Model For The Purchase Probability Of Anonymous Customers To Support Real Time Web Marketing: A Case Study, *Expert Systems with Applications*, 27, 245–255.
- Svetina, M. ve Zupancic, J. (2005). How to Increase Sales in Retail with Market Basket Analysis, Systems Integration Conference 2005, Prag, Çek Cumhuriyeti.
- Swift, R. S. (2000). *Accelerating Customer Relationship: Using CRM And Relationship Technologies*, Texas, USA.
- Şimşek, U.T. (2006). *Veri Madenciliği ve Müşteri İlişkileri Yönetiminde Bir Uygulama*. (Yayınlanmamış Doktora Tezi). İstanbul Üniversitesi Sosyal Bilimler Enstitüsü, İstanbul.
- Tang, P.N., Steinbach, M. ve Kumar, V. (2006). *Introduction to Data Mining*. USA: Addison Wesley.
- Tantuğ, A. C. (2002). *Veri Madenciliği ve Demetleme*. (Yayınlanmamış Yüksek Lisans Tezi). İstanbul Teknik Üniversitesi Fen Bilimleri Enstitüsü, İstanbul.
- Taşkın, E. (2002). Kategori Yönetimi ve Perakende Pazarlama, *Yönetim ve Ekonomi*, cilt: 9 (1-2): 103-117.
- Tatlıdil, H. (2002). *Uygulamalı Çok Değişkenli İstatistiksel Analiz*, Akademi Matbaası, Ankara.
- Tek, Ö.B. (1984). *Perakendede Pazarlama Yönetimi*, Üçel Yayıncılık, İzmir.
- Tekin, M., Güleş, H.K. ve Öğüt, A. (2003). *Değişim Çağında Teknoloji Yönetimi*, Nobel Yayın Dağıtım, Ankara.
- Teknomo, K. (2008). *Market Basket Analysis*, 20.05.2008, <http://people.revoledu.com/kardi/tutorial/MarketBasket/index.htm>.



- Timur, N., Öztürk, S.A. ve Oyman M. (1996). *Pazarlama Kanalları*, Eskişehir:Anadolu Üniversitesi Yayınları.
- Ulaş, M.A. (2001). *Market Basket Analysis For Data Mining*. (Yayınlanmamış Yüksek Lisans Tezi). Boğaziçi Üniversitesi Fen Bilimleri Enstitüsü, İstanbul.
- Ürkmez, İ. (2008). *Tanzim ve Teşhir (Merchandising)*. İstanbul: Hayat Yayıncılık, 1.basım.
- Vahaplar, A. ve İnceoğlu, M. M. (2006). *Veri Madenciliği ve Elektronik Ticaret*. 20.02.2007, [www.bayar.edu.tr/bid/dokumanlar/inceoglu.doc](http://www.bayar.edu.tr/bid/dokumanlar/inceoglu.doc).
- Varinli, İ. (2005). *Marketlerde Pazarlama Yönetimi*. Ankara: Detay Yayınları, 1.basım.
- Yaylı, A. ve Alabay, N. (2004). *İşletmelerde Barkod Uygulamaları*, Ankara: Detay Yayıncılık, Birinci Baskı.
- Yen, S.J. ve Lee, Y.S. (2006). An Efficient Data Mining Approach For Discovering Interesting Knowledge From Customer Transactions, *Expert Systems with Applications*, 30, 650-657.
- Yıldırım, L. P. (2005). *Veri Madenciliği Ve Bir Uygulaması*. (Yayınlanmamış Yüksek Lisans Tezi). Dokuz Eylül Üniversitesi Sosyal Bilimler Enstitüsü, İzmir.
- Yılmaz, E. (2006). *Kütahya İlinde Sosyal Sınıfların Belirlenmesi ve Veri Madenciliği İle Tüketici Profilinin Çıkarılmasına Yönelik Bir Uygulama*. (Yayınlanmamış Yüksek Lisans Tezi). Dumlupınar Üniversitesi Sosyal Bilimler Enstitüsü, Kütahya.
- Zhang, C. ve Zhang S. (2002). *Association rule mining: Models and Algorithms*, Berlin, Almanya: Springer-Verlag.

**EKLER**

<b>Ek Tablo 1.</b> Ürün sınıfları arasındaki ilişkileri gösteren yakınlık matrisi	183
<b>Ek Şekil 1.</b> Ürün sınıfları arasındaki ilişkiyi gösteren dendogram grafiği	190

## ÖZGEÇMİŞ

ERTUĞRUL ERGÜN  
İŞLETME ANABİLİM DALI  
Doktora

### **Kişisel Bilgiler:**

Doğum Yeri ve Yılı: Ankara, 1976

### **Eğitim:**

Yüksek Lisans : 2002 Anadolu Üniversitesi Eğitim Bilimleri Enstitüsü

Lisans: 1999 Gazi Üniversitesi Teknik Eğitim Fakültesi

Lise: 1994 Uşak Anadolu Lisesi

### **İş/İstihdam:**

2000- Öğretim Görevlisi, Afyon Kocatepe Üniversitesi Afyon MYO

### **Mesleki Birlik/Dernek/Kuruluş Üyelikleri**

Türkiye Bilişim Derneği

### **Yabancı Dil ve Puanı**

İngilizce, 85 (KPDS)

Ek Tablo 1. Ürün sınıfları arasındaki ilişkileri gösteren yakınlık matrisi 1/6

	çikolata	bisküvi	çips	bebekmama	dondurma	misirgevrek	süt	meyve suyu	diyet ürünler	kahve	çay	enerji içeceği	kola	makarna	tuz	şeker	un vb.	baharat	pekmez helva	zeytin	zeytinyağı	ayıcicek	reçel bal	margarin	yumurta	kuryemis	bakliyat	ekmek	dondurulmus
çikolata	1.000	0.287	0.137	0.049	0.037	0.121	0.186	0.154	0.092	0.118	0.124	0.026	0.145	0.168	0.090	0.116	0.133	0.136	0.081	0.090	0.054	0.096	0.077	0.136	0.100	0.031	0.130	0.066	0.071
bisküvi	0.287	1.000	0.116	0.065	0.036	0.109	0.190	0.190	0.138	0.108	0.126	0.012	0.163	0.178	0.092	0.116	0.141	0.136	0.065	0.087	0.051	0.096	0.070	0.141	0.103	0.014	0.133	0.060	0.067
çips	0.137	0.116	1.000	0.015	0.016	0.049	0.062	0.109	0.025	0.043	0.038	0.031	0.129	0.059	0.031	0.031	0.039	0.050	0.018	0.021	0.017	0.032	0.018	0.047	0.028	0.049	0.041	0.021	0.032
bebekmama	0.049	0.065	0.015	1.000	0.007	0.039	0.100	0.038	0.031	0.028	0.040	0.000	0.019	0.061	0.028	0.043	0.063	0.047	0.017	0.027	0.023	0.037	0.028	0.050	0.058	0.000	0.052	0.023	0.026
dondurma	0.037	0.036	0.016	0.007	1.000	0.024	0.045	0.029	0.025	0.020	0.025	0.003	0.036	0.041	0.021	0.028	0.032	0.035	0.011	0.030	0.017	0.026	0.017	0.035	0.037	0.002	0.033	0.025	0.050
misirgevrek	0.121	0.109	0.049	0.039	0.024	1.000	0.164	0.074	0.072	0.092	0.086	0.008	0.049	0.123	0.064	0.077	0.101	0.099	0.045	0.059	0.035	0.067	0.055	0.111	0.072	0.018	0.098	0.056	0.057
süt	0.186	0.190	0.062	0.100	0.045	0.164	1.000	0.166	0.117	0.115	0.154	0.010	0.104	0.216	0.109	0.164	0.169	0.171	0.069	0.148	0.091	0.138	0.109	0.197	0.206	0.029	0.192	0.113	0.097
meyve suyu	0.154	0.190	0.109	0.038	0.029	0.074	0.166	1.000	0.077	0.072	0.089	0.028	0.119	0.128	0.066	0.086	0.087	0.106	0.043	0.078	0.051	0.077	0.062	0.102	0.092	0.042	0.102	0.059	0.051
diyet ürünler	0.092	0.138	0.025	0.031	0.025	0.072	0.117	0.077	1.000	0.067	0.081	0.006	0.042	0.096	0.049	0.060	0.076	0.082	0.031	0.056	0.041	0.049	0.052	0.074	0.066	0.024	0.083	0.084	0.044
kahve	0.118	0.108	0.043	0.028	0.020	0.092	0.115	0.072	0.067	1.000	0.142	0.008	0.070	0.146	0.084	0.131	0.108	0.125	0.043	0.074	0.056	0.097	0.070	0.130	0.082	0.019	0.124	0.050	0.043
çay	0.124	0.126	0.038	0.040	0.025	0.086	0.154	0.089	0.081	0.142	1.000	0.004	0.090	0.260	0.171	0.337	0.157	0.202	0.062	0.135	0.111	0.214	0.093	0.207	0.132	0.014	0.235	0.065	0.061
enerji içeceği	0.026	0.012	0.031	0.000	0.003	0.008	0.010	0.028	0.006	0.048	0.004	1.000	0.027	0.003	0.001	0.003	0.000	0.007	0.007	0.004	0.004	0.002	0.005	0.203	0.002	0.017	0.004	0.005	0.005
kola	0.145	0.163	0.129	0.019	0.036	0.049	0.104	0.119	0.042	0.070	0.090	0.027	1.000	0.112	0.066	0.093	0.070	0.097	0.033	0.067	0.036	0.077	0.035	0.096	0.083	0.016	0.090	0.039	0.044
makarna	0.168	0.178	0.059	0.061	0.041	0.123	0.216	0.128	0.096	0.146	0.260	0.003	0.112	1.000	0.221	0.256	0.232	0.305	0.081	0.158	0.134	0.249	0.115	0.266	0.172	0.019	0.403	0.088	0.094
tuz	0.090	0.092	0.031	0.028	0.021	0.064	0.109	0.066	0.049	0.084	0.171	0.001	0.066	0.221	1.000	0.170	0.160	0.204	0.047	0.092	0.082	0.158	0.064	0.166	0.088	0.010	0.195	0.041	0.050
şeker	0.116	0.116	0.031	0.043	0.028	0.077	0.164	0.086	0.060	0.131	0.337	0.003	0.093	0.256	0.170	1.000	0.174	0.188	0.060	0.134	0.109	0.246	0.088	0.211	0.146	0.014	0.256	0.054	0.060
un vb.	0.133	0.141	0.039	0.063	0.032	0.101	0.169	0.087	0.076	0.108	0.157	0.000	0.070	0.232	0.160	0.174	1.000	0.216	0.066	0.095	0.076	0.146	0.071	0.210	0.127	0.013	0.203	0.072	0.096
baharat	0.136	0.136	0.050	0.047	0.035	0.099	0.171	0.106	0.082	0.125	0.202	0.007	0.097	0.305	0.204	0.188	0.216	1.000	0.069	0.123	0.112	0.176	0.098	0.218	0.131	0.019	0.263	0.078	0.091
pekmez helva	0.081	0.065	0.018	0.017	0.011	0.045	0.069	0.043	0.031	0.043	0.062	0.007	0.033	0.081	0.047	0.060	0.066	0.069	1.000	0.074	0.036	0.054	0.078	0.088	0.065	0.010	0.076	0.039	0.029
zeytin	0.090	0.087	0.021	0.027	0.030	0.059	0.148	0.078	0.056	0.074	0.135	0.004	0.067	0.158	0.092	0.134	0.095	0.123	0.074	1.000	0.089	0.113	0.119	0.164	0.173	0.013	0.140	0.073	0.060
zeytinyağı	0.054	0.051	0.017	0.023	0.017	0.035	0.091	0.051	0.041	0.056	0.111	0.004	0.036	0.134	0.082	0.109	0.076	0.112	0.036	0.089	1.000	0.150	0.069	0.099	0.072	0.014	0.138	0.040	0.036
ayıcicek	0.096	0.096	0.032	0.037	0.026	0.067	0.138	0.077	0.049	0.097	0.214	0.002	0.077	0.249	0.158	0.246	0.146	0.176	0.054	0.113	0.150	1.000	0.079	0.193	0.134	0.010	0.239	0.047	0.053
reçel bal	0.077	0.070	0.018	0.028	0.017	0.055	0.109	0.062	0.052	0.070	0.093	0.005	0.035	0.115	0.064	0.088	0.071	0.098	0.078	0.119	0.069	0.079	1.000	0.114	0.101	0.018	0.104	0.050	0.046
margarin	0.136	0.141	0.047	0.050	0.035	0.111	0.197	0.102	0.074	0.130	0.207	0.003	0.096	0.266	0.166	0.211	0.210	0.218	0.088	0.164	0.099	0.193	0.114	1.000	0.174	0.014	0.230	0.076	0.085
yumurta	0.100	0.103	0.028	0.058	0.037	0.072	0.206	0.092	0.066	0.082	0.132	0.002	0.083	0.172	0.088	0.146	0.127	0.131	0.065	0.173	0.072	0.134	0.101	0.174	1.000	0.021	0.157	0.088	0.095
kuryemis	0.031	0.014	0.049	0.000	0.002	0.018	0.029	0.042	0.024	0.019	0.014	0.017	0.016	0.019	0.010	0.014	0.013	0.019	0.010	0.013	0.014	0.010	0.018	0.014	0.021	1.000	0.025	0.023	0.013
bakliyat	0.130	0.133	0.041	0.052	0.033	0.098	0.192	0.102	0.083	0.124	0.235	0.004	0.090	0.403	0.195	0.256	0.203	0.263	0.076	0.140	0.138	0.239	0.104	0.230	0.157	0.025	1.000	0.076	0.081
ekmek	0.066	0.060	0.021	0.023	0.025	0.056	0.113	0.059	0.084	0.050	0.065	0.005	0.039	0.088	0.041	0.054	0.072	0.078	0.039	0.073	0.040	0.047	0.050	0.076	0.088	0.023	0.076	1.000	0.081
dondurulmus	0.071	0.067	0.032	0.026	0.050	0.057	0.097	0.051	0.044	0.043	0.061	0.005	0.044	0.094	0.050	0.060	0.096	0.091	0.029	0.060	0.036	0.053	0.046	0.085	0.095	0.013	0.081	0.081	1.000
peynir	0.143	0.134	0.040	0.060	0.045	0.107	0.244	0.129	0.106	0.111	0.174	0.008	0.099	0.215	0.115	0.168	0.148	0.167	0.100	0.354	0.098	0.138	0.145	0.229	0.230	0.026	0.184	0.145	0.106
yogurt	0.106	0.106	0.036	0.057	0.042	0.082	0.238	0.108	0.087	0.092	0.136	0.002	0.093	0.197	0.099	0.144	0.123	0.151	0.060	0.154	0.085	0.131	0.093	0.192	0.207	0.022	0.176	0.097	0.075
ayran	0.042	0.041	0.025	0.008	0.015	0.027	0.068	0.054	0.035	0.022	0.023	0.006	0.039	0.036	0.020	0.022	0.023	0.034	0.023	0.040	0.015	0.019	0.021	0.039	0.039	0.010	0.031	0.045	0.027
unlu-hazır yemek	0.076	0.076	0.024	0.017	0.016	0.048	0.115	0.084	0.052	0.041	0.056	0.000	0.085	0.095	0.047	0.053	0.061	0.075	0.087	0.125	0.032	0.048	0.058	0.073	0.122	-0.010	0.072	0.096	0.058
sürme çikolata	0.137	0.136	0.038	0.040	0.029	0.098	0.122	0.074	0.050	0.110	0.139	0.002	0.075	0.173	0.101	0.132	0.125	0.134	0.069	0.105	0.051	0.113	0.097	0.174	0.091	0.010	0.135	0.047	0.059
çocuk gel. yog	0.153	0.144	0.068	0.134	0.031	0.127	0.200	0.111	0.079	0.073	0.113	0.010	0.078	0.146	0.073	0.111	0.108	0.121	0.046	0.081	0.042	0.084	0.064	0.126	0.106	0.019	0.118	0.063	0.066
salça konserve	0.123	0.122	0.052	0.045	0.034	0.095	0.185	0.115	0.092	0.109	0.171	0.006	0.091	0.266	0.157	0.169	0.162	0.252	0.067	0.158	0.108	0.161	0.100	0.185	0.147	0.024	0.231	0.092	0.099
ketçap mayonez	0.100	0.102	0.054	0.028	0.020	0.079	0.115	0.080	0.045	0.090	0.129	0.006	0.075	0.214	0.111	0.122	0.128	0.169	0.044	0.081	0.084	0.152	0.061	0.153	0.086	0.009	0.160	0.058	0.077
sucuk sosis	0.096	0.084	0.034	0.030	0.027	0.076	0.138	0.077	0.048	0.065	0.092	0.013	0.073	0.131	0.064	0.086	0.086	0.112	0.087	0.168	0.052	0.081	0.092	0.145	0.144	0.011	0.102	0.099	0.080
et kıyma	0.050	0.042	0.009	0.031	0.036	0.035	0.101	0.050	0.037	0.039	0.064	-0.001	0.068	0.097	0.053	0.068	0.055	0.107	0.033	0.111	0.048	0.069	0.059	0.080	0.128	0.014	0.095	0.067	0.051
tavuk	0.091	0.089	0.019	0.036	0.042	0.061	0.138	0.072	0.059	0.060	0.104	0.001	0.090	0.154	0.079	0.104	0.100	0.146	0.051										

Ek Tablo 1. Ürün sınıfları arasındaki ilişkileri gösteren yakınlık matrisi 2/6

	çikolata	bisküvi	çips	bebekmama	dondurma	msirgevrek	süt	meyvesuyu	diyet ürünler	kahve	çay	enerji içeceği	kola	makarna	tuz	şeker	un vb.	baharat	pekmez helva	zeytin	zeytinyağı	ayçiçek	reçel bal	margarin	yumurta	kuruyemis	bakliyat	ekmek	dondurulus
camasır suyu	0.091	0.091	0.023	0.033	0.025	0.066	0.123	0.068	0.056	0.082	0.142	0.003	0.060	0.185	0.117	0.144	0.127	0.146	0.040	0.082	0.075	0.130	0.056	0.141	0.087	0.007	0.161	0.050	0.051
yumusatici	0.078	0.072	0.020	0.035	0.021	0.059	0.109	0.057	0.051	0.065	0.113	0.005	0.046	0.150	0.095	0.119	0.109	0.126	0.030	0.075	0.072	0.109	0.048	0.117	0.073	0.007	0.141	0.048	0.047
makyaj	0.055	0.043	0.024	0.016	0.012	0.043	0.067	0.035	0.043	0.043	0.057	0.010	0.022	0.068	0.039	0.045	0.055	0.063	0.019	0.041	0.031	0.041	0.032	0.052	0.043	0.009	0.059	0.030	0.032
sampuan	0.098	0.083	0.039	0.026	0.022	0.065	0.107	0.066	0.056	0.080	0.122	0.012	0.057	0.150	0.090	0.110	0.098	0.121	0.033	0.077	0.067	0.100	0.057	0.115	0.069	0.012	0.125	0.043	0.049
kadın pedi	0.082	0.073	0.042	0.030	0.017	0.058	0.093	0.056	0.052	0.066	0.091	0.008	0.042	0.122	0.070	0.083	0.092	0.099	0.028	0.050	0.045	0.072	0.041	0.092	0.054	0.012	0.102	0.038	0.045
agiz bakim	0.082	0.064	0.031	0.022	0.013	0.059	0.092	0.057	0.053	0.066	0.091	0.009	0.036	0.114	0.066	0.079	0.078	0.091	0.028	0.061	0.050	0.071	0.047	0.087	0.063	0.016	0.096	0.040	0.038
sabun-dus	0.125	0.120	0.042	0.032	0.026	0.076	0.143	0.089	0.071	0.099	0.177	0.007	0.079	0.215	0.133	0.163	0.145	0.168	0.052	0.113	0.091	0.143	0.074	0.159	0.104	0.012	0.183	0.057	0.063
sac boya jöle	0.043	0.031	0.022	0.011	0.011	0.032	0.049	0.027	0.024	0.036	0.048	0.008	0.023	0.063	0.037	0.045	0.045	0.053	0.013	0.030	0.026	0.039	0.023	0.051	0.033	0.006	0.053	0.025	0.023
bebek temizlik	0.035	0.038	0.008	0.081	0.005	0.023	0.047	0.022	0.011	0.019	0.035	-0.001	0.014	0.046	0.021	0.035	0.034	0.034	0.012	0.018	0.014	0.030	0.018	0.035	0.030	-0.005	0.036	0.010	0.013
gunes süt krem	-0.007	-0.007	-0.005	0.009	-0.002	0.000	-0.003	0.000	0.000	-0.002	-0.004	0.004	-0.005	-0.006	-0.002	-0.005	-0.005	-0.005	-0.007	-0.004	-0.001	-0.004	-0.001	-0.007	-0.006	0.000	-0.005	-0.001	-0.003
kagit temizlik	0.150	0.147	0.040	0.059	0.032	0.099	0.192	0.113	0.089	0.125	0.209	0.004	0.093	0.259	0.161	0.207	0.179	0.205	0.060	0.134	0.113	0.181	0.092	0.198	0.142	0.015	0.228	0.077	0.076
bebek malz	0.135	0.123	0.046	0.114	0.024	0.096	0.164	0.095	0.066	0.088	0.129	0.007	0.058	0.171	0.099	0.121	0.126	0.137	0.047	0.085	0.066	0.108	0.068	0.127	0.096	0.017	0.145	0.057	0.059
plastik ahsap	0.042	0.051	0.041	0.008	0.021	0.017	0.031	0.063	0.021	0.023	0.028	0.004	0.114	0.026	0.031	0.027	0.023	0.033	0.024	0.037	0.017	0.016	0.022	0.021	0.020	0.008	0.025	0.022	0.012
oda parf.	0.071	0.059	0.022	0.027	0.021	0.058	0.094	0.056	0.047	0.069	0.092	0.008	0.038	0.115	0.068	0.089	0.088	0.108	0.026	0.056	0.056	0.073	0.048	0.091	0.062	0.010	0.105	0.044	0.054
sinek öld.	0.017	0.016	0.003	0.015	0.017	0.013	0.030	0.017	0.013	0.022	0.029	0.001	0.016	0.034	0.023	0.029	0.022	0.031	0.007	0.026	0.020	0.024	0.018	0.029	0.022	0.003	0.028	0.013	0.010
kitap	0.037	0.019	0.010	0.014	0.002	0.043	0.046	0.019	0.024	0.019	0.023	0.004	-0.001	0.031	0.013	0.017	0.027	0.027	0.013	0.016	0.014	0.015	0.020	0.020	0.023	0.005	0.027	0.022	0.015
icki	-0.012	-0.018	0.099	-0.007	-0.002	0.005	0.013	0.012	0.004	0.016	0.012	0.010	0.013	0.015	0.004	0.008	0.002	0.013	-0.011	0.011	0.008	0.012	0.005	0.010	0.019	0.062	0.010	0.011	0.013
züccaciye	0.084	0.065	0.023	0.025	0.017	0.053	0.085	0.049	0.045	0.063	0.087	0.006	0.040	0.100	0.063	0.081	0.080	0.096	0.044	0.064	0.044	0.060	0.050	0.071	0.060	0.004	0.086	0.045	0.040
evcil hayvan	0.008	0.002	0.004	-0.002	0.005	0.009	0.018	0.006	0.016	0.014	0.012	0.002	0.004	0.010	0.007	0.012	0.010	0.016	0.004	0.007	0.008	0.007	0.007	0.008	0.014	0.005	0.014	0.021	0.009
araba bakim	0.012	0.000	0.005	0.004	-0.001	0.010	0.015	0.008	0.004	0.012	0.016	0.007	0.004	0.017	0.009	0.015	0.015	0.017	0.005	0.007	0.008	0.011	0.010	0.017	0.012	-0.004	0.016	0.004	0.008
pil	0.025	0.019	0.008	0.022	0.005	0.021	0.030	0.015	0.012	0.017	0.023	0.002	0.009	0.028	0.015	0.019	0.023	0.025	0.010	0.015	0.012	0.017	0.014	0.022	0.019	0.001	0.023	0.014	0.014
kirtasiye	0.054	0.039	0.018	0.014	0.007	0.049	0.062	0.030	0.024	0.035	0.041	0.007	0.016	0.058	0.031	0.039	0.044	0.050	0.024	0.028	0.022	0.032	0.028	0.045	0.032	0.004	0.053	0.025	0.022
oyuncak	0.032	0.011	0.005	0.047	-0.003	0.037	0.043	0.011	0.004	0.007	0.009	0.004	-0.007	0.021	0.009	0.010	0.019	0.016	0.015	0.006	0.009	0.009	0.013	0.011	0.017	-0.010	0.017	0.010	0.011
dvd vcd	0.020	0.009	0.006	0.011	0.006	0.028	0.032	0.012	0.010	0.014	0.010	0.007	0.001	0.017	0.006	0.008	0.014	0.019	0.008	0.010	0.009	0.008	0.013	0.012	0.015	0.003	0.014	0.017	0.014
müzik kaset	-0.008	-0.018	-0.005	0.000	-0.003	0.002	-0.001	-0.006	0.002	0.000	-0.004	0.013	-0.010	-0.007	-0.004	-0.006	-0.007	-0.005	0.002	-0.003	0.001	-0.005	0.003	-0.006	-0.004	0.000	-0.005	0.004	-0.001
hazır tatli	0.138	0.143	0.045	0.055	0.032	0.103	0.169	0.088	0.083	0.109	0.140	0.004	0.073	0.218	0.126	0.145	0.308	0.197	0.067	0.086	0.067	0.122	0.074	0.176	0.106	0.017	0.179	0.081	0.092
su maden suyu	0.033	0.031	0.025	0.006	0.011	0.015	0.032	0.061	0.020	-0.001	-0.001	0.021	0.053	0.009	0.006	-0.001	0.004	0.003	0.000	0.010	0.003	0.002	0.007	0.001	0.009	0.003	0.006	0.005	0.001
hazır corba	0.112	0.116	0.045	0.031	0.038	0.065	0.144	0.115	0.067	0.092	0.130	0.009	0.116	0.184	0.100	0.125	0.108	0.174	0.040	0.094	0.064	0.113	0.067	0.136	0.113	0.011	0.145	0.056	0.057
sebze meyve	0.083	0.080	0.028	0.037	0.028	0.046	0.149	0.093	0.074	0.045	0.073	0.011	0.079	0.099	0.055	0.071	0.071	0.094	0.055	0.153	0.057	0.062	0.077	0.083	0.172	0.027	0.096	0.104	0.071
sakiz	0.116	0.082	0.055	0.016	0.018	0.054	0.079	0.064	0.044	0.043	0.041	0.022	0.048	0.053	0.029	0.036	0.049	0.046	0.025	0.028	0.020	0.027	0.025	0.044	0.035	0.021	0.044	0.032	0.022
deoda. kolonya	0.073	0.052	0.026	0.012	0.015	0.045	0.064	0.043	0.039	0.054	0.068	0.013	0.033	0.081	0.049	0.058	0.057	0.072	0.022	0.046	0.037	0.050	0.035	0.061	0.041	0.009	0.067	0.027	0.030
erkek tras	0.052	0.038	0.020	0.015	0.009	0.033	0.060	0.039	0.027	0.044	0.060	0.012	0.028	0.072	0.042	0.058	0.051	0.060	0.018	0.040	0.035	0.049	0.039	0.057	0.044	0.012	0.063	0.023	0.027
sigara	-0.156	-0.127	-0.082	-0.044	-0.037	-0.051	-0.103	-0.100	-0.054	-0.057	-0.073	-0.028	-0.128	-0.078	-0.042	-0.061	-0.058	-0.068	-0.066	-0.068	-0.029	-0.050	-0.043	-0.064	-0.065	-0.036	-0.062	-0.046	-0.035
diger	0.079	0.062	0.021	0.034	0.017	0.059	0.105	0.066	0.052	0.062	0.091	0.007	0.046	0.118	0.076	0.088	0.095	0.114	0.040	0.072	0.053	0.075	0.059	0.093	0.081	0.010	0.107	0.052	0.054
lokum	0.032	0.002	-0.004	-0.007	-0.010	0.000	-0.003	0.004	0.000	-0.011	-0.011	0.013	-0.002	-0.012	-0.008	-0.015	-0.009	-0.007	0.158	0.015	-0.001	-0.011	0.017	-0.011	-0.003	-0.002	-0.009	0.009	-0.005
prezervatif	0.016	0.012	0.009	0.011	0.003	0.011	0.022	0.014	0.008	0.008	0.011	0.006	0.006	0.018	0.007	0.011	0.013	0.013	0.002	0.012	0.007	0.009	0.006	0.015	0.017	0.004	0.017	0.009	0.009
dergi	-0.042	-0.043	-0.023	-0.011	-0.011	-0.010	-0.022	-0.030	-0.010	-0.015	-0.017	-0.006	-0.041	-0.017	-0.011	-0.015	-0.013	-0.015	-0.016	-0.016	-0.006	-0.012	-0.009	-0.015	-0.014	-0.005	-0.013	-0.006	-0.007
gazete	-0.085	-0.078	-0.044	-0.023	-0.024	-0.026	-0.050	-0.057	-0.026	-0.030	-0.035	-0.013	-0.075	-0.038	-0.020	-0.030	-0.029	-0.034	-0.031	-0.033	-0.014	-0.024	-0.021	-0.031	-0.030	-0.019	-0.030	-0.021	-0.017

Ek Tablo 1. Ürün sınıfları arasındaki ilişkileri gösteren yakınlık matrisi 3/6

	peynir	yoğurt	ayran	unlu-hazır yemek	şürme cokolata	çocuk gel. yog	salca konserve	ketcap mayonez	sucuk sosis	et kıyma	tavuk	genel temizlik	bulasik el	bulasik mak	camasir el	camasir mak	camasir suyu	yumusatıcı	makyaaj	şampuan	kadın pedi	agiz bakim	sebutin-dus	şarc boyaj (öle	bebek temizlik	gunes sut krem	kagit temizlik	bebek malz	plastik ahşap
cikolata	0.143	0.106	0.042	0.076	0.137	0.153	0.123	0.100	0.096	0.050	0.091	0.146	0.081	0.085	0.025	0.121	0.091	0.078	0.055	0.098	0.082	0.082	0.125	0.043	0.035	-0.007	0.150	0.135	0.042
bisküvi	0.134	0.106	0.041	0.076	0.136	0.144	0.122	0.102	0.084	0.042	0.089	0.138	0.081	0.075	0.026	0.114	0.091	0.072	0.043	0.083	0.073	0.064	0.120	0.031	0.038	-0.007	0.147	0.123	0.051
cips	0.040	0.036	0.025	0.024	0.038	0.068	0.052	0.054	0.034	0.009	0.019	0.041	0.021	0.019	0.008	0.034	0.023	0.020	0.024	0.039	0.042	0.031	0.042	0.022	0.008	-0.005	0.040	0.046	0.041
bebekmama	0.060	0.057	0.008	0.017	0.040	0.134	0.045	0.028	0.030	0.031	0.036	0.056	0.031	0.039	0.004	0.061	0.033	0.035	0.016	0.026	0.030	0.022	0.032	0.011	0.081	0.009	0.059	0.114	0.008
dondurma	0.045	0.042	0.015	0.016	0.029	0.031	0.034	0.020	0.027	0.036	0.042	0.037	0.020	0.024	0.003	0.025	0.025	0.021	0.012	0.022	0.017	0.013	0.026	0.011	0.005	0.002	0.032	0.024	0.021
misirgevrek	0.107	0.082	0.027	0.048	0.098	0.127	0.095	0.079	0.076	0.035	0.061	0.104	0.052	0.074	0.016	0.081	0.066	0.059	0.043	0.065	0.058	0.059	0.076	0.032	0.023	0.000	0.099	0.096	0.017
süt	0.244	0.238	0.068	0.115	0.122	0.200	0.185	0.115	0.138	0.101	0.138	0.183	0.114	0.121	0.028	0.150	0.123	0.109	0.067	0.107	0.093	0.092	0.143	0.049	0.047	-0.003	0.192	0.164	0.031
meyve suyu	0.129	0.108	0.054	0.084	0.074	0.111	0.115	0.080	0.077	0.050	0.072	0.102	0.063	0.061	0.017	0.080	0.068	0.057	0.035	0.066	0.056	0.057	0.089	0.027	0.022	0.000	0.113	0.095	0.063
diyet ürünler	0.106	0.087	0.035	0.052	0.050	0.079	0.092	0.045	0.048	0.037	0.059	0.090	0.044	0.058	0.013	0.062	0.056	0.051	0.043	0.056	0.052	0.053	0.071	0.024	0.011	0.000	0.089	0.066	0.021
kahve	0.111	0.092	0.022	0.041	0.110	0.073	0.109	0.090	0.065	0.039	0.060	0.124	0.080	0.071	0.023	0.095	0.082	0.065	0.043	0.080	0.066	0.066	0.099	0.036	0.019	-0.002	0.125	0.088	0.023
çay	0.174	0.136	0.023	0.056	0.139	0.113	0.171	0.129	0.092	0.064	0.104	0.194	0.149	0.113	0.043	0.179	0.142	0.113	0.057	0.122	0.091	0.091	0.177	0.048	0.035	-0.004	0.209	0.129	0.028
enerji içeceği	0.008	0.002	0.006	0.000	0.002	0.010	0.006	0.006	0.013	-0.001	0.001	0.005	0.002	0.002	0.000	0.004	0.003	0.005	0.010	0.012	0.008	0.009	0.007	0.008	-0.001	0.004	0.004	0.007	0.004
kola	0.099	0.093	0.039	0.085	0.075	0.078	0.091	0.075	0.073	0.068	0.090	0.080	0.057	0.042	0.019	0.071	0.060	0.046	0.022	0.057	0.042	0.036	0.079	0.023	0.014	-0.005	0.093	0.058	0.114
makarna	0.215	0.197	0.036	0.095	0.173	0.146	0.266	0.214	0.131	0.097	0.154	0.253	0.177	0.157	0.056	0.214	0.185	0.150	0.068	0.150	0.122	0.114	0.215	0.063	0.046	-0.006	0.259	0.171	0.026
tuz	0.115	0.099	0.020	0.047	0.101	0.073	0.157	0.111	0.064	0.053	0.079	0.153	0.118	0.088	0.038	0.128	0.117	0.095	0.039	0.090	0.070	0.066	0.133	0.037	0.021	-0.002	0.161	0.099	0.031
şeker	0.168	0.144	0.022	0.053	0.132	0.111	0.169	0.122	0.086	0.068	0.104	0.187	0.158	0.117	0.044	0.172	0.144	0.119	0.045	0.110	0.083	0.079	0.163	0.045	0.035	-0.005	0.207	0.121	0.027
un vb.	0.148	0.123	0.023	0.061	0.125	0.108	0.162	0.128	0.086	0.055	0.100	0.190	0.104	0.121	0.033	0.141	0.127	0.109	0.055	0.098	0.092	0.078	0.145	0.045	0.034	-0.005	0.179	0.126	0.023
baharat	0.167	0.151	0.034	0.075	0.134	0.121	0.252	0.169	0.112	0.107	0.146	0.210	0.141	0.120	0.043	0.170	0.146	0.126	0.063	0.121	0.099	0.091	0.168	0.053	0.034	-0.005	0.205	0.137	0.033
pekmez helva	0.100	0.060	0.023	0.087	0.069	0.046	0.067	0.044	0.087	0.033	0.051	0.057	0.036	0.032	0.012	0.047	0.040	0.030	0.019	0.033	0.028	0.028	0.052	0.013	0.012	-0.007	0.060	0.047	0.024
zeytin	0.354	0.154	0.040	0.125	0.105	0.081	0.158	0.081	0.168	0.111	0.137	0.124	0.098	0.075	0.031	0.098	0.082	0.075	0.041	0.077	0.050	0.061	0.113	0.030	0.018	-0.004	0.134	0.085	0.037
zeytinyagi	0.098	0.085	0.015	0.032	0.051	0.042	0.108	0.084	0.052	0.048	0.059	0.105	0.085	0.077	0.026	0.096	0.075	0.072	0.031	0.067	0.045	0.050	0.091	0.026	0.014	-0.001	0.113	0.066	0.017
ayccicek	0.138	0.131	0.019	0.048	0.113	0.084	0.161	0.152	0.081	0.069	0.094	0.166	0.146	0.107	0.044	0.165	0.130	0.109	0.041	0.100	0.072	0.071	0.143	0.039	0.030	-0.004	0.181	0.108	0.016
reçel bal	0.145	0.093	0.021	0.058	0.097	0.064	0.100	0.061	0.092	0.059	0.074	0.092	0.068	0.053	0.021	0.070	0.056	0.048	0.032	0.057	0.041	0.047	0.074	0.023	0.018	-0.001	0.092	0.068	0.022
margarin	0.229	0.192	0.039	0.073	0.174	0.126	0.185	0.153	0.145	0.080	0.126	0.194	0.137	0.110	0.044	0.163	0.141	0.117	0.052	0.115	0.092	0.087	0.159	0.051	0.035	-0.007	0.198	0.127	0.021
yumurta	0.230	0.207	0.039	0.122	0.091	0.106	0.147	0.086	0.144	0.128	0.157	0.127	0.097	0.078	0.025	0.100	0.087	0.073	0.043	0.069	0.054	0.063	0.104	0.033	0.030	-0.006	0.142	0.096	0.020
kuruyemis	0.026	0.022	0.010	-0.010	0.010	0.019	0.024	0.009	0.011	0.014	0.008	0.015	0.007	0.010	0.002	0.006	0.007	0.007	0.009	0.012	0.012	0.016	0.012	0.006	-0.005	0.000	0.015	0.017	0.008
bakliyat	0.184	0.176	0.031	0.072	0.135	0.118	0.231	0.160	0.102	0.095	0.138	0.223	0.164	0.146	0.048	0.190	0.161	0.141	0.059	0.125	0.102	0.096	0.183	0.053	0.036	-0.005	0.228	0.145	0.025
ekmek	0.145	0.097	0.045	0.096	0.047	0.063	0.092	0.058	0.099	0.067	0.087	0.080	0.044	0.053	0.011	0.054	0.050	0.048	0.030	0.043	0.038	0.040	0.057	0.025	0.010	-0.001	0.077	0.057	0.022
dondurulmus	0.106	0.075	0.027	0.058	0.059	0.066	0.099	0.077	0.080	0.051	0.087	0.086	0.045	0.057	0.010	0.057	0.051	0.047	0.032	0.049	0.045	0.038	0.063	0.023	0.013	-0.003	0.076	0.059	0.012
peynir	1.000	0.229	0.072	0.215	0.141	0.150	0.204	0.119	0.254	0.144	0.178	0.175	0.114	0.112	0.031	0.133	0.113	0.102	0.063	0.102	0.081	0.087	0.147	0.046	0.028	-0.006	0.185	0.129	0.053
yoğurt	0.229	1.000	0.066	0.121	0.100	0.117	0.173	0.097	0.120	0.132	0.160	0.143	0.100	0.086	0.029	0.115	0.096	0.085	0.045	0.081	0.062	0.063	0.115	0.034	0.026	-0.005	0.154	0.101	0.031
ayran	0.072	0.066	1.000	0.094	0.014	0.042	0.048	0.021	0.052	0.044	0.047	0.032	0.021	0.019	0.005	0.020	0.018	0.016	0.017	0.016	0.014	0.017	0.030	0.007	0.003	-0.001	0.035	0.029	0.046
unlu-hazır yemek	0.215	0.121	0.094	1.000	0.050	0.060	0.110	0.050	0.134	0.122	0.133	0.062	0.037	0.041	0.009	0.041	0.039	0.031	0.018	0.028	0.025	0.023	0.053	0.012	0.006	-0.013	0.071	0.049	0.055
şürme cokolata	0.141	0.100	0.014	0.050	1.000	0.103	0.115	0.114	0.094	0.043	0.078	0.127	0.083	0.070	0.028	0.116	0.089	0.074	0.034	0.087	0.065	0.062	0.113	0.032	0.035	-0.003	0.132	0.094	0.015
çocuk gel. yog	0.150	0.117	0.042	0.060	0.103	1.000	0.117	0.084	0.096	0.053	0.082	0.120	0.073	0.078	0.018	0.108	0.080	0.070	0.039	0.070	0.062	0.057	0.087	0.031	0.058	-0.003	0.116	0.133	0.027
salca konserve	0.204	0.173	0.048	0.110	0.115	0.117	1.000	0.163	0.122	0.099	0.137	0.182	0.119	0.120	0.029	0.140	0.121	0.110	0.059	0.106	0.087	0.082	0.144	0.047	0.026	-0.005	0.179	0.119	0.036
ketcap mayonez	0.119	0.097	0.021	0.050	0.114	0.084	0.163	1.000	0.104	0.046	0.082	0.135	0.091	0.083	0.030	0.118	0.097	0.085	0.045	0.088	0.073	0.067	0.115	0.038	0.022	-0.003	0.137	0.094	0.016
sucuk sosis	0.254	0.120	0.052	0.134	0.094	0.096	0.122	0.104	1.000	0.109	0.141	0.092	0.064	0.062	0.019	0.076	0.062	0.058	0.036	0.061	0.046	0.050	0.080	0.031	0.017	-0.004	0.099	0.073	0.033
et kıyma	0.144	0.132	0.044	0.122	0.043	0.053	0.099	0.046	0.109	1.000	0.240	0.072	0.053	0.046	0.016	0.054	0.051	0.045	0.020	0.041	0.025	0.031	0.056	0.019	0.010	-0.005	0.081	0.050	0.024
tavuk	0.178	0.160	0.047	0.133	0.078	0.082	0.137	0.082	0.141	0.240	1.000	0.111	0.074	0.071	0.022														

Ek Tablo 1. Ürün sınıfları arasındaki ilişkileri gösteren yakınlık matrisi 4/6

	peynir	yoğurt	ayran	unlu-hazır yemek	şurme cikolata	çocuk gel. yog	selca konserve	ketçap mayonez	sucuk sosis	et kıyma	lavuk	genel temizlik	bulaşık el	bulaşık mak	camasır el	camasır mak	camasır suyu	yumusatici	makyaj	sampuan	kadın pedi	agiz bakım	sabun-dus	sac boya jöle	bebek temizlik	gunes sut krem	kagit temizlik	bebek malz	plastik ahsap
camasır suyu	0.113	0.096	0.018	0.039	0.089	0.080	0.121	0.097	0.062	0.051	0.079	0.258	0.176	0.146	0.055	0.201	1.000	0.176	0.047	0.110	0.076	0.071	0.159	0.045	0.031	-0.001	0.206	0.122	0.020
yumusatici	0.102	0.085	0.016	0.031	0.074	0.070	0.110	0.085	0.058	0.045	0.066	0.207	0.159	0.147	0.044	0.226	0.176	1.000	0.047	0.095	0.074	0.067	0.137	0.040	0.026	-0.002	0.178	0.109	0.021
makyaj	0.063	0.045	0.017	0.018	0.034	0.039	0.059	0.045	0.036	0.020	0.035	0.097	0.044	0.052	0.015	0.066	0.047	0.047	1.000	0.110	0.081	0.111	0.116	0.080	0.009	0.033	0.082	0.093	0.029
sampuan	0.102	0.081	0.016	0.028	0.087	0.070	0.106	0.088	0.061	0.041	0.063	0.164	0.103	0.095	0.035	0.153	0.110	0.095	0.110	1.000	0.113	0.162	0.222	0.087	0.027	0.035	0.163	0.144	0.025
kadın pedi	0.081	0.062	0.014	0.025	0.065	0.062	0.087	0.073	0.046	0.025	0.045	0.125	0.071	0.085	0.024	0.108	0.076	0.074	0.081	0.113	1.000	0.099	0.114	0.057	0.039	0.012	0.129	0.117	0.019
agiz bakım	0.087	0.063	0.017	0.023	0.062	0.057	0.082	0.067	0.050	0.031	0.048	0.124	0.071	0.072	0.023	0.103	0.071	0.067	0.111	0.162	0.099	1.000	0.158	0.087	0.014	0.031	0.121	0.117	0.023
sabun-dus	0.147	0.115	0.030	0.053	0.113	0.087	0.144	0.115	0.080	0.056	0.094	0.242	0.170	0.129	0.060	0.202	0.159	0.137	0.116	0.222	0.114	0.158	1.000	0.075	0.033	0.012	0.245	0.184	0.039
sac boya jöle	0.046	0.034	0.007	0.012	0.032	0.031	0.047	0.038	0.031	0.019	0.029	0.071	0.040	0.040	0.012	0.056	0.045	0.040	0.080	0.087	0.057	0.087	0.075	1.000	0.007	0.022	0.062	0.061	0.012
bebek temizlik	0.028	0.026	0.003	0.006	0.035	0.058	0.026	0.022	0.017	0.010	0.019	0.043	0.027	0.025	0.006	0.068	0.031	0.026	0.009	0.027	0.039	0.014	0.033	0.007	1.000	0.001	0.054	0.092	0.000
gunes sut krem	-0.006	-0.005	-0.001	-0.013	-0.003	-0.003	-0.005	-0.003	-0.004	-0.005	-0.007	0.002	-0.002	-0.001	0.002	0.000	-0.001	-0.002	0.033	0.035	0.012	0.031	0.012	0.022	0.001	1.000	0.002	0.024	0.002
kagit temizlik	0.185	0.154	0.035	0.071	0.132	0.116	0.179	0.137	0.099	0.081	0.122	0.282	0.196	0.174	0.063	0.244	0.206	0.178	0.082	0.163	0.129	0.121	0.245	0.062	0.054	0.002	1.000	0.235	0.045
bebek malz	0.129	0.101	0.029	0.049	0.094	0.133	0.119	0.094	0.073	0.050	0.079	0.193	0.112	0.113	0.035	0.171	0.122	0.109	0.093	0.144	0.117	0.117	0.184	0.061	0.092	0.024	0.235	1.000	0.040
plastik ahsap	0.053	0.031	0.046	0.055	0.015	0.027	0.036	0.016	0.033	0.024	0.020	0.050	0.023	0.020	0.009	0.023	0.020	0.021	0.029	0.025	0.019	0.023	0.039	0.012	0.000	0.002	0.045	0.040	1.000
oda parf.	0.092	0.070	0.018	0.025	0.061	0.064	0.098	0.070	0.051	0.042	0.058	0.178	0.085	0.105	0.019	0.113	0.112	0.100	0.053	0.082	0.060	0.066	0.107	0.043	0.021	0.002	0.135	0.096	0.028
sinek öld.	0.031	0.028	0.009	0.006	0.021	0.022	0.027	0.017	0.021	0.021	0.018	0.046	0.030	0.021	0.012	0.032	0.028	0.025	0.022	0.028	0.019	0.024	0.032	0.020	0.006	0.031	0.036	0.032	0.012
kitap	0.032	0.021	0.005	0.002	0.019	0.039	0.025	0.019	0.022	0.012	0.016	0.040	0.012	0.024	0.004	0.029	0.020	0.021	0.035	0.029	0.025	0.036	0.029	0.018	0.006	0.013	0.033	0.047	0.012
ıcki	0.020	0.023	0.000	-0.006	0.003	0.005	0.026	0.012	0.011	0.027	0.021	0.005	0.007	0.005	0.000	0.004	0.004	0.006	0.001	0.005	0.004	0.002	0.006	0.003	-0.003	-0.008	0.007	-0.003	-0.003
züccaciye	0.092	0.060	0.024	0.051	0.055	0.060	0.080	0.054	0.055	0.044	0.054	0.127	0.064	0.062	0.019	0.078	0.065	0.059	0.053	0.066	0.054	0.056	0.094	0.034	0.018	-0.002	0.104	0.091	0.091
evcil hayvan	0.020	0.017	0.003	0.007	0.004	0.002	0.016	0.006	0.010	0.014	0.013	0.021	0.007	0.011	0.002	0.012	0.013	0.011	0.017	0.011	0.011	0.013	0.015	0.008	-0.002	-0.001	0.012	0.007	0.006
araba bakım	0.010	0.009	-0.001	-0.011	0.013	0.012	0.013	0.011	0.007	0.003	0.007	0.030	0.012	0.013	0.002	0.018	0.011	0.013	0.017	0.017	0.009	0.017	0.018	0.013	0.005	0.000	0.020	0.017	0.009
pil	0.022	0.019	0.003	0.008	0.018	0.024	0.023	0.017	0.016	0.009	0.015	0.031	0.015	0.021	0.005	0.024	0.017	0.012	0.019	0.019	0.015	0.021	0.022	0.010	0.008	0.004	0.026	0.032	0.010
kirtasiye	0.045	0.034	0.009	0.015	0.040	0.052	0.046	0.032	0.033	0.018	0.030	0.067	0.032	0.039	0.010	0.048	0.035	0.033	0.037	0.043	0.041	0.045	0.051	0.026	0.008	0.003	0.058	0.064	0.027
oyuncak	0.017	0.010	0.001	-0.007	0.016	0.060	0.014	0.013	0.016	0.003	0.003	0.031	0.010	0.019	0.003	0.023	0.014	0.013	0.022	0.016	0.017	0.020	0.017	0.009	0.018	0.036	0.021	0.061	0.015
dvd vcd	0.020	0.014	0.004	0.002	0.012	0.031	0.019	0.011	0.017	0.009	0.009	0.022	0.010	0.017	0.001	0.017	0.010	0.012	0.016	0.012	0.011	0.015	0.012	0.010	0.003	0.001	0.016	0.022	0.006
müzik kaset	-0.004	-0.005	-0.003	-0.020	-0.006	-0.002	-0.004	-0.005	0.005	-0.004	-0.009	0.001	-0.003	-0.001	-0.003	-0.003	-0.002	0.000	0.016	0.009	0.001	0.013	-0.002	0.010	-0.001	0.026	-0.003	0.008	0.001
hazır tatlı	0.140	0.112	0.025	0.069	0.122	0.107	0.160	0.123	0.082	0.053	0.094	0.171	0.092	0.105	0.030	0.129	0.112	0.094	0.052	0.094	0.088	0.075	0.131	0.043	0.029	-0.007	0.161	0.120	0.021
su maden suyu	0.013	0.011	0.028	-0.015	0.000	0.025	0.005	0.005	0.014	0.006	-0.002	0.005	0.004	0.005	0.001	-0.003	0.004	0.003	0.012	0.008	0.017	0.009	0.012	0.002	-0.004	0.008	0.016	0.025	0.053
hazır corba	0.126	0.121	0.035	0.042	0.088	0.092	0.148	0.101	0.075	0.067	0.097	0.130	0.090	0.072	0.029	0.107	0.090	0.078	0.038	0.079	0.063	0.060	0.113	0.030	0.022	-0.008	0.141	0.090	0.021
sebze meyve	0.216	0.159	0.067	0.184	0.043	0.072	0.135	0.046	0.117	0.156	0.146	0.083	0.056	0.052	0.015	0.057	0.048	0.046	0.035	0.042	0.034	0.042	0.068	0.017	0.011	-0.004	0.094	0.063	0.053
sakiz	0.048	0.036	0.018	0.016	0.033	0.068	0.041	0.030	0.038	0.013	0.019	0.058	0.026	0.033	0.007	0.042	0.033	0.028	0.047	0.052	0.051	0.056	0.051	0.032	0.010	0.036	0.058	0.079	0.031
deoda. kolonya	0.065	0.044	0.016	0.016	0.047	0.044	0.060	0.047	0.038	0.023	0.035	0.102	0.053	0.048	0.017	0.078	0.057	0.055	0.116	0.123	0.075	0.124	0.134	0.084	0.012	0.031	0.096	0.107	0.023
erkek tras	0.059	0.044	0.010	0.011	0.046	0.035	0.057	0.043	0.035	0.024	0.035	0.082	0.048	0.049	0.016	0.070	0.048	0.042	0.065	0.111	0.050	0.142	0.102	0.066	0.013	0.022	0.080	0.071	0.013
sigara	-0.110	-0.072	-0.052	-0.167	-0.052	-0.067	-0.072	-0.043	-0.088	-0.054	-0.074	-0.085	-0.045	-0.042	-0.016	-0.073	-0.046	-0.041	-0.056	-0.077	-0.056	-0.074	-0.084	-0.045	-0.030	-0.025	-0.084	-0.087	-0.066
diger	0.102	0.084	0.039	0.063	0.058	0.066	0.101	0.066	0.064	0.073	0.089	0.167	0.082	0.090	0.024	0.096	0.087	0.082	0.080	0.084	0.069	0.077	0.117	0.046	0.021	0.009	0.133	0.110	0.077
lokum	0.023	-0.010	0.013	0.078	-0.005	-0.004	-0.011	-0.009	0.098	0.005	-0.004	-0.017	-0.008	-0.008	-0.003	-0.015	-0.011	-0.007	-0.001	-0.010	-0.007	-0.005	-0.009	-0.008	-0.009	-0.003	-0.011	-0.001	0.023
prezervatif	0.017	0.014	0.002	0.001	0.009	0.015	0.015	0.012	0.012	0.006	0.008	0.023	0.007	0.013	0.003	0.016	0.010	0.013	0.029	0.024	0.016	0.027	0.028	0.015	0.006	0.017	0.020	0.029	0.004
dergi	-0.025	-0.016	-0.013	-0.045	-0.013	-0.014	-0.016	-0.010	-0.020	-0.012	-0.017	-0.019	-0.010	-0.008	-0.005	-0.017	-0.010	-0.008	-0.010	-0.017	-0.012	-0.016	-0.020	-0.009	-0.007	-0.003	-0.020	-0.020	-0.020
gazete	-0.053	-0.034	-0.024	-0.078	-0.026	-0.034	-0.035	-0.022	-0.041	-0.025	-0.035	-0.041	-0.022	-0.021	-0.007	-0.036	-0.022	-0.021	-0.026	-0.037	-0.028	-0.035	-0.040	-0.022	-0.016	-0.010	-0.042	-0.044	-0.035

Ek Tablo 1. Ürün sınıfları arasındaki ilişkileri gösteren yakınlık matrisi 5/6

	oda parf.	şinek öld.	kitap	ıçki	zuccaciyeye	evcil hayvan	araba bakım	pi	kırasiyeye	pyunçak	dvd vcd	muzik kaset	hazır tatlı	su madden suyu	hazır corba	sebze meyve	şakız	deoda. kolomya	erkek tras	şigara	diğer	lokum	prezervatif	dergi	gazete	
cikolata	0.071	0.017	0.037	-0.012	0.084	0.008	0.012	0.025	0.054	0.032	0.020	-0.008	0.138	0.033	0.112	0.083	0.116	0.073	0.052	-0.156	0.079	0.032	0.016	-0.042	-0.085	
bisküvi	0.059	0.016	0.019	-0.018	0.065	0.002	0.000	0.019	0.039	0.011	0.009	-0.018	0.143	0.031	0.116	0.080	0.082	0.052	0.038	-0.127	0.062	0.002	0.012	-0.043	-0.078	
cips	0.022	0.003	0.010	0.099	0.023	0.004	0.005	0.008	0.018	0.005	0.006	-0.005	0.045	0.025	0.045	0.028	0.055	0.026	0.020	-0.082	0.021	-0.004	0.009	-0.023	-0.044	
bebekmama	0.027	0.015	0.014	-0.007	0.025	-0.002	0.004	0.022	0.014	0.047	0.011	0.000	0.055	0.006	0.031	0.037	0.016	0.012	0.015	-0.044	0.034	-0.007	0.011	-0.011	-0.023	
dondurma	0.021	0.017	0.002	-0.002	0.017	0.005	-0.001	0.005	0.007	-0.003	0.006	-0.003	0.032	0.011	0.038	0.028	0.018	0.015	0.009	-0.037	0.017	-0.010	0.003	-0.011	-0.024	
misirgevrek	0.058	0.013	0.043	0.005	0.053	0.009	0.010	0.021	0.049	0.037	0.028	0.002	0.103	0.015	0.065	0.046	0.054	0.045	0.033	-0.051	0.059	0.000	0.011	-0.010	-0.026	
süt	0.094	0.030	0.046	0.013	0.085	0.018	0.015	0.030	0.062	0.043	0.032	-0.001	0.169	0.032	0.144	0.149	0.079	0.064	0.060	-0.103	0.105	-0.003	0.022	-0.022	-0.050	
mevye suyu	0.056	0.017	0.019	0.012	0.049	0.006	0.008	0.015	0.030	0.011	0.012	-0.006	0.088	0.061	0.115	0.093	0.064	0.043	0.039	-0.100	0.066	0.004	0.014	-0.030	-0.057	
diyet ürünler	0.047	0.013	0.024	0.004	0.045	0.016	0.004	0.012	0.024	0.004	0.010	0.002	0.083	0.020	0.067	0.074	0.044	0.039	0.027	-0.054	0.052	0.000	0.008	-0.010	-0.026	
kahve	0.069	0.022	0.019	0.016	0.063	0.014	0.012	0.017	0.035	0.007	0.014	0.000	0.109	-0.001	0.092	0.045	0.043	0.054	0.044	-0.057	0.062	-0.011	0.008	-0.015	-0.030	
çay	0.092	0.029	0.023	0.012	0.087	0.012	0.016	0.023	0.041	0.009	0.010	-0.004	0.140	-0.001	0.130	0.073	0.041	0.068	0.060	-0.073	0.091	-0.011	0.011	-0.017	-0.035	
enerji içeceği	0.008	0.001	0.004	0.010	0.006	0.002	0.007	0.002	0.007	0.004	0.007	0.000	0.103	0.004	0.121	0.009	0.011	0.022	0.013	0.012	-0.028	0.007	0.013	0.006	-0.006	-0.013
kola	0.038	0.016	-0.001	0.013	0.040	0.004	0.004	0.009	0.016	-0.007	0.001	-0.010	0.073	0.053	0.116	0.079	0.048	0.033	0.028	-0.128	0.046	-0.002	0.006	-0.041	-0.075	
makarna	0.115	0.034	0.031	0.015	0.100	0.010	0.017	0.028	0.058	0.021	0.017	-0.007	0.218	0.009	0.184	0.099	0.053	0.081	0.072	-0.078	0.118	-0.012	0.018	-0.017	-0.038	
tuz	0.068	0.023	0.013	0.004	0.063	0.007	0.009	0.015	0.031	0.009	0.006	-0.004	0.126	0.006	0.100	0.055	0.029	0.049	0.042	-0.042	0.076	-0.008	0.007	-0.011	-0.020	
seker	0.089	0.029	0.017	0.008	0.081	0.012	0.015	0.019	0.039	0.010	0.008	-0.006	0.145	-0.001	0.125	0.071	0.036	0.058	0.058	-0.061	0.088	-0.015	0.011	-0.015	-0.030	
un vb.	0.088	0.022	0.027	0.002	0.080	0.010	0.015	0.023	0.044	0.019	0.014	-0.007	0.308	0.004	0.108	0.071	0.049	0.057	0.051	-0.058	0.095	-0.009	0.013	-0.013	-0.029	
baharat	0.108	0.031	0.027	0.013	0.096	0.016	0.017	0.025	0.050	0.016	0.019	-0.005	0.197	0.003	0.174	0.094	0.046	0.072	0.060	-0.068	0.114	-0.007	0.013	-0.015	-0.034	
pekmez helva	0.026	0.007	0.013	-0.011	0.044	0.004	0.005	0.010	0.024	0.015	0.008	0.002	0.067	0.000	0.040	0.055	0.025	0.022	0.018	-0.066	0.040	0.158	0.002	-0.016	-0.031	
zeytin	0.056	0.026	0.016	0.011	0.064	0.007	0.007	0.015	0.028	0.006	0.010	-0.003	0.086	0.010	0.094	0.153	0.028	0.046	0.040	-0.068	0.072	0.015	0.012	-0.016	-0.033	
zeytinyagi	0.056	0.020	0.014	0.008	0.044	0.008	0.008	0.012	0.022	0.009	0.009	0.001	0.067	0.003	0.064	0.057	0.020	0.037	0.035	-0.029	0.053	-0.001	0.007	-0.006	-0.014	
ayccicek	0.073	0.024	0.015	0.012	0.060	0.007	0.011	0.017	0.032	0.009	0.008	-0.005	0.122	0.002	0.113	0.062	0.027	0.050	0.049	-0.050	0.075	-0.011	0.009	-0.012	-0.024	
reçel bal	0.048	0.018	0.020	0.005	0.050	0.007	0.010	0.014	0.028	0.013	0.013	0.003	0.074	0.007	0.067	0.077	0.025	0.035	0.039	-0.043	0.059	0.017	0.006	-0.009	-0.021	
margarin	0.091	0.029	0.020	0.010	0.071	0.008	0.017	0.022	0.045	0.011	0.012	-0.006	0.176	0.001	0.136	0.083	0.044	0.061	0.057	-0.064	0.093	-0.011	0.015	-0.015	-0.031	
yumurta	0.062	0.022	0.023	0.019	0.060	0.014	0.012	0.019	0.032	0.017	0.015	-0.004	0.106	0.009	0.113	0.172	0.035	0.041	0.044	-0.065	0.081	-0.003	0.017	-0.014	-0.030	
kuryemisi	0.010	0.003	0.005	0.062	0.004	0.005	-0.004	0.001	0.004	-0.010	0.003	0.000	0.017	0.003	0.111	0.027	0.021	0.009	0.012	-0.036	0.010	-0.002	0.004	-0.005	-0.019	
bakliyat	0.105	0.028	0.027	0.010	0.086	0.014	0.016	0.023	0.053	0.017	0.014	-0.005	0.179	0.006	0.145	0.096	0.044	0.067	0.063	-0.062	0.107	-0.009	0.017	-0.013	-0.030	
ekmek	0.044	0.013	0.022	0.011	0.045	0.021	0.004	0.014	0.025	0.010	0.017	0.004	0.081	0.005	0.056	0.104	0.032	0.027	0.023	-0.046	0.052	0.009	0.009	-0.006	-0.021	
dondurulmus	0.054	0.010	0.015	0.013	0.040	0.009	0.008	0.014	0.022	0.011	0.014	-0.001	0.092	0.001	0.057	0.071	0.022	0.030	0.027	-0.035	0.054	-0.005	0.009	-0.007	-0.017	
peynir	0.092	0.031	0.032	0.020	0.092	0.020	0.010	0.022	0.045	0.017	0.020	-0.004	0.140	0.013	0.126	0.216	0.048	0.065	0.059	-0.110	0.102	0.023	0.017	-0.025	-0.053	
yogurt	0.070	0.028	0.021	0.023	0.060	0.017	0.009	0.019	0.034	0.010	0.014	-0.005	0.112	0.011	0.121	0.159	0.036	0.044	0.044	-0.072	0.084	-0.010	0.014	-0.016	-0.034	
ayran	0.018	0.009	0.005	0.000	0.024	0.003	-0.001	0.003	0.009	0.001	0.004	-0.003	0.025	0.028	0.035	0.067	0.018	0.016	0.010	-0.052	0.039	0.013	0.002	-0.013	-0.024	
unlu-hazır yemek	0.025	0.006	0.002	-0.006	0.051	0.007	-0.011	0.008	0.015	-0.007	0.002	-0.020	0.069	-0.015	0.042	0.184	0.016	0.016	0.011	-0.167	0.063	0.078	0.001	-0.045	-0.078	
sürme cikolata	0.061	0.021	0.019	0.003	0.055	0.004	0.013	0.018	0.040	0.016	0.012	-0.006	0.122	0.000	0.088	0.043	0.033	0.047	0.046	-0.052	0.058	-0.005	0.009	-0.013	-0.026	
çocuk gel. yog	0.064	0.022	0.039	0.005	0.060	0.002	0.012	0.024	0.052	0.060	0.031	-0.002	0.107	0.025	0.092	0.072	0.068	0.044	0.035	-0.067	0.066	-0.004	0.015	-0.014	-0.034	
salça konserve	0.098	0.027	0.025	0.026	0.080	0.016	0.013	0.023	0.046	0.014	0.019	-0.004	0.160	0.005	0.148	0.135	0.041	0.060	0.057	-0.072	0.101	-0.011	0.015	-0.016	-0.035	
ketçap mayonez	0.070	0.017	0.019	0.012	0.054	0.006	0.011	0.017	0.032	0.013	0.011	-0.005	0.123	0.005	0.101	0.046	0.030	0.047	0.043	-0.043	0.066	-0.009	0.012	-0.010	-0.022	
sucuk sosis	0.051	0.021	0.022	0.011	0.055	0.010	0.007	0.016	0.033	0.016	0.017	0.005	0.082	0.014	0.075	0.117	0.038	0.038	0.035	-0.088	0.064	0.098	0.012	-0.020	-0.041	
et kıyma	0.042	0.021	0.012	0.027	0.044	0.014	0.003	0.009	0.018	0.003	0.009	-0.004	0.053	0.006	0.067	0.156	0.013	0.023	0.024	-0.054	0.073	0.005	0.006	-0.012	-0.025	
tavuk	0.058	0.018	0.016	0.021	0.054	0.013	0.007	0.015	0.030	0.003	0.009	-0.009	0.094	-0.002	0.097	0.146	0.019	0.035	0.035	-0.074	0.089	-0.004	0.008	-0.017	-0.035	
genel temizlik	0.178	0.046	0.040	0.005	0.127	0.021	0.030	0.031	0.067	0.031	0.022	0.001	0.171	0.005	0.130	0.083	0.058	0.102	0.082	-0.085	0.167	-0.017	0.023	-0.019	-0.041	
bulaşık el	0.085	0.030	0.012	0.007	0.064	0.007	0.012	0.015	0.032	0.010	0.010	-0.003	0.092	0.004	0.090	0.056	0.026	0.053	0.048	-0.045	0.082	-0.008	0.007	-0.010	-0.022	
bulaşık mak	0.105	0.021	0.024	0.005	0.062	0.011	0.013	0.021	0.039	0.019	0.017	-0.001	0.105	0.005	0.072	0.052	0.033	0.048	0.049	-0.042	0.090	-0.008	0.013	-0.008	-0.021	
camasır el	0.019	0.012	0.004	0.000	0.019	0.002	0.002	0.005	0.010	0.003	0.001	-0.003	0.030	0.001	0.029	0.015	0.007	0.017	0.016	-0.016	0.024	-0.003	0.003	-0.005	-0.007	
camasır mak	0.113	0.032	0.029	0.004	0.078	0.012	0.018	0.024	0.048	0.023	0.017	-0.003	0.129	-0.003	0.107	0.057	0.042	0.078	0.070	-0.073	0.096	-0.015	0.016	-0.017	-0.036	



Ek Tablo 1. Ürün sınıfları arasındaki ilişkileri gösteren yakınlık matrisi 6/6

	oda parf.	sinek öld.	kitap	icki	züccaciye	evcil hayvan	araba bakım	pil	kirtasiye	oyuncak	dvd vcd	müzik kaset	hazır tatlı	su maden suyu	hazır corba	sebze meyve	sakız	deoda. kolonya	erkek tras	sigara	diğer	lokum	prezervatif	dergi	gazete
camasır suyu	0.112	0.028	0.020	0.004	0.065	0.013	0.011	0.017	0.035	0.014	0.010	-0.002	0.112	0.004	0.090	0.048	0.033	0.057	0.048	-0.046	0.087	-0.011	0.010	-0.010	-0.022
yumusatıcı	0.100	0.025	0.021	0.006	0.059	0.011	0.013	0.012	0.033	0.013	0.012	0.000	0.094	0.003	0.078	0.046	0.028	0.055	0.042	-0.041	0.082	-0.007	0.013	-0.008	-0.021
makyaj	0.053	0.022	0.035	0.001	0.053	0.017	0.011	0.019	0.037	0.022	0.016	0.016	0.052	0.012	0.038	0.035	0.047	0.116	0.065	-0.056	0.080	-0.001	0.029	-0.010	-0.026
sampuan	0.082	0.028	0.029	0.005	0.066	0.011	0.017	0.019	0.043	0.016	0.012	0.009	0.094	0.008	0.079	0.042	0.052	0.123	0.111	-0.077	0.084	-0.010	0.024	-0.017	-0.037
kadın pedi	0.060	0.019	0.025	0.004	0.054	0.011	0.009	0.015	0.041	0.017	0.011	0.001	0.088	0.017	0.063	0.034	0.051	0.075	0.050	-0.056	0.069	-0.007	0.016	-0.012	-0.028
agız bakım	0.066	0.024	0.036	0.002	0.056	0.013	0.017	0.021	0.045	0.020	0.015	0.013	0.075	0.009	0.060	0.042	0.056	0.124	0.142	-0.074	0.077	-0.005	0.027	-0.016	-0.035
sabun-dus	0.107	0.032	0.029	0.006	0.094	0.015	0.018	0.022	0.051	0.017	0.012	-0.002	0.131	0.012	0.113	0.068	0.051	0.134	0.102	-0.084	0.117	-0.009	0.028	-0.020	-0.040
sac boya jöle	0.043	0.020	0.018	0.003	0.034	-0.008	0.013	0.010	0.026	0.009	0.010	0.010	0.043	0.002	0.030	0.017	0.032	0.084	0.066	-0.045	0.046	-0.008	0.015	-0.009	-0.022
bebek temizlik	0.021	0.006	0.006	-0.003	0.018	-0.002	0.005	0.008	0.008	0.018	0.003	-0.001	0.029	-0.004	0.022	0.011	0.010	0.012	0.013	-0.030	0.021	-0.009	0.006	-0.007	-0.016
günes süt krem	0.002	0.031	0.013	-0.008	-0.002	-0.001	0.000	0.004	0.003	0.036	0.001	0.026	-0.007	0.008	-0.008	-0.004	0.036	0.031	0.022	-0.025	0.009	-0.003	0.017	-0.003	-0.010
kagıt temizlik	0.135	0.036	0.033	0.007	0.104	0.012	0.020	0.026	0.058	0.021	0.016	-0.003	0.161	0.016	0.141	0.094	0.058	0.096	0.080	-0.084	0.133	-0.011	0.020	-0.020	-0.042
bebek malz	0.096	0.032	0.047	-0.003	0.091	0.007	0.017	0.032	0.064	0.061	0.022	0.008	0.120	0.025	0.090	0.063	0.079	0.107	0.071	-0.087	0.110	-0.001	0.029	-0.020	-0.044
plastik ahsap	0.028	0.012	0.012	-0.003	0.091	0.006	0.009	0.010	0.027	0.015	0.006	0.001	0.021	0.053	0.021	0.053	0.031	0.023	0.013	-0.066	0.077	0.023	0.004	-0.020	-0.035
oda parf.	1.000	0.033	0.032	0.007	0.067	0.014	0.057	0.024	0.047	0.020	0.022	0.012	0.089	0.001	0.066	0.044	0.041	0.068	0.049	-0.051	0.098	-0.005	0.018	-0.011	-0.026
sinek öld.	0.033	1.000	0.013	0.003	0.019	0.007	0.005	0.011	0.016	0.014	0.004	0.007	0.016	0.006	0.022	0.019	0.018	0.027	0.020	-0.022	0.039	-0.002	0.006	-0.004	-0.010
kitap	0.032	0.013	1.000	-0.007	0.040	0.010	0.016	0.016	0.087	0.055	0.043	0.029	0.028	0.004	0.010	0.017	0.038	0.033	0.021	-0.062	0.041	0.005	0.010	-0.007	-0.027
icki	0.007	0.003	-0.007	1.000	-0.005	0.008	0.001	-0.002	-0.009	-0.018	-0.001	-0.004	-0.001	-0.018	0.009	0.021	-0.007	-0.001	0.006	-0.049	0.002	-0.031	0.004	-0.016	-0.028
züccaciye	0.067	0.019	0.040	-0.005	1.000	0.008	0.016	0.023	0.054	0.044	0.021	0.007	0.076	0.009	0.050	0.059	0.036	0.058	0.032	-0.090	0.114	0.024	0.009	-0.021	-0.043
evcil hayvan	0.014	0.007	0.010	0.008	0.008	1.000	0.006	0.006	0.015	0.000	0.006	0.001	0.012	-0.003	0.008	0.017	0.006	0.010	0.005	-0.016	0.019	-0.002	0.000	-0.001	-0.007
araba bakım	0.057	0.005	0.016	0.001	0.016	0.006	1.000	0.013	0.027	0.014	0.009	0.017	0.014	-0.004	0.006	0.003	0.015	0.019	0.020	-0.049	0.030	-0.004	0.008	-0.013	-0.025
pil	0.024	0.011	0.016	-0.002	0.023	0.006	0.013	1.000	0.021	0.059	0.015	0.009	0.020	-0.001	0.015	0.012	0.019	0.017	0.017	-0.029	0.029	-0.002	0.007	-0.007	-0.014
kirtasiye	0.047	0.016	0.087	-0.009	0.054	0.015	0.027	0.021	1.000	0.045	0.029	0.011	0.048	0.010	0.029	0.020	0.042	0.043	0.028	-0.064	0.076	0.006	0.007	-0.011	-0.031
oyuncak	0.020	0.014	0.055	-0.018	0.044	0.000	0.014	0.059	0.045	1.000	0.045	0.034	0.019	-0.002	-0.005	0.004	0.035	0.018	0.008	-0.081	0.043	0.014	0.014	-0.018	-0.039
dvd vcd	0.022	0.004	0.043	-0.001	0.021	0.006	0.009	0.015	0.029	0.045	1.000	0.061	0.017	-0.002	0.009	0.017	0.022	0.011	0.010	-0.036	0.022	0.000	0.005	-0.002	-0.016
müzik kaset	0.012	0.007	0.029	-0.004	0.007	0.001	0.017	0.009	0.011	0.034	0.061	1.000	-0.005	-0.004	-0.012	-0.001	0.028	0.015	0.008	-0.053	0.004	0.014	0.005	-0.006	-0.021
hazır tatlı	0.089	0.016	0.028	-0.001	0.076	0.012	0.014	0.020	0.048	0.019	0.017	-0.005	1.000	-0.001	0.119	0.067	0.046	0.056	0.047	-0.060	0.091	-0.006	0.014	-0.013	-0.029
su maden suyu	0.001	0.006	0.004	-0.018	0.009	-0.003	-0.004	-0.001	0.010	-0.002	-0.002	-0.004	-0.001	1.000	0.043	0.027	0.060	0.010	0.003	-0.113	0.006	0.011	0.005	-0.039	-0.069
hazır corba	0.066	0.022	0.010	0.009	0.050	0.008	0.006	0.015	0.029	-0.005	0.009	-0.012	0.119	0.043	1.000	0.075	0.042	0.042	0.044	-0.060	0.064	-0.023	0.008	-0.026	-0.051
sebze meyve	0.044	0.019	0.017	0.021	0.059	0.017	0.003	0.012	0.020	0.004	0.017	-0.001	0.067	0.027	0.075	1.000	0.033	0.029	0.027	-0.115	0.072	0.039	0.008	-0.026	-0.052
sakız	0.041	0.018	0.038	-0.007	0.036	0.006	0.015	0.019	0.042	0.035	0.022	0.028	0.046	0.060	0.042	0.033	1.000	0.059	0.032	-0.090	0.040	0.024	0.021	-0.024	-0.048
deoda. kolonya	0.068	0.027	0.033	-0.001	0.058	0.010	0.019	0.017	0.043	0.018	0.011	0.015	0.056	0.010	0.042	0.029	0.059	1.000	0.096	-0.074	0.072	0.003	0.030	-0.016	-0.036
erkek tras	0.049	0.020	0.021	0.006	0.032	0.005	0.020	0.017	0.028	0.008	0.010	0.008	0.047	0.003	0.044	0.027	0.032	0.096	1.000	-0.055	0.050	-0.007	0.029	-0.013	-0.027
sigara	-0.051	-0.022	-0.062	-0.049	-0.090	-0.016	-0.049	-0.029	-0.064	-0.081	-0.036	-0.053	-0.060	-0.113	-0.060	-0.115	-0.090	-0.074	-0.055	1.000	-0.078	-0.095	-0.025	-0.014	-0.005
diğer	0.098	0.039	0.041	0.002	0.114	0.019	0.030	0.029	0.076	0.043	0.022	0.004	0.091	0.006	0.064	0.072	0.040	0.072	0.050	-0.078	1.000	0.007	0.021	-0.020	-0.045
lokum	-0.005	-0.002	-0.005	-0.031	0.024	-0.002	-0.004	-0.002	0.006	0.014	0.000	0.014	-0.006	0.011	-0.023	0.039	0.024	0.003	-0.007	-0.095	0.007	1.000	-0.002	-0.025	-0.044
prezervatif	0.018	0.006	0.010	0.004	0.009	0.000	0.008	0.007	0.007	0.014	0.005	0.005	0.014	0.005	0.008	0.008	0.021	0.030	0.029	-0.025	0.021	-0.002	1.000	-0.006	-0.012
dergi	-0.011	-0.004	-0.007	-0.016	-0.021	-0.001	-0.013	-0.007	-0.011	-0.018	-0.002	-0.006	-0.013	-0.039	-0.026	-0.026	-0.024	-0.016	-0.013	-0.014	-0.020	-0.025	-0.006	1.000	0.041
gazete	-0.026	-0.010	-0.027	-0.028	-0.043	-0.007	-0.025	-0.014	-0.031	-0.039	-0.016	-0.021	-0.029	-0.069	-0.051	-0.052	-0.048	-0.036	-0.027	-0.005	-0.045	-0.044	-0.012	0.041	1.000

**Ek Şekil 1.** Ürün sınıfları arasındaki ilişkiyi gösteren dendogram grafiği

