

T.C.
MUĞLA ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ

KİMYA ANABİLİM DALI

OBJELERİN BENZER YA DA FARKLI
ÖZELLİKLERİNE GÖRE SINIFLANDIRILMASINDA
KÜMELEME TEKNİKLERİNİN KULLANILMASI

YÜKSEK LİSANS TEZİ

SİNEM ÇAĞLAR ODABAŞ

HAZİRAN 2012

MUĞLA

T.C.
MUĞLA ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ

KİMYA ANABİLİM DALI

OBJELERİN BENZER YA DA FARKLI
ÖZELLİKLERİNE GÖRE SINIFLANDIRILMASINDA
KÜMELEME TEKNİKLERİNİN KULLANILMASI

YÜKSEK LİSANS TEZİ

SİNEM ÇAĞLAR ODABAŞ

HAZİRAN 2012

MUĞLA

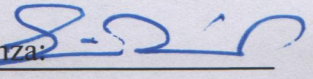
MUGLA ÜNİVERSİTESİ
Fen Bilimleri Enstitüsü

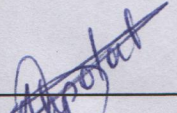
TEZ ONAYI

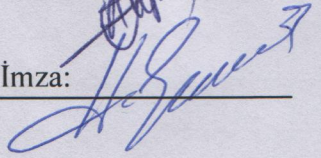
SİNEM ÇAĞLAR ODABAŞ tarafından hazırlanan **OBJELERİN BENZER YA DA FARKLI ÖZELLİKLERİNE GÖRE SINIFLANDIRILMASINDA KÜMELEME TEKNİKLERİNİN KULLANILMASI BAŞLIKLİ** tezinin, 01/06/2012 tarihinde aşağıdaki jüri tarafından Kimya Anabilim Dalı'nda yüksek lisans derecesi için gerekli şartları sağladığı oybirliği/oyçokluğu ile kabul edilmiştir.

TEZ SINAV JURİSİ

Doç. Dr. Mehmet Emin Duru (Jüri Başkanı)
Kimya Anabilim Dalı, Muğla Üniversitesi, Muğla
Yrd. Doç. Dr. Oğuz Akpolat (Üye)
Kimya Anabilim Dalı, Muğla Üniversitesi, Muğla
Yrd. Doç. Dr. Hasan Ertas (Üye)
Kimya Anabilim Dalı, Ege Üniversitesi, İzmir

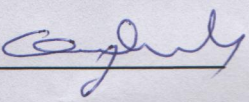
İmza: 

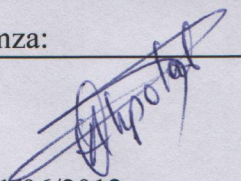
İmza: 

İmza: 

BÖLÜM BAŞKANLIĞI VE FEN BİLİMLERİ ENSTİTÜSÜ ONAYI

Prof. Dr. Gül Asiye Ayçık
Kimya Anabilim Dalı Bölüm Başkanı, Muğla Üniversitesi, Muğla
Yrd. Doç. Dr. Oğuz AKPOLAT
Danışman, Kimya Anabilim Dalı Muğla Üniversitesi, Muğla

İmza: 

İmza: 

Savunma Tarihi: 01/06/2012

Tez çalışmalarım sırasında elde ettiğim ve sunduğum tüm sonuç, doküman, bilgi ve belgelerin tarafımdan bizzat ve bu tez çalışması kapsamında elde edildiğini; akademik ve bilimsel etik kurallarına uygun olduğunu beyan ederim. Ayrıca, akademik ve bilimsel etik kuralları gereği bu tez çalışması sırasında elde edilmemiş başkalarına ait tüm orijinal bilgi ve sonuçlara atıf yapıldığını da beyan ederim.

Sinem Çağlar Odabaş

01/06/2012



ÖZET
OBJELERİN BENZER YA DA FARKLI ÖZELLİKLERİNE GÖRE
SINIFLANDIRILMASINDA KÜMELEME TEKNİKLERİNİN
KULLANILMASI

Sinem Çağlar ODABAŞ

Yüksek Lisans Tezi

Kimya Anabilim Dalı

Danışman: Yrd. Doç. Dr. Oğuz AKPOLAT

Haziran 2012, 92 sayfa

Bu çalışmada sığla, sigala veya günlük ağacı olarak adlandırılan *Liquidambar orientalis* Mill. Bitkisinin yapraklarından elde edilen uçucu yağ bileşenleri incelenmiştir. Bu amaçla Marmaris, Fethiye, Köyceğiz olmak üzere üç farklı bölgeden Temmuz ayında toplanan bitki materyallerinin ekstraktları hidrodestilasyon metodu ile ayrıştırıldıktan sonra üç farklı bölgeden toplanan örneklerin uçucu yağ bileşenleri GC-MS analizi ile belirlenerek karışımların yaklaşık %70-80’lik büyük kısmı 11 major bileşenin oluşturduğu tespit edilmiştir.

Kromatografik analizle bulunan bileşen dağılımlarının istatistiksel analizinde kümeleme yöntemi kullanılmıştır. Kümeleme yöntemi SPSS ve MATLAB yazılımları kullanılarak gerçekleştirilmiştir.

Bileşenler açısından incelendiğinde Fethiye ve Marmaris’ten gelen örneklerin bir küme oluşturacak şekilde çok benzer özellikler taşıdığı, Köyceğiz’den gelenlerin ise farklı bir küme oluşturduğu saptanmıştır.

Anahtar Kelimeler: *Liquidambar orientalis* Miller, uçucu yağ, SPSS, MATLAB, kümeleme analizi

ABSTRACT
USE OF CLUSTURING TECHNICS SIMILAR OR DIFFERENT IN TERMS
OF THE SPECIFICATIONS OF OBJECTS

Sinem Çağlar ODABAŞ

Master of Science (M.Sc)

Department of Chemistry

Supervisor: Asst. Yrd. Doç. Dr. Oğuz AKPOLAT

June 2012, 92 pages

In this study essential oil components were reviewed that obtained from plant leaves of *Liquidambar orientalis* Milller called as sığla, sigala or günlük tree.

For this purpose, after extraction by the method of hydrodistillation Extracts of plant materials collected from three different regions including Marmaris, Fethiye, Köyceğiz in July, approximately 70-80% greater part of the mixtures that essential oil components of the samples from three different regions, determined by GC-MS analysis were formed by the 11 major component.

Clustering method was used at statistical analysis of the component distributions found with the chromatographic analysis. Clustering method was carried out using the software SPSS and MATLAB.

In terms of components, determined as a cluster of samples came from Fethiye and Marmaris have many same characteristics as for came from Köyceğiz generate a different cluster.

Keywords: *Liquidambar orientalis* Milller, essential oil, SPSS, MATLAB, cluster analysis

ÖNSÖZ

Tüm çalışma sürecim boyunca bilimsel katkılarıyla bana destek olan, emeğini ve kıymetli zamanını esirgemedi yüksek lisans eğitimim süresince her konuda yardımcı olan tez danışmanım Yrd. Doç. Dr. Oğuz Akpolat'a saygı ve teşekkürlerimi sunarım.

Tezimi hazırlarken bilgi birikimini benden esirgemeyen Prof. Dr. Mehmet Emin DURU' ya,

Laboratuvar çalışmalarım boyunca bilgi ve fikirleriyle bana yardımcı olan Doç. Dr. Mehmet ÖZTÜRK ve Arş. Gör. Gülsen TEL'e, teşekkür ederim.

Dostluğu ve bilgileriyle bana destek olan Gülçin ÖZEVCI' ye,

Çalışmamın tüm aşamalarında ilgisi sevgisi ve desteğiyle hep yanımda olan canım annem ve babam Emine ve Ali ODABAŞ' a, bana her zaman pozitif motivasyon sağlayan ablam ve eniştem Gizem ve Mehmet ÇEKİM' e, ve yoldaşım Ceyhun ÇINAR' a çok teşekkür ederim.

Bu tez çalışması, BAP-2011-50 numaralı Bilimsel Araştırma Projesiyle desteklenmiştir.

İÇİNDEKİLER

ÖNSÖZ	vi
İÇİNDEKİLER	vii
ÇİZELGELER DİZİNİ	ix
ŞEKİLLER DİZİNİ	x
KISALTMALAR DİZİNİ	xii
1. GİRİŞ	1
1.1. Kimyada Sayısal Verilerin Analizi	4
1.1.1. İstatistiğin Tanımı ve Anlamı	4
1.1.2. Veri Madenciliği	5
1.1.3. Kümeleme Teknikleri.....	6
1.1.4. Bilgisayar Programları	7
1.1.5. Kümeleme Çalışmaları.....	8
1.2. Sığla Ağacı	11
1.3. Uçucu Yağlar	12
1.3.1. Fiziksel Özellikleri	13
1.3.2. Uçucu Yağların Sınıflandırılması	14
1.3.3. Uçucu Yağların Elde Edilme Yöntemleri	15
1.3.3.1. Destilasyon Yöntemi	17
1.3.3.2. Mekanik Yöntem (presleme yoluyla uçucu yağ elde edilmesi)	18
1.3.3.3. Anfloranj Yöntemi (ekstraksiyon yoluyla uçucu yağ elde edilmesi) ..	18
1.3.4. Uçucu Yağ Eldesindeki Değişmeler	19
1.3.5. Uçucu Yağlarda Bulunan Bileşiklere Örnekler.....	20
1.3.6. Uçucu Yağların Analizi.....	21
2. MALZEME VE YÖNTEM	22
2.1. Deneysel Kısım	22
2.1.1. Kullanılan Bitkisel Materyal, Kimyasal Maddeler ve Aletler	22
2.1.1.1 Bitkisel Materyal	22
2.1.1.2 Çalışmada Kullanılan Kimyasal Maddeler.....	23
2.1.1.3 Aletler	23
2.1.2. Metot	23
2.1.2.1. Clavenger Cihazı ile Hidrodestilasyon	23
2.1.3. GC-MS Sonuçları.....	24

2.2. İstatistiksel Değerlendirme.....	27
2.2.1. Kümeleme Analizinin Yapılması.....	27
2.2.2. En Yakın Komşu Kümeleme Algoritması	27
2.2.3. Bilgisayarda Programlama	28
2.2.3.1. SPSS ile Kümeleme Analizi	28
2.2.3.1. MATLAB ile Kümeleme Analizi.....	29
3. BULGULAR VE İRDELEME.....	31
3.1. Gaz Kromatografisi Kütle Spektroskopisi Analizleri	31
3.2. İstatistiksel Değerlendirme.....	37
3.2.1. SPSS İle Kümeleme Analizi	37
3.2.2. MATLAB İle Kümeleme Analizi	39
4. SONUÇLAR VE ÖNERİLER	43
KAYNAKLAR	46
EKLER.....	51
Ek 1. İstatistiğe İlişkin Bazı Temel Tanım Ve Kavramlar.....	53
Ek 2. Veri Madenciliği.....	60
Ek 3. Kümeleme Analizi Ve Teknikleri.....	63
Ek 4. MATLAB İle Programlama.....	79
Ek 5. SPSS İle Programlama.....	87
ÖZGEÇMİŞ.....	91

ŞEKİLLER DİZİNİ

Şekil 1.1. Terpenlerin sınıflandırılması ve bazı monoterpen, seskuitepen ve triterpen molekülleri.....	16
Şekil 2.1. GC tayinleri amacıyla alınan kromatogramlara ilişkin bir pik raporu örneği (LOF1HD00)	25
Şekil 2.2. Majör bileşenlerden α -Pine'e ait kütle spektroskopisi örneği menüleri	26
Şekil 2.3. Kümeler analizi için SPSS140 menüleri.....	29
Şekil 3.1.a. GC tayinleri sonucu alınan kromatogramlara ilişkin pik raporu örneği (LOF1HD00)	33
Şekil 3.1.b. GC tayinleri sonucu alınan kromatogramlara ilişkin pik raporu örneği (LOK1HD00)	34
Şekil 3.1.c. GC tayinleri sonucu alınan kromatogramlara ilişkin pik raporu örneği (LOM1HD00)	35
Şekil 3.2. Üç örneğin kütle spektrumları (LOF1HD00, LOK1HD00 ve LOM1HD00)	36
Şekil 3.3. Üç örneğin kütle spektrumları (LOF1HD00, LOK1HD00 ve LOM1HD00) üzerinde majör bileşenlerin ayrıntılı görünümü.....	36
Şekil 3.4. Üç örneğin (LOF1HD00 (1), LOK1HD00 (2) ve LOM1HD00 (3)) majör bileşenlerinin dağılımına göre yakınlık ilişkilerini veren kümeleme analizi	42
Şekil Ek 1.1. Çürük dış ölçümsel ve teorik frekans (sıklık) grafiği.....	56
Şekil Ek 1.2. Doğrusal modelli bir regresyon korelasyon eğrisi	59
Şekil Ek 2.1. Veri Madenciliği Aşamaları	62
Şekil Ek 3.1. Kümeleme İşleminin Adımları	63
Şekil Ek 3.2. Bütünleyici algoritmaya göre kümeleme işlemi.....	64
Şekil Ek 3.3. Kümeleme Tekniklerinin Genel Sınıflandırılması	65
Şekil Ek 3.4.a. Hiyerarşik yapıda kümeleri iki boyutlu düzlemde gösterme.....	66
Şekil Ek 3.4.b. Hiyerarşik yapıda kümeleri birleştirme ya da ayırıştırma	66
Şekil Ek 3.4.c. Hiyerarşik yapıda kümeleri birleştirme ağacı (dengdongram)	66
Şekil Ek 3.5. Koordinat sisteminde iki nokta arasındaki uzaklık	68
Şekil Ek 3.6. Öklit uzaklık ölçütü	69
Şekil Ek 3.7. Tek-Bağ kümelemede iki küme arası uzaklık	73
Şekil Ek 3.8. Kümeleme analizi için SPSS menü seçenekleri	77

Şekil Ek 3.9. Kümeleme analizi sonucu SPSS te oluşturulan dendogram.....	78
Şekil Ek 4.1. MATLAB Arayüzü	78
Şekil Ek 4.2. M-File Arayüzü	78
Şekil Ek 4.3. Figure Arayüzü	79
Şekil Ek 5.1.Verilerin istatistiksel değerlendirilmesi için SPSS11.5 te yazılmış kayıt dosyası	88
Şekil Ek 5.2. Verilerin istatistiksel değerlendirilmesi için SPSS11.5 ta var olan menü alternatifleri	90

ÇİZELGELER DİZİNİ

Çizelge 2.1. Örneklerin bileşen dağılımları ve geldikleri yöreler.....	29
Çizelge 3.1. Örneklerle göre major bileşenlerin dağılımı	32
Çizelge 3.2. SPSS14 veri dosyası	37
Çizelge 3.3. Verilerin saklandığı EXCEL dosyası.....	39
Çizelge Ek 1.1. İstatistiksel değerlendirme İçin uygun test seçim klavuzu.....	54
Çizelge Ek 1.2. Student-T dağılışına ilişkin kritik değerler (tek yönlü)	58

KISALTMALAR DİZİNİ

KDD	Knowledge Discovery in Databases
VTBK	Veri madenciliği birçok kaynakta Veri Tabanlarında Bilgi Keşfi
RLS	Restless Legs Syndrome
MDS	Çok Boyutlu Ölçekleme
PCA	Temel Bileşen Analizi
HCA	Hierarchical Kümeleme Analizi
LOF1HD00	Fethiye Bölgesinden Alınan Örnek
LOK1HD00	Köyceğiz Bölgesinden Alınan Örnek
LOM1HD00	Marmaris Bölgesinden Alınan Örnek
GC	Gaz Kromatografisi
MS	Kütle Spektrometresi Sistemi
GC/MS	Gaz Kromatografisi-Kütle Spektrometresi Sistemi
HD	Su Destilasyonu (Hydrodistillation)
mg	Miligram
μ L	Mikrolitre
mL	Mililitre
L	Litre

1. GİRİŞ

Veri tabanlarında farklı özelliklerdeki belli bir anlamı olmadan kayıtlı olan verilerin yorumlanarak anlamlı ve kullanışlı bilgiler haline getirilmesi veri madenciliğinin temelini oluşturmaktadır. Bilimsel kaynaklarda “Veri Madenciliği” teriminin farklı tanımları ile karşılaşılmaktadır. Bunlardan biri veri ambarlarındaki büyük veri setlerinin yapay zeka ve istatistiksel matematik tekniklerle işlenerek uyumlu ilinti ağlarının belirlenmesi iken, bir diğeri büyük miktardaki verilerin otomatik veya yarı-otomatik yöntemlerle analiz edilerek anlamlı örüntü ve kuralların bulunmasıdır. Veri madenciliği kullanımının yaygınlaşmasının nedenleri ise teknolojinin ve bilimin ilerlemesiyle gereksinim duyulan ve saklanan veri boyutunun her geçen gün artması ve bu verilerin anlamlandırılmadan sadece saklanması bilimsel yorum getirme, geleceğe yönelik tahminlerde bulunma, sınıflandırma, kümeleme ve tanımlama gibi konularda yeterli olmamaktadır.

Sınıflandırma veri madenciliğinin en çok kullanıldığı alandır. Amacı yeni bir nesnenin özelliklerini açıklamak ve bu yeni nesnenin daha önceden tanımlanmış sınıf dizilerinden birine atamasını yapmaktır. Sınıflandırmanın yapılabilmesi için sınıflandırılmamış veriye uygun bir model uygulanabilir. Veri madenciliğinin sınıflandırma grubu içerisinde en sık kullandığı teknik de karar ağaçlarıdır. Aynı zamanda lojistik regresyon, diskriminant analizi, sinir ağları ve bulanık mantık setleri de kullanılmaktadır. İnsanlar verileri daima sınıflandırdıkları, kategorize ettikleri ve derecelendirdikleri için sınıflandırma, hem veri madenciliğinin temeli olarak hem de veri hazırlama aracı olarak da kullanılabilir. Sınıflandırmada bir başka yaklaşım olarak ele alınan kümelemenin amacı ise heterojen olan ana kütleyi homojen gruplara ayırmaktır. Kümeleme önceden tanımlanmış sınıfları esas alarak gruplamaz yani kümeleme gruplamada benzerlikleri kullanırken sınıflandırma önceden tanımlanmış sınıflar modelini temel alır. Ayrıca kümeleme, veri madenciliği sürecinden önce veri hazırlama esnasında da kullanılabilir.

Kaynaklarda objelerin birbirlerine benzeyen ya da farklı olan özelliklerine göre gruplandırılması kümeleme olarak adlandırılmaktadır. Benzer ya da farklı özelliklerin öbeklendirilebilmesi için öncelikle veri tabanlarında birleştirilmesi söz konusudur. Bu benzerlik ya da farklılıklar bir veya daha fazla boyutta olabilir ve bunlar uzaklık cinsinden metrik olarak ölçümlendirilebilir. Değişik kaynaklarda kümeleme metotları farklı biçimlerde sınıflandırılmaktadır. Yapılan herhangi bir araştırma sonrasında elde edilen verilerin analizinde ham veri çok büyük öneme sahiptir ve bu verilerin değerlendirilmelerine geçmeden önce, özellikleri dikkatle incelenmeli ve elde edilen veriler arasında birbirinden çok farklı değerler bulunuyorsa ya da değerler arası büyük dengesizlikler varsa, verilerin benzer birimlere çevrilmeleri veya ortalama değer etrafında toplanması gerekmektedir. Bunun gibi durumlarda verilere merkezileştirme, standartlaştırma ve normalizasyon gibi verilerin analize hazırlanmasına yönelik işlemler uygulanır. Gelişen bilgisayar teknolojisi ve hazır istatistik paket programlarıyla bu konuda yapılacak olan sayısal işlemler oldukça kolaylaşmıştır. Bu amaca yönelik olarak da çok sayıda paket programda geliştirilmiştir ve bunlardan en çok bilinenleri ise SAS, SPSS, MATLAB, MINITAB, JMP, NTSYS, SYSTAT, MEGA, POPGEN ve MVSP' dir.

Günlük yaşamda ise ürünlerin pazarlanması ya da hizmetlerin sunulması gibi pratik alanlarda veya yapılan bilimsel ve teknolojik araştırmalarda toplanan verilerin değerlendirilmesinde karşılaşılan kümeleme tekniklerinin, gerek fizik, kimya ve biyoloji gibi temel bilimlerde gerekse tıp, mühendislik, nanoteknoloji, bilgi teknolojileri, genetik, çevre ya da biyoteknoloji gibi pek çok uygulamalı bilimde, madenlerin, ürünlerin, canlıların fiziksel ya da kimyasal özelliklerine göre orijinlerin belirlenmesi, yapılarının gruplanması, özelliklerinin zamana bağlı olarak değişimlerinin incelenmesi gibi alanlarda çok geniş kullanıma sahip olduğu pek çok kaynaktan açıkça anlaşılmaktadır. Kimya alanında da benzer olarak, özellikle yapılan kemometrik analizlerde sağlanan ham veri artık hem çok miktarda hem de büyük öneme sahip olmaktadır, ve yapılacak olan bir deneysel çalışma sonrasında elde edilen verilerin hesaplamalarına ve değerlendirilmelerine geçmeden önce, özelliklerinin dikkatle incelenmesi, anlaşılır ve karşılaştırılabilir olması için istatistiksel olarak araştırılması, ve veri madenciliği ilkeleri çerçevesinde bütünleştirilmesi gerekmektedir.

Bu çalışmanın amacı da; örnek obje olarak seçilen bir bitkinin yetiştirildiği ortam koşullarına veya hasat zamanına bağlı olarak değişen kimyasal özelliklerinin yapılacak analizlerle belirlenmesi ve bu özelliklere bağlı olarak, yetiştiği yörenin çevresel özelliklerinin veya mevsimsel olarak hasat zamanının saptanabilmesi için sınıflandırılmasında, kullanılacak olan kümeleme tekniklerinin incelenmesi ve bu bağlamda Muğla bölgesinin çeşitli yörelerinden gelen ve farklı zamanlarda toplanan sığla ağacının yapraklarından farklı yöntemlerle elde edilen uçucu yağların bileşimlerine göre kümelenmesi ile geldikleri yörenin iklim toprak gibi bölgesel özelliklerine veya toplanma zamanlarına göre ya da uçucu yağların bitkiden ayrıştırılması yöntemlerine göre farklılık gösterdiği düşünülen özelliklerinin bu kümeler ile bağlantısının araştırılmasıdır.

Muğla yöresinde yetişen Sığla ağacı, geçmiş zamanlarda yapılan çeşitli ayinlerde yakılan kuru yongalarının içerdiği yağdan gelen hoş kokusu nedeniyle tütsü olarak kullanıldığından "Günlük Ağacı" olarak da adlandırılmaktadır ve ağacın önemi, elde edilen bu yağdan kaynaklanmaktadır. Ağacın kabuğunun çizilmesiyle, özünden elde edilen bir çeşit balsam olan "*Sığla Yağı*" artık özellikle parfüm sanayinde kullanılan önemli ve çok değerli bir hammadde haline gelmiştir. Ayrıca, bu yağ iyi bir antiseptik olup sağlık ve eczacılıkta, parfümeride ve buhur olarak kilise vb. yerlerde de kullanım alanı bulmaktadır.

Çalışmanın amacına uygun olarak hazırlanan bu tez kapsamında da, kabukları yerine değerlendirilmek amacıyla farklı coğrafi alanlardan farklı zamanlarda toplanan sığla ağacı yapraklarından alınan örneklerin, hidrodestilasyon, yöntemiyle ayrıştırılan uçucu yağlarında ki bileşenlerinin kalitatif ve kantitatif kromatografik analizleri yapılarak örneklerdeki uçucu yağların majör bileşenlerdeki değişimi tayin edilmiş olup, farklı yörelerde doğal olarak yetişen Günlük Ağacının eş zamanlı hasat edilen yapraklarının majör bileşenlerindeki değişimler kümelenmiş ve bu kümelerin geldikleri yöreye göre gruplanmış ve uçucu yağ bileşimi ile ilgili bağlantısı saptanmıştır.

1.1. Kimyada Sayısal Verilerin Analizi

Kimyada deneysel verilerin değerlendirilmesi neden-sonuç ilişkilerinin belirlenmesi açısından istatistiğin ilkelerine, veri gruplarının oluşturulması ve anlamlandırılması açısından veri madenciliği yöntemlerine, kabul edilebilir deney tasarımlarının gerçekleştirilebilmesi açısından modelleme ve optimizasyon tekniklerine ve tüm bu ilke, yöntem ve tekniklerin uygulanabilirliği açısından da bilgisayar donanımı, yazılımı, algoritma ve programlama ya da hazır olarak kullanılabilen paket programlarına ihtiyaç duyar ve bunlar da kısaca Kemometri biliminin temel alanlarını oluşturur. Günümüzde veri tabanlarının boyutları terabaytlarla ifade edilecek kadar büyüktür ve önemli bilgiler bu büyük veri yığınları içinde saklı kalmaktadır. Farklı özelliklerdeki veri tabanlarında belli bir anlamı olmadan kayıtlı olan verilerin yorumlanarak anlamlı ve kullanışlı bilgiler haline getirilmesi ile teknolojinin ve bilimin ilerlemesiyle gereksinim duyulan ve saklanan veri boyutunun her geçen gün artması ve bu verilerin anlamlandırılmadan sadece saklanması bilimsel yorum getirme, geleceğe yönelik tahminlerde bulunma, sınıflandırma, kümeleme ve tanımlama gibi konularda yeterli olmaması ve bundan dolayı bu konulara olan ilginin artması kaçınılmaz görünmektedir. Bunlardan biri olan kümeleme işleminin, ekonomi, tıp hizmetler gibi pek çok alandakine benzer olarak, bilimsel ve teknolojik araştırmalarda da toplanan verilerin değerlendirilmesinde, örneğin kimyada ürünlerin, madenlerin, bileşiklerin, canlıların fiziksel ya da kimyasal özelliklerine göre orijinlerin belirlenmesi, yapılarının gruplanması, özelliklerinin zamana bağlı olarak değişimlerinin incelenmesi gibi konularda da çok geniş kullanıma sahip olduğu pek çok kaynaktan açıkça anlaşılmaktadır (Brereton, 2003; Arifoğlu, 2005; Özkan, 2008; Demir vd., 2009; Akpolat, 2010).

1.1.1. İstatistiğin tanımı ve anlamı

İstatistik çeşitli anlamlarda kullanılmıştır, istatistik sözcüğünün kökü hakkında çeşitli görüşler vardır. İstatistik sözcüğünün kökeni konusunda kesin bir görüş birliği yoktur. Bazı dil bilimciler sözcüğün Latince’de “durum “ ya da “vaziyet” anlamına gelen status kökünden türediğine inanırken, bazıları stato kökünden geldiğine ya da

İtalyanca’da devlet anlamında kullanılan “istatista” sözcüğünden türediğine inanırlar. Başka bir grup bilimci ise, sözcüğün “gözlem” anlamına gelebilen statizein sözcüğünden türemiş olabileceğini belirtmektedirler (Arıcı, 2001). Bu görüşler fazla önemli değildir. Önemli olan, bugün kullandığımız kavramın oldukça eski bir geçmişinin olduğunun bilinmesidir. İstatistik düşünce ve mantığında temel amaç, rasgele toplanmış verilerle rasgele sayısal değerlendirmeler yapmak değil, gerekli olan tüm sürecin belirli ilkeler çerçevesinde bilinçli olarak gerçekleştirilmesidir. Bu tür bir gereksinimin ortaya çıkışı esasta şu temele dayanır: herhangi bir olayda bir genelleme yapabilmek için, olaya ilişkin tüm öğelerin elde edilmesi her zaman güç hatta zaman zaman imkansızdır, dolayısıyla elde edilebilecek belli sayıda örneğe dayanarak tüm yorumların yapılabilmesi için bu konudaki yöntemsel bakış açılarının geliştirilmesi ve düzenlenmesi her zaman büyük öneme sahiptir. İstatistiğin tanımı ve anlaşılması ile ilgili bazı temel kavram ve tanımlar EK1’de kısaca verilmiştir (Şenocak, 1990; Akpolat, 2010).

1.1.2. Veri madenciliği

Literatürde “Veri Madenciliği” teriminin farklı tanımlarının yapıldığı makale ve çalışmalar mevcuttur. Berson ve Smith veri madenciliğini veri ambarlarındaki büyük veri setlerinin yapay zeka (AI-Artificial Intelligence) ve istatistiksel matematik tekniklerle işlenerek yeni korelasyonlar ve örüntülerin bulunması işlemi olarak tanımlamışlardır (Berson, 1997). Veri madenciliğinin başka bir tanımı, büyük miktardaki verilerin otomatik veya yarı-otomatik yöntemlerle analiz edilerek anlamlı örüntü ve kuralların bulunmasıdır. Fayyad’ın tanımlamasına göre veri madenciliği, veri tabanlarından geçerli, önceden bilinmeyen, güvenilir, potansiyel olarak kullanışlı ve nihayetle anlaşılabilir örüntülerin çıkarılması işlemidir (Fayyad, 1996). Veri madenciliği birçok kaynakta Veri Tabanlarında Bilgi Keşfi (VTBK) (KDD-Knowledge Discovery in Databases) kavramı ile eş anlamlı olarak kullanılmaktadır. 1995 yılında Montreal’de gerçekleşen birinci uluslararası KDD konferansında VTBK, veritabanındaki verilerden bilgilerin (örüntüler, ilişkiler, kurallar, vb..) elde edilmesi işleminin tamamı olarak tanımlanmıştır. Veri madenciliğinin ilgilendiği temel konular aşağıda kısaca verilmiştir. Bunlar;

- Sınıflama
- Tahminleme
- Segmentasyon/Kümeleme
- Tanımlama

şeklinde olup, veri madenciliği ile ilgili ayrıntılar EK2’de sunulmuştur (Fayyad vd., 1996; Cao, 1998):

1.1.3. Kümeleme teknikleri

Sınıflandırma veri madenciliğinin en çok kullanıldığı alandır. Amacı yeni bir nesnenin özelliklerini açıklamak ve bu yeni nesnenin daha önceden tanımlanmış sınıf setlerinden birine atamasını yapmaktır. Sınıflandırmanın yapılabilmesi için sınıflandırılmamış veriye uygun bir model uygulanabilir. Veri madenciliğinin sınıflandırma grubu içerisinde en sık kullandığı teknik karar ağaçlarıdır. Aynı zamanda lojistik regresyon, diskriminant analizi, sinir ağları ve fuzzy setleri de kullanılmaktadır. İnsanlar verileri daima sınıflandırdıkları, kategorize ettikleri ve derecelendirdikleri için sınıflandırma, hem veri madenciliğinin temeli olarak hem de veri hazırlama aracı olarak da kullanılabilir.

Kümelemenin amacı ise heterojen olan ana kütleli homojen gruplara veya kümeler ayırmaktır. Kümeleme önceden tanımlanmış sınıfları baz alarak gruplamaz yani kümeleme gruplamada benzerlikleri kullanırken, sınıflandırma önceden tanımlanmış sınıflar modelini temel alır. Kümeleme, veri madenciliği sürecinden önce veri hazırlama esnasında da kullanılabilir. k-means ve toplayıcı algoritma (agglomerative algorithm) teknikleri ile onların varyasyonları kümeleme yöntemleri olarak kullanılmaktadır. Kümeleme analizi en yaygın olarak kullanılan veri madenciliği tekniklerinden biri olup, literatürde kümeleme analizini açıklayan pek çok tanım bulunmaktadır (Han ve Kamber, 2001; Karypis vd., 1999; Jain ve Dubes, 1988).

Kümeleme, en basit tanımıyla benzer özellik gösteren veri elemanlarının kendi aralarında gruplara ayrılmasıdır. Kümeleme, bir dizi örüntüyü (gözlem sonuçları, veri nesneleri ve özellik vektörleri) ayrık ve homojen gruplar oluşturacak şekilde

gruplandırma işlemidir. Bu işlem örüntülerin benzerlik derecelerine göre sınıflara veya kümelere ayrılmasıyla gerçekleştirilir. Kümeleme işlemi sonucunda elde edilen her başarılı (geçerli) küme içinde yer alan nesneler arasında maksimum benzerlik ve kümeleme sonucu elde edilen her kümedeki nesneler arasında ise maksimum farklılık oluşması sağlanır (Han ve Kamber, 2001; Karypis vd.,1999; Jain, vd., 1999). Karypis'in kümeleme tanımı da benzerdir: Kümeleme işlemi nesneleri aralarındaki benzerliğe bağlı olarak gruplara ayırır (Karypis vd., 1999). Böylece benzer nesneler aynı grup içinde toplanırken farklı nesneler farklı gruplarda yer alır. Benzerlik nesneleri tanımlayan özelliklerdir. Kümeleme analizi yüksek veri yoğunluğu sağlama tekniği olarak da tanımlanmaktadır. Bu tanımlamaya göre daha az bellek kullanılarak daha çok veriye ulaşılabilmektedir. Witten ve Frank kümeleme analizine daha karmaşık bir tanım getirmişlerdir (Witten ve Frank, 1999). Bu tanıma göre. Kümeleme işlemi ile elde edilen kümeler, örneklerin ve veri noktalarının çizildiği alanda bir mekanizma oluşturur. Bu mekanizma, bazı nesnelerin kendi aralarında diğer nesnelerden daha güçlü bir benzerlik oluşturmasını sağlar.

Mercer ise, kümeleme algoritmalarını “ölçüm” uzayını “anlam” uzayına yönlendiren işlemler olarak tanımlamıştır (Mercer, 2005). Bu tanıma göre, ölçüm uzayı çok boyutlu olup sürekli, ayrık veya kategorik veriler içerebilir. Anlam uzayı ise sonlu ve ayrık bir etiketler dizisidir. Bir nesnenin ölçüm uzayındaki değeri bulunduktan sonra, nesneyi bir kümeye dahil etmek için bu nesneye anlam uzayından bir etiket verilir. Kümeleme için yapılan ve literatürde önemli bir yere sahip bir diğer tanımlama ise şöyledir: Kümeleme analizi nesnelerin kendine özgü bir anlamı olan alt gruplara sınıflandırılması işlemidir. Kümeler temsil ettikleri nesneleri en iyi şekilde ifade edecek şekilde düzenlenir (Jain vd., 1988). Bu tanıma göre, kümeleme sonlu bir nesne dizisi üzerine uyarlanmış, bir sınıflandırma yöntemidir. Kümeleme tekniklerinde karşılaşılan temel kavram ve tanımlar ile ilgili ayrıntılar EK3'de verilmiştir.

1.1.4. Bilgisayar programları

Kemometrik uygulamalarda gelişen bilgisayar teknolojisi ve bilgisayar programları bu konuda yapılacak olan sayısal işlemleri oldukça kolaylaştırmıştır. Bu amaca yönelik olarak da çok sayıda programda geliştirilmiş olup ve bunlardan en çok

karşılaşılanları SAS, SPSS, MATLAB, MINITAB, JMP, NTSYS, SYSTAT, MEGA, POPGEN ve MVSP' dir. Bu çalışmada yapılacak olan kümeleme tekniklerine ilişkin olarak hesaplama algoritmalarının yazılımı için MATLAB paket program uygulaması için ise SPSS seçilmiştir. Bu iki bilgisayar programına ilişkin ayrıntılı açıklamalar EK4 ve EK5'de sunulmuştur.

1.1.5. Kümeleme çalışmaları

Pazarlama ya da hizmetler gibi günlük yaşamdaki uygulamalarda veya yapılan bilimsel ve teknolojik araştırmalarda toplanan verilerin değerlendirilmesinde karşılaşılan kümeleme tekniklerinin kullanıldığı çalışmalarla, gerek fizik, kimya ve biyoloji gibi temel bilimlerde gerekse tıp, mühendislik, bilgi teknolojileri, çevre ya da biyoteknoloji gibi pek çok uygulamalı bilimde karşılaşılmaktadır. Bundan dolayı çok geniş bir uygulama alanı bulan bu tekniklerin özellikle kimya alanında kimyasal analizlerde sağlanan ham verilerin değerlendirildiği kemometrik çalışmalara ilişkin olarak bu araştırma için yapılan kaynak değerlendirilmesinde ulaşılanlardan konuyla ilgili olacağı düşünülerek seçilenlerin bir özeti aşağıdaki sunulmuştur.

Bu konuda yazılmış pek çok kaynaktan seçilen ikisi alanlarında başlangıç yaptıkları için değerli bulunmuştur. 1988'de yapılan bir çalışmada iki araştırmacı yayınladıkları "İstatiksel Ekoloji" isimli kitaplarında temel bileşenler, ayırma ve kümeleme analizlerinin teorilerini izah etmiş ve bu yöntemlerin ekolojik verilere kullanımı ile ilgili örnekler vermişlerdir (Ludwig ve Reynolds, 1988). 1996'da yayınlanan "Uygulamalı Çok Değişkenli Teknikler" adlı kitapta temel bileşenler, ayırma ve kümeleme analizlerinin teorilerine değinilmiş ve bu analizlerle ilgili birer örnek verilerek SAS ve SPSS paket programlarında kullanımı gösterilmiştir (Sharma, 1991).

Sağlık bilimleri alanında örnek olarak seçilen Mersin Üniversitesi, Tıp Fakültesi, Nöroloji bölümünün 206 denek üzerinde yaptığı anket çalışmasının sonuçlarına göre, deneklerin RLS hastası olup olmama durumunu belirleyen değişkenler sınıflama ağaçları analizi ile tespit edilmiş ve RLS (Restless Legs Syndrome) hastalarının risk faktörleri farklı bir yaklaşımla incelenmiştir (Örekici vd., 2005).

Hayvancılık alanında yapılan arařtırmalara iliřkin olarak seilen iki rnek koyun yetiřtiricilięi zerinedir. Buların ilkinde, ukurova niversitesi Ziraat Fakltesi St Keicilięi Arařtırma ve Uygulama nitesi'ndeki bazı arařtırmacılar, yapmıř oldukları alıřmalarında, kmeleme analizi yntemlerinin kuramına deęinmiřler ayrıca kltr, yerli ve melezlerinden oluřan keilerden almıř oldukları zelliklere farklı uzaklık lleri kullanarak hem ham veri zerinde hem de veriler Z-puanlarına dnřtrldkten sonra kmeleme analizi yntemlerini uygulayıp elde ettikleri sonuları karřılařtırmalı olarak yorumlamıřlardır. Ham veriler iin hem Pearson korelasyon katsayısı kullanıldığında en iyi sonucu gruplar arası ortalama baęlantı kmeleme ynteminin, hem de Z-puanına dnřtrlmř veriler iin klit uzaklıęı kullanıldığında yine gruplar arası ortalama baęlantı kmeleme ynteminin en iyi olduęunu belirlemiřlerdir (Yazgan ve Kayaalp, 2002). Dięerinde ise, Batı Anadolu'daki 11 ilin koyuncululuęunun yapısı karřılařtırmalı olarak incelendięi bir arařtırmada, 2003-2008 yıllarına ait Trkiye İstatistik Kurumu'ndan elde edilen koyun sayısı, saęılan koyun sayısı, koyun st verimi, kesilen koyun-kuzu sayısı, koyun-kuzu eti retimi (ton), kırkılan koyun sayısı, yapaęı retimi verileri kullanılarak ok Boyutlu lkleme (MDS) ve Kmeleme analizleri uygulanarak illerin koyunculuk aısından sınıflandırılmalar yapılmıřtır (Gevreki vd., 2011).

Trlere iliřkin yapılan bir arařtırmada ise, Ege denizindeki deęiřik istasyonlardan aldıkları rneklere ayırma analizi uygulanmıř ve analiz sonucunda istasyon grupları arasında sayı bakımından ve aęırlık bakımından trlar sıralanmıřtır. Sayı bakımından en ok bulunan trlar, *Pagellus erythrinus*, *Mullus barbatus*, *Mullus surmeletus*, *Diplodus annularis*, *Spicara flexuosa*, *Sardina pilchardus* ve *Parapenaeus longirostris* olduęu, aęırlık bakımından ise *Mullus barbatus*, *Pagellus erythrinus*, *Spicara maena*, *Scylliorhinus canicula*, *Loligo vulgaris* ve *Parapenaeus longirostris* olduęu tespit etmiřlerdir (Tserper vd., 1999). Costa Rica'da yapılmıř olan bir alıřmada da aynı aileden iki trden (*Cryptotis merriami* ve *Cryptotis nigrescent*) almıř olduęu llebilir zelliklere Temel Bileřen Analizini uygulanarak, bu trlerin coęrafik daęılımlarını incelenmiřtir. Analiz sonucuna gre *Cryptotis nigrescent*'in *Cryptotis merriami*'den daha geniř bir alana yayılım gsterdięi tespit edilmiřtir (Woodman, 2000).

1998 de yapılan bir n kleer santralin  evresiyle olan iliřkisine dair yapılan bir  alıřmada, Bel ika'nın Doel santralının suyunu bıraktıėı b lgede 1994-1995 yılları arasında 1 yıl boyunca 2 karides, 55 balık 4 yenge  t r  yakalamıřlar ve bu t rlerin mevsimsel deėiřimlerini incelemek i in temel bileřenler analizini uygulamıřlardır. Bu analiz sonucunda t rleri etkilenmeye g re beř grupta sınıflandırmıřlar. Buradaki balıkların 36 tanesi deniz balıėı, 16 tanesinin tatlı su balıėı olduėu tespit edilmiřtir (Maes, 1998).  evre ile ilgili olarak He ve ark. tarafından yapılan bir bařka  alıřmada, Hawaii adalarında 4 yıl boyunca (1991–1994) dip trol  ile yapmıř oldukları  alıřmada elde ettikleri verilere K meleme Analizi uygulanmıř ve birim  abada harcanan g ce g re trolde  ıkan balık t rleri sınıflandırmıřtır (He vd., 1997).

Gıda ve kimya alanında yapılan  alıřmalar incelendiėinde muhtelif  alıřmalarla karřılařılmaktadır. Bunların ilkinde, Kanarya adalarının farklı b lgelerinden toplanan bal  rneklerinin  zellikleri ve sınıflandırılmasında farklı  ok deėiřkenli istatistik tekniklerinin kullanıldıėı bir arařtırmada, temel bileřen ve k meleme analizi ile alınan  rneklerin Na, K, Sr, Mg, Ca ve Cu i erikleri a ısından karakterize edilebildiėini a ıklamıřlardır. Ayrıca ayırma analizini uygulayarak, Kanarya balının metal i eriėi bakımından Kanarya balı olamayan bal  rneklerinden ayırt edilebildiklerini de a ıklamıřlardır (Hernandez, 1999). Yaė asitlerinin kimyasal analizlerine g re benzer t rlerin belirlendiėi bir bařka arařtırma sonu larına g re ise, yakalanan somon balıklarından  ıkan parazitlerin  eřitli  zellikleri alınarak bu  zelliklere temel bileřen analizi uygulanmıř ve bu analiz sonucunda parazitlerin    farklı t r *Gyrodactylus salaris*, *Gyrodactylus derjavini* ve *Gyrodactylus truttae* olduėu tespit edilmiřtir (Shinn, 1999).

Voncina ve arkadařları yapmıř oldukları bir  alıřmada yaė asit i eriklerinin karakterize edilmesi ile bitkisel yaėların sınıflandırılmasında ayırma ve temel bileřenler analizinden yararlanmıřlardır.  alıřmada 7 element i in 132 yaė  rneėi incelenmiř ve ilk 2 temel bileřen ile toplam varyansın % 97,8'inin a ıklandıėını ve ayırma analizi sonucunda 5 ayırım fonksiyonu ile 7 element arasındaki ayrımı % 96.84 doėruluk oranıyla ortaya koyduklarını ifade etmiřlerdir (Voncina vd., 2005). Bitki  rneklerine ait kimyasal farklılıkları ortaya koymak amacıyla yapılan benzer bir  alıřmada da temel bileřenler analizinden yararlanılmıř ve analizi sonucunda

incelenen kimyasal özelliklere ait toplam varyasyonun % 78,6'sının ilk 2 bileşen tarafından açıklandığını ortaya konulmuştur (Pierce vd., 2006). Bir başka çalışmada ise, Kuzey Ege yöresinin hakim zeytin çeşidi olan Ayvalık zeytin çeşidine ait erken hasat yağ örnekleri Ege Bölgesi'nin farklı yörelerinden sağlanmış olup, yağ asitleri profili ve skualen düzeyleri açısından kemometrik yöntemlerle (Temel Bileşen Analizi PCA, Hierarchical Kümeleme Analizi HCA) sınıflandırılmıştır.

1.2. Sığla Ağacı

Liquidambar Orientalis Liquidambar L. Hamamelidaceae familyasındaki tek cins olup, Batı Asya, Doğu Asya ve Kuzey Amerika'da dağılım gösteren 4 ayrı türü ile bilinmektedir. L. Orientalis'in deniz seviyesinden başlayıp, 1000 metre yüksekliğe kadar çıkabildiği görülmektedir. Geniş tepeli, çok dallı, 20 metreye kadar uzayabilen, 200-300 yıl yaşayabilen bir ağaçtır. Sert kışlara ve dona karşı duyarlıdır. Kökleri sığdır. Nemli, serince toprakları sever ışık isteği çoktur. Çınara benzeyen, genellikle beş dilimli olan yapraklarını kışın döker. Çiçekleri bir evciklidir. Dişi çiçekler ilk oluştukları zaman renkleri yeşil olup daha sonraları kırmızımsı renk alır. Meyvesi odunsu-kapsül yapıda olup uzun bir sapın ucunda, aşağıya doğru sarkık olarak durur. Olgunlaştıkları zaman sertleşir, kapsüller açılır ve tohumlar dökülür. Çok küçük kanatlı olan tohumun rengi koyu kahverengidir, basık, dip tarafı yuvarlak, uç kısmı sivridir. Tohum kabuğu parlak, ince ve serttir (http://tr.wikipedia.org/wiki/Anadolu/Sığla_ağacı, 2011). Ekonomik bir ağaç olan *Liquidambar orientalis* gövdesinden sığla balzamu elde edilmektedir. Farmakognoside bu balzam *Styrax liquidis*, *Styrax cradus*, *Styrax reçinesi* ve sığla yağı olarak adlandırılmaktadır Balzam gövde de belirli usullere göre yapılan yaralanmalar sonucu sızıntıdan elde edilen patolojik bir üründür. Toplanan bu ürün sonradan sıkılmakta ve yabancı maddelerden ayrılmaktadır. Geri kalan posa fümigan olarak kullanılmaktadır. Balzam yarı akıcı, kahverengi, yapışkan, opak ve belirli aroması olan bir maddedir. Sığla ağacının kabuklarında yara açılarak elde edilen sığla yağı içeriğinde tarçın asidi, styracin, styrol ve storesinol gibi maddeler bulunmaktadır. Kullanım alanları olarak basta parfümeri sanayinde fiksator (kalıcı olma) olmak üzere sabun yapımında, verniklerde, tütüne tat vermede, bazı tıbbi

uygulamalarda; haricen uyuz ve mantar gibi cilt hastalıklarında, antiseptik olarak, parazit öldürücü olarak, dahilen astım ve bronşit gibi üst solunum yolu hastalıklarının tedavisinde kullanılır. Ayrıca sığla yağı elde edildikten sonra arta kalan buhur (Cortex Thymiamitis), cami ve kiliselerde tütsü ve yakı olarak kullanılmaktadır. Tarihsel süreçteki ticaretinde Finikeliler önemli rol oynamışlar ve eski Mısırlılar sığla yağını mumyaların hazırlanmasında kullanmışlardır (Pamukcuoğlu, 1964).

Sığla ağacı ülkemizde en geniş yayılışını Muğla bölgesinde özellikle bunun kıyı kesimindeki Marmaris, Dalaman, Köyceğiz ve Fethiye dolaylarında sıcak ve ılık Akdeniz iklim katında yapmaktadır. Ayrıca iç kısımlarda küçük adacıklar şeklinde örneğin Denizli-Acıpayam-Gölcük köy, Antalya-Sütçüler-Çandır ve Muğla bölgesinde Namnam çayına karışan Turgut deresinin üzer bölgelerine kadar yayılmaktadır. 1940 larda ülkemizdeki toplam sığla orman alanı 6300 hektar iken günümüzde bu ancak 1350 hektara kadar düşmüştür. Bunun sonucu olarak bu ormanlardan elde edilen sığla balzamu da buna paralel olarak azalmıştır. Örneğin 1950 yılında 181.279 kg iken bu miktar 1980 yılında 18.000 kg'a ve 1990 yılında 1000 kg'dan aşağıya inmiştir (Huş, 1964; Kaya, 2003).

1.3. Uçucu Yağlar

Uçucu yağlar, aromatik bitkilerden veya bitkisel ekstraktlardan elde edilen, oda sıcaklığında sıvı halde olan, bazen donabilen, uçucu, kuvvetli kokulu, su buharı ile sürüklenebilen yağimsı karışımlardır. Daha çok koku özelliklerinden yararlanılan uçucu yağlar, aynı zamanda anti mikrobiyal aktiviteye de sahiptirler. Uçucu yağ, katı ya da sıvı değişik birçok kimyasal bileşiğin birbirinde çözünerek homojen bir çözelti oluşturduğu kompleks bir karışımdır. Kimyasal ve fiziksel özellikleri açısından sabit yağlardan farklı özelliklere sahiptirler. Açıkta bırakıldıklarında oda sıcaklığında bile buharlaşabildiklerinden ‘uçucu yağ’, ‘eterik yağ’ veya ‘esans’ adını alırlar. Güzel kokulu olmaları ve parfüm sanayinde kullanılmaları nedeniyle önemli bir ekonomik değer taşırlar. Genellikle yurdumuzda yerli olarak elde edilenlere ‘yağ’ (defne yağı, kekik yağı, gül yağı vs.), yurt dışından ithal edilenlere ise ‘esans’ (limon esansı, lavanta esansı vs.) denilmektedir. Buna karşılık ithal edildiği halde yağ olarak

isimlendirilenler de vardır (tarçın yağı, karanfil yağı vs.). Bunlar su ile karışmadıklarından ve su yüzeyinde tabaka oluşturdıklarından ‘yağ’ olarak adlandırılırlar. Bugün doğada yetişen 300’e yakın bitki familyasından yaklaşık üçte biri uçucu yağ içermektedir. Ülkemizin büyük bir kısmını da içine alan Akdeniz bitki coğrafyası bölgesi uçucu yağ taşıyan bitkiler açısından en zengin floralardan biridir (Tümen vd., 1992; Tanker vd., 1990). Uçucu yağlar ya bitkinin belirli organlarında örneğin taç yaprak, yaprak, meyve, kabuk, meyve sapı, odunsu doku gibi ya da bitkinin tüm organlarında ayrıca bazen bir organın belirli dokularında da bulunabilirler. Bu yağlar bitkilerin bağlı bulunduğu familyalara göre salgı tüyünde, salgı ceplerinde, salgı kanallarında veya salgı hücrelerinde bulunmaktadır (Ceylan, 1987).

1.3.1. Fiziksel özellikleri

1. Genellikle %1’den az elde edilirler. Karanfil tomurcukları ise istisnadır ve yaklaşık %15 oranında uy taşır.
2. Genelde soluk sarı veya sarı renklidirler. İlk elde edilirken tamamı renksizdir. Bekledikçe kromofor grupları oluşur ve renk şiddeti artar. Nadiren kırmızı (O.Caryophylli) veya mavidir (O.Chamomillae).
3. Tamamı özel kokuludur.
4. Yoğunlukları genellikle 0,8 –1,0 arasındadır. Nadiren yoğunluğu sudan ağır(1,0-1,5) uçucu yağlara rastlanır. (O.Caryophylli, O.Cinnamomi, O. Sassafras)
5. Suda çözünmezler, suyla karışmazlar ancak sürüklenirler. Suyla bekletildiğinde taşıdığı bazı maddeler suya geçerek eczacılıkta önemli olan aromatik suları (=Aqua) meydana getirirler.
6. Nadiren oda sıcaklığında katı haldedir(O.Anisi). Alkol ilave edildiğinde tüm uçucu yağların donma noktası düşer.
7. Doğal olarak elde edilen yağların hepsi optikçe aktiftir ve değerleri spesifikdir. İlaveten her uyun alkoldeki çözünürlüğü, yoğunluğu, asitlik, sabunlaşma, kırılma gibi indisleri yağa ait farklılıklar gösterir.

1.3.2. Uçucu yağların sınıflandırılması

Uçucu yağlar değişik özelliklerine göre gruplara ayrılabilir. Bunlar kimyasal bileşimleri, aromatik özellikleri, farmakolojik ve terapik etkileri göz önünde bulundurularak gruplandırılabilirler. Bunlardan en önemlisi kimyasal bileşimlerine göre olmaktadır; Bitkisel uçucu yağlar oldukça farklı yapıda olup, önemli antioksidan bileşikler içerirler. Bu yağların kimyasal yapıları iki kısımdan meydana gelmektedir. Bunlar hidrokarbonlar (C_5H_8)_n ve oksijen içeren hidrokarbon türevleridir. Hidrokarbonlar ise kendi içinde hemiterpenler, terpenler ve seskiterpenler olarak üç gruba ayrılmaktadır. Terpenlerin oksitlenmesi ile meydana gelen oksijenli türevler uçucu yağın kendine özgü kokusunu, tadını ve terapik özelliklerini verirler. Uçucu yağlarda asıl önemli olan bileşikler oksitlenmiş türevlerdir (Altuğ ve Elmacı, 1998; Ceylan, 1997). Aşağıda terpenlere ilişkin daha ayrıntılı bilgi sunulmuştur.

Terpenler, doğal ürünlerin en yaygın gruplarından biridir. Bitkilerde ve hayvanlarda birçok farklı işlevleri bulunurken gıdalarda da aroma bileşenleri olarak önemlidirler. Örneğin turuncgiller, tarçın ve diğer baharat aromaları birkaç terpen ile karakterize edilir. Limonene ve citral (her ikisi de limonda bulunur), camphor, pinene (çam ağaçları), eugenol (karanfil), anethol, thymol, geraniol (gül) ve menthol en yaygın bilinen terpenlerdir. Kimyasal anlamda terpenler, yapısı çeşitli fakat belli sayıda isoprene birimlerine sahip olan bir moleküller grubu olarak tanımlanır. (methylbuta-1,3-diene, hemiterpene olarak isimlendirilen 5 karbonlu atomdur). Kimyasal anlamda terpenler, yapısı çeşitli fakat belli sayıda Şekil 2.1.'de verildiği gibi izopren birimlerine sahip olan bir moleküller grubu olarak tanımlanır. Terpenik ve aromatik maddelerin oksijensiz yada oksijenli türevlerinden bir çoğu uçucu yağda karışım halinde bulunmaktadır. Oksijensiz olanlar çoğunlukla kolay uçucudurlar. Uçucu yağlar düşük sıcaklıklarda bile sıvı halde kalabilirler. Oksijenli türevleri ise daha az uçucudurlar ve uçucu yağ soğutulduğunda birçoğu çökerek oksijensiz bileşiklerden ayrılırlar. Bazı uçucu yağlarda çöken kısmına stearopten, bu koşullarda sıvı halde kalan kısmına da elaopten adı verilir. Uçucu yağlara fraksiyonlu destilasyon uygulandığında ilk ele geçen fraksiyonlar elaoptenden oluşan oksijensiz bileşiklerdir. Terpenlerin oksitlenmesi ile meydana gelen oksijenli türevler uçucu yağın kendine özgü kokusunu, tadını verirler. Uçucu yağlarda asıl önemli olan bileşikler oksitlenmiş türevlerdir (Çoban, 2009).

1.3.3. Uçucu yağların elde edilme yöntemleri

Uygulanan yöntemler aşağıda verilmiştir (Ranghau, 2001);

- Destilasyon Yöntemi
- Mekanik Yöntem (Presleme Yoluyla Uçucu Yağ Elde Edilmesi)
- Anfloranj Yöntemi (Ekstraksiyon Yoluyla Uçucu Yağ Elde Edilmesi)
- Tüketme Yöntemi

Monoterpenler		İzopren sayısı: 2	C atomu sayısı: 10	
	Mirsen		Limonen	
			Geranol	
	Mentol		Sitronella	
			Kampor	
Seskuiterpenler		İzopren sayısı: 3	C atomu sayısı: 15	
	Farnesol		Bisabolen	
			Selinen	
	Absisik asit (fitohormon)			
Triterpenler		İzopren sayısı: 6	C atomu sayısı: 30	
	Squalen		Kolesterol (zar bileşeni)	
	Östradiol (hormon)		Testeron (hormon)	

Şekil 1.1. Terpenlerin sınıflandırılması ve bazı monoterpen, seskuiterpen ve triterpen molekülleri

1.3.3.1. Destilasyon yöntemi

Bilindiği gibi destilasyon, sıvıların kaynama noktalarındaki farklardan yararlanılarak gerçekleştirilen bir ayırma işlemidir. Bu yöntem ile elde edilen uçucu yağlar yüksek oranda kaynama noktası düşük bileşikler ve az miktarda kaynama noktası yüksek ve suda çözünen bileşikler içermektedirler. Uçucu yağ eldesinde kullanılan modern destilasyon teknikleri de 3 başlık altında incelenebilir;

- Su Destilasyonu
- Su ve Buhar Destilasyonu
- Buhar Destilasyonu
- ***Su Destilasyonu (Hydrodistillation - HD)***

Uçucu bileşiklerin eldesin de yaygın olarak kullanılan geleneksel bir yöntemdir. Küçük ölçekli üretimlerde Clevenger tipi bir aparatla yapılan destilasyon işlemi endüstriyel uygulamalarda büyük destilasyon kazanlarında (İmbik) gerçekleştirilmektedir. Yöntemin esası; soğutucu ile irtibatlandırılan bir cam balon içerisinde su ve bitki materyalinin 2-8 saat süre ile kaynatılarak, su buharı ile birlikte hareket eden yağ moleküllerinin soğutucuda yoğunlaştırılıp sudan ayrıştırılmasına dayanmaktadır. Elde edilen uçucu yağ miktarı volumetrik olarak ifade edilir. Su destilasyonu en iyi toz halindeki materyallerde (örneğin; kök ya da odun unu) sonuç vermektedir. Elde edilen yağ miktarı çok olmakla birlikte suyun kaynatılması esnasında uygulanan yüksek sıcaklık, termal bazı reaksiyonlara neden olmaktadır. Bunun sonucu olarak artıfak oluşumu, hidroliz ve izomerizasyon olayları meydana gelmektedir. Uçucu yağların bileşimi pH'a bağlı olarak değişse de su destilasyonu yönteminde genellikle sıvının pH değeri kontrol edilmemektedir (Linskens ve Jackson, 1997; Fakhari vd., 2005).

- ***Su ve Buhar Destilasyonu***

İmbik içerisinde bulunan bir ızgara üzerine yerleştirilen bitkisel materyal, imbiğin altında bulunan suyun ısıtılması yoluyla buhara maruz kalır. Yöntemin diğer bir tipinde, su ve baharat karışımından buhar geçirilir. Esansiyel yağın destilasyon hızı, hammaddeye ve hammaddeden yağın serbest hale geçiş hızına bağlıdır (Polat ve Ötleş, 1997; Reineccius., 2005).

- **Buhar Destilasyonu**

Su destilasyonunda karşılaşılabilecek riskler buhar destilasyonu uygulaması ile ortadan kaldırılabilir. Buhar destilasyonunda kullanılacak bitki delikli bir metal levhaya ya da destilasyon kazanı ile aynı çaptaki etrafı kapalı üstü açık bir sepete konularak başka yerde üretilmiş buhar alttan bitkisel materyale püskürtülerek destilasyon uygulanır. Böylelikle buhar, baharata iyice nüfuz eder (Polat ve Ötleş, 1997; Reineccius., 2005).

1.3.3.2. Mekanik yöntem (presleme yoluyla uçucu yağ elde edilmesi)

Bazı droglardan destilasyon yöntemi ile uçucu yağ elde edilmek istendiğinde bu droglardaki uçucu yağ bozunmaktadır. Bu durumda bir kısım droğa bu yöntem uygulanır. Presleme yöntemiyle elde edilen yağlar ve ekstreleri genellikle berrak değildir. Bu ekstreleri berraklaştırmak için süzme santrifüj, alkol ile seyreltme (fermantasyonu engellemek için), ısıtma (albuminleri çöktürmek için) gibi işlemler uygulanır.

1.3.3.3. Anfloranj yöntemi (ekstraksiyon yoluyla uçucu yağ elde edilmesi)

Bu yöntem uçucu yağı az, ancak kıymetli droglar için kullanılmaktadır. Bazı bitkilerin esansları su buharının yüksek ısıyla bozulabilir ve bazı bitkilerin uçucu yağı çok az olduğundan esanslarını distilasyonla çekmek güçtür. Bu gibi hallerde ekstraksiyon metodu uygulanır. Bu metotta esanslar bazı çözücüler yardımıyla bitkiden alınır. Daha sonra çözücüye geçen esans distilasyon yoluyla çözücünden ayrılır. Uçucu yağ eldesinde kullanılan ekstraksiyon teknikleri aşağıda özetlenmiştir (Ceylan,1987).

- **Soğuk Suyula Ekstraksiyon**

Bu metot yasemin ve çuha çiçeği gibi bitkilerde uygulanır. Eritici olarak iç yağ ve domuz yağı kullanılmaktadır. Taze materyal (özellikle çiçekler) ince bir yağ sürülmüş cam plakalar üzerine konulur. Birkaç saat veya gün sonra konulan çiçeğin uçucu yağını alan cam plaketa üzerine sürülmüş yağdan alkol ekstraksiyonu ile uçucu yağ elde edilir.

- ***Sıcak Suyla Ekstraksiyon***

Bu metotta kullanılan eritici yağlar sıcaktır. Bazı materyaller toplandıktan kısa bir süre sonra canlılıklarını kaybederler. Az zaman alan bir yöntemdir. Yasemin ve çuha çiçeği dışındaki hemen bütün çiçekler bu metotla esanslarını verebilirler. Günümüz teknolojisinde bu metot yerini uçucu eriticilerle ekstraksiyona bırakmıştır. Fakat özellikle Fransız gülü ve bazen de portakal çiçeği, mimoza ve sümbül için halen bu metot uygulanmaktadır.

- ***Sıvılaştırılmış Gazlarla Ekstraksiyon***

Genel bir ekstraksiyon yöntemidir. CO₂ gibi sıvılaştırılmış gazlar kullanılarak gerçekleştirilir.

- ***Tüketme Yöntemi***

Drog uygun bir organik çözücüyle (hekzan, heotan vs. gibi) ekstre edilir. Bu esnada uçucu yağ, sabit yağ, mum, boya maddeleri v.s. gibi organik çözücüde çözünebilen maddeler organik çözücüye geçer. Süzülerek alınan organik çözücü vakumda tamamen uçurulur. Uçucu yağ etanol ya da sulu etanol ile tüketilir. -15 °C’ de bir süre, genellikle bir gece bekletilir, çöken kısımları soğukta süzildükten sonra çözücü vakumda uçurularak uçucu yağ elde edilir. Bu yöntemle elde edilen uçucu yağa ‘absolü’ adı verilmektedir.

1.3.4. Uçucu yağ eldesindeki değişimler

Destilasyon işlemi sırasında uçucu yağın değişikliğe uğrayıp uğramadığının tespiti için birçok araştırma yapılmıştır. Bu araştırmalarda destilasyona bağlı olarak, ortam şartlarında meydana gelen değişiklikler ve bunların uçucu yağının kimyasal yapısı üzerine olan etkisi incelenmiştir. Elde edilen sonuçlara göre, uçucu yağ ya destilasyon şartlarından veya destilasyon cihazına bağlı olarak değişmektedir. Değişmelere sebep olan faktörleri başlıca destilasyon süresi, tatbik edilen ısı, ortamın pH’sı ve destilasyon aletinin yapısı olmak üzere dört kısımda düşünülebilir. Uçucu yağın bulunduğu dokuya bağlı olarak, tatbik edilen süre değişir. Uçucu yağ dış salgı türlerinde bulunuyorsa, süre kısa, buna karşılık içteki dokularda bulunuyorsa süre uzundur. Bu süreyi azaltmak için bitki toz edilir. Ancak toz etme esnasında

materyalin taşıdığı uçucu yağın bir kısmının kaybolması ve yapısının değişmesi gibi bir dezavantajı da vardır. Bu dezavantajı ortadan kaldırmak için toz etme işlemi, sıvı azot veya karbondioksit karı ile sağlanan düşük ısı kullanılır. Koedam metot farklılığına göre uçucu yağların kimyasal bileşimlerdeki değişmeler üzerine araştırma yapmıştır. Bu çalışmada ekstraksiyon sonucu elde edilen uçucu yağ ile, destilasyon sonucu elde edilen uçucu yağ arasında önemli farklılıklar gözlenmiştir. Ekstraksiyon yönteminin de mahsurlar vardır. Bu yöntemde uçucu yağlar tamamen elde edebilmek için ekstraksiyon süresini uzatmak gerekmektedir. Bu da başka maddelerin, özellikle sabit yağların, uçucu yağa karışmasına neden olur. (Koadem, 1981).

1.3.5. Uçucu yağlarda bulunan bileşiklere örnekler

I. Hidrokarbonlar (Alifatik : Mirsen, osimen., Aromatik : Agropiren, naftalen, p-simen, stiren., Siklik : Felandren, fenken, kamfen, karen, limonen, pinen, sabinen, terpinen, terpinolen, tüyen. , Seskiterpen : Aromadendren, bizabolen, humulen, kadinen, karyofilen, kopaen, kurkumen, longifolen, selinen, ingiberen. , Diterpen : Kamforen., Azulen : Kamazulen).

II. Alkoller (Alifatik : Jeraniol, linalol, nerol, sitronellol., Siklik : Borneol, karveol, mentol, mirtenol, pulegol, terpi-nen- 4-ol, terpineol, tuyol., Seskiterpen : Elemol, farnesol, guayol, kübeol, nerolidol, paçulol, zingiberol., Aromatik : Feniletil alkol, küminalkol, sinamil alkol).

III. Aldehitler (Alifatik : Desilaldehit, sitral, sitronellal., Siklik : Mirtenal, perilaldehit, safranlal., Aromatik : Anisalaldehit, benzaldehit, küminaldehit, sinamaldehit, vanilin., Heterosiklik : Furfural).

IV. Ketonlar (Alifatik : Artemisia keton, diasetil, metil heptenon, metilno-nil keton, tageton., Siklik : Fenkon, kâfur, karvon, menton, piperiton, piperite-non, pulegon, tuyon, verbenon., Seskiterpen : Atlanton, turmeron., Aromatik ., setofenon., İronlar., İyononlar., Diğer : Jasmon, muskon, santalon, siveton).

V. Fenoller ve Fenol Eterler : (Alliltetrametoksibenzen, anetol, apiol, asaron, dilapiol, elemisin, fenikulin, karvakrol, metil kavikol, metilöjenol, miristisin, öjenol, safrol, timol).

VI. Kinonlar : (Hidrokinon, timohidrokinon, timokinon).

VII. Asitler (Alifatik : Ambretolik, anjelik, bütirik, jeranik, miristik, oleik, tiglich., Aromatik : Anisik, benzoik, sinamik., Diğer : Kamforik, nepetalik, sedanolik, sedanonik).

VIII. Esterler (Alifatik : Etil asetat., Terpenik : Bornil izovalerat, jeranil asetat, linalil asetat, sitronellil bütirat, sitronellil format, terpinil asetat., Aromatik : Benzil asetat, etil asetat, metil benzoat, sinamil sinamat., Azotlu : Damasenin, metil antranilat),

IX. Laktonlar, Kumarinler ve Kumaronlar : (Ambretolit, bergamotin, bergapten, bergaptol, ftalit, imperatonin, kumarin, limetin, sedanolit, umbeliferon) (Anonim, 2003).

1.3.6. Uçucu yağların analizi

Hidrodestilasyon, su-buharı destilasyonu ve ekstraksiyon yöntemleriyle ayrıştırılan uçucu yağların bileşenlerinin kalitatif ve kantitatif kromatografik analizleri GC-MS de yapılmıştır.

2. MALZEME VE YÖNTEM

Bu çalışma iki bölümden oluşmaktadır. Birinci bölümde deneysel aşama, verilerin elde edilişi anlatılmaktadır. İkinci bölümde ise elde edilen deneysel verilerdeki majör bileşenlerin istatistiksel analizinin yapılmasıdır. Çalışmanın aşamaları aşağıda kısaca sunulmuştur.

1. Farklı yörelerde ancak aynı dönemde toplanan ve doğal olarak yetişen Günlük ağacı yapraklarının majör bileşenlerindeki olabilecek değişmelere bakılması.
2. Elde edilen verilere göre majör bileşenler arasındaki değişimlerin örneklerin alındığı yörelere göre SPSS ve MATLAB programları yardımıyla kümeleme analizinin yapılması.

2.1. Deneysel Kısım

Bu bölümde, öncelikle çalışmalarda kullanılan bitkisel materyal, kimyasal maddeler ve aletler belirtilmiş olup yapılan deneysel çalışmalar hakkında bilgi verilmiştir. Ayrıca araştırma konusu olarak seçilen Muğla yöresine özgün bir bitki olan sığla (Günlük) ağacı yapraklarının kimyasal analizi yapılarak onun özelliklerini veren kimyasal bileşimi belirlenmiştir.

2.1.1. Kullanılan bitkisel materyal, kimyasal maddeler ve aletler

2.1.1.1. Bitkisel Materyal

Bu çalışmada kullanılan *Liquidambar orientalis* Mill. ağacının yaprakları Marmaris, Fethiye, Köyceğiz olmak üzere üç farklı bölgeden Temmuz ayında toplanmıştır.

2.1.1.2. Çalışmada kullanılan kimyasal maddeler

- n-Hekzan (Merck) veya Petrol Eteri (Carlo Erba)
- Metanol veya - Etanol

2.1.1.3. Aletler

- Clevenger Düzeneği (Amerikan Farmokopisine Göre)
- Gaz Kromatografisi- Kütle Spektroskopisi Sistemi (GC-MS) (Shimadzu Fission Model 8000-Fission MD 800)
- Rotary Evaporatör(Dönen Buharlaştırıcı)

2.1.2. Metot

Bu çalışmada uçucu yağların miktar ve bileşenlerini belirlemek için aynı farklı yörelerden aynı dönemde alınan bitki materyallerinin hidrodestilasyon metodu ile elde edilen uçucu yağlarında ki bileşenlerinin kalitatif ve kantitatif değişimleri incelenmiştir. Araştırmada, hidrodestilasyon işlemi (HD) için Clavenger cihazı kullanılmıştır.

2.1.2.1. Clavenger cihazı ile hidrodestilasyon

500 ml'lik cam balon içerisinde her bir bölgeden gelen bitki (Marmaris, Fethiye, Köyceğiz) ayrı ayrı olacak şekilde; Günlük ağacı bitkisinin yaprakları iyice öğütülmüş ve 500 gram tartılmıştır. Ardından, yaklaşık 500 mL saf su ile geri soğutucu altında HD işlemi ile uçucu yağlarının elde edilmesi için 2 saat kaynatma işlemi yapılmıştır. Her bölgeden gelen bitki için bu işlem 2 paralel olarak tekrarlanmış ve uçucu yağının toplandığı kapiler borudaki ölçekten okuma yapıldıktan sonra ayırma hunisine aktarılmıştır.. Ayırma hunisine ekstraktla birlikte ve 5 mL n-hekzan yerleştirilmiş ve emülsiyona yaklaşık olarak 1g NaCl ilave edilmiştir. Daha sonrada içerisinde kalan su yeterince Na₂SO₄ ilavesi ile tutulmuştur.. Ekstraktların darası alınan viallere aktarılması sonrasında basınçlı azot gazı kullanımı ile uçucu yağ miktarları belirlenerek kütle/kütle olarak yüzde verimleri

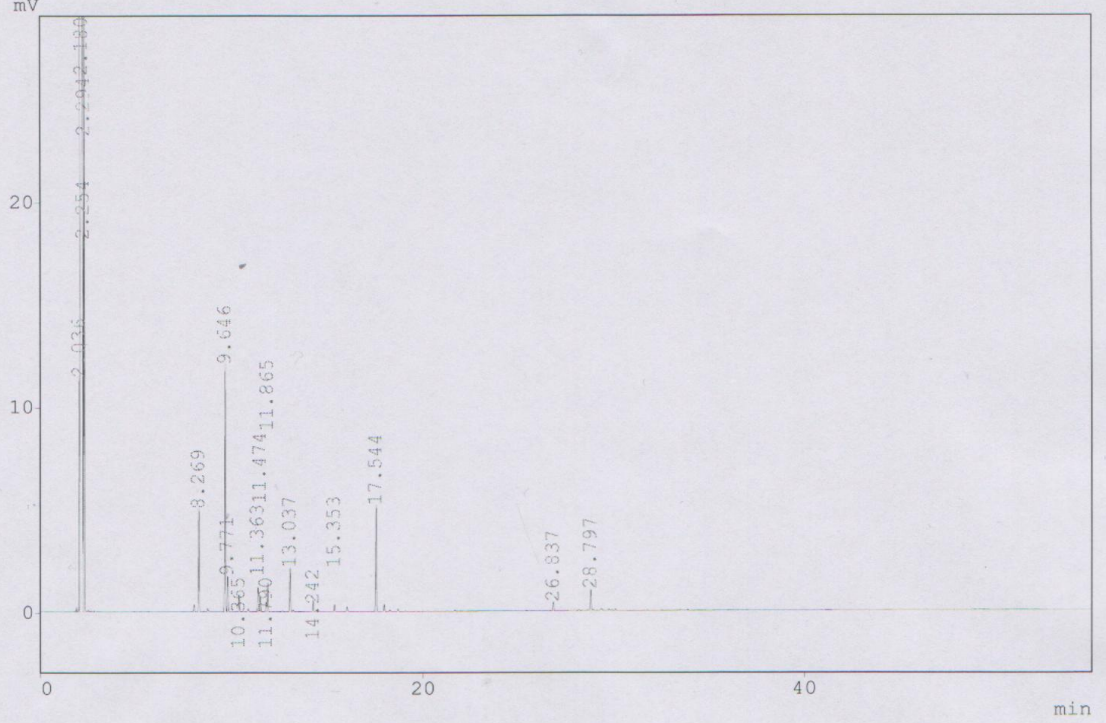
hesaplanmıřtır. Son olarak da elde edilen ekstraktlar uygun zcler ile GC-MS de analiz edilmiřlerdir.

2.1.3. GC-MS sonular

Bu alıřmada uucu yaėların miktar ve bileřenlerini belirlemek iin MS destekli GC tayinleri amacıyla alınan kromatogramlara iliřkin bir pik raporu rneėi (LOF1HD00) řekil 2.1’de verilmiřtir. Analiz sonucu alınan karıřımların bileřimi MS destekli olarak yapılmıř ve majr bileřenler ayrıntılı olarak saptanmıřtır. řekil 2.2’de ise majr bileřenlerden α -Pine’e ait ktle spektroskopisi rneėi verilmiřtir. İstatistiksel deėerlendirmede ise majr bileřenlerin daėılım oranlarının sayısal deėerleri kmeleme analizine tutularak arařtırılan bitkisel materyelin orjini hakkında bir grř oluřturulması saėlanmıřtır.

CLASS-GC10 Ver.=2.01 SYS=1 Ch=1 REPORT.NO=4 DATA=LOF1H.D00 11/12/01 13:48:34
Sample : LOFthye1.ksm.h
ID : LOFthye1.ksm.h
Dilution Factor: 1
Type : Unknown
Detector : FID
Operator : GulsenTEL

*** Chromatogram *** Filename:LOF1H.C00



*** Peak Report ***

PKNO	TIME	AREA	HEIGHT	MK	IDNO	CONC	NAME
1	2.036	20916	11482	V		0.2429	
2	2.180	8291159	1103307	VE		96.2901	
3	2.254	33916	18345	V		0.3939	
4	2.294	165573	165456	V		1.9229	
5	8.269	14292	5027			0.1660	
6	9.646	36349	12075			0.4221	
7	9.771	4832	1722	V		0.0561	
8	10.365	2306	831			0.0268	
9	11.363	3413	1170			0.0396	
10	11.474	2175	745	V		0.0253	
11	11.790	1286	452			0.0149	
12	11.865	4049	1389	V		0.0470	
13	13.037	6260	2119			0.0727	
14	14.242	1675	556			0.0195	
15	15.353	1037	346			0.0120	
16	17.544	16625	5183			0.1931	
17	26.837	1372	414			0.0159	
18	28.797	3374	1033			0.0392	

8610609 1331651 100.0000

*** Calibration Result ***

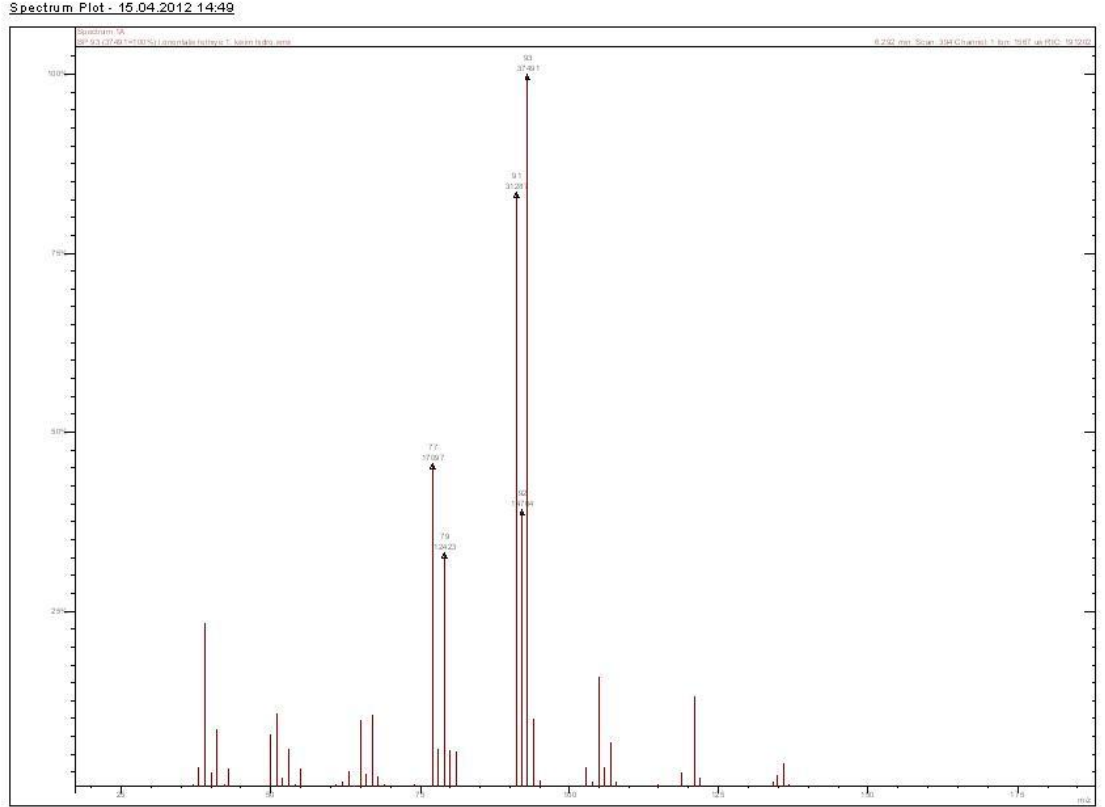
** ID Table **

IDNO	NAME	TIME	CONC	RF1	RF2
------	------	------	------	-----	-----

4 - 1/1

12/04/19 19:05:50

Şekil 2.1. GC tayinleri amacıyla alınan kromatogramlara ilişkin bir pik raporu örneği
(LOF1HD00)



Şekil 2.2. Majör bileşenlerden α -Pine'e ait kütle spektroskopisi örneği

Ancak GC_MS gibi özel donanımlara ilişkin yazılımlarda alınan sayısal verilerin doğrudan kullanımı hemen mümkün olmadığından bu sayısallaştırma işlemi sırasında CLASS-GC10 özel donanımdan alınan sayısal değerlerin SPSS, EXCEL veri tabanlarında kullanışlı hale getirilebilmesi için aşağıdaki geçişler yapılmıştır. Bunlar;

1. GLASS-GC10
2. Postrun Analysis
3. Single Chromatogram Processing
4. Display: PeakReport
5. PeakReport (Başlık Alanı) Sağtuş-OpenNotePad
6. Dosya-Farklı Kaydet ↓ (↓0F1S.D00.TXT)↓_KAPAT
7. Aç- (L0F15.D00.TXT) (WORD de)
8. Dosya-Farklı kaydet ↓ (L0F1S.D00wrđ.WRD)
9. Dosyadaki sayfanın tüm bilgilerinden sayfa altına yeni bir kopyasını çıkar. Başlık ve sayıları 7 kolonlu pik no sayılı tabloya çevir

PK NO	TIME	AREA	HEIGHT	MK	CONC	LOF1S. D00ex
1						
....			...	...	...	...
PKNO						

Tabloyu kopyala;

10. Boş Excel dosyası aç. Farklı kaydet (L0F1S.D00exl. XLS) ve tabloyu yapıştır
11. Dosyayı farklı kaydet (L0F1S.D00duz.xls)
12. Dosyada ilk 3 dakikadan sonra bileşenlerin alan dağılımını yeniden düzenle

2.2. İstatistiksel Değerlendirme

2.2.1. Kümeleme analizinin yapılması

Aşamalı olmayan kümeleme yöntemlerinin gerek teorik dayanaklarının aşamalı kümeleme yöntemlerine göre daha güçlü olması gerekse küme sayısı konusunda ön bilgi olması ya da araştırmacının anlamlı olacak küme sayısına karar verebilmesi açısından tercih edilmesini sağlamaktadır. Ancak bu yöntem belli bir ordinat noktası olmayan verilerde kullanılmaz. Bu nedenle bu çalışmada Aşamalı Kümeleme Yöntemi ve *Öklit uzaklığının* hesaplanmasıyla *Nearest Neighbor* (En Yakın Komşu) metodunu kullanılmıştır.

2.2.2. En yakın komşu kümeleme algoritması

Kümelerin belirlenmesinde kullanılan en doğal yol, en yakın komşu (Nearest Neighbor) özelliğinin kullanılmasıdır. Bir örnek kendine en yakın komşusu ile aynı kümeye konulmaktadır. Eğer iki örnek aynı komşuyu paylaşıyorsa birbirine benzer olarak kabul edilir. Çizge kuramlı metotlar en yakın komşu özelliğiyle yakından ilişkilidir, bununla birlikte, kümelerin oluşturulma şekli ve en önemlisi en son

paylaşımın belirlenmesi açısından birbirlerinden açık şekilde farklı özelliktedirler. $H = \{x_1, x_2, \dots, x_n\}$ örnek dizisini kümeye ayırmak amacıyla aşağıdaki adımlar yapılmalıdır. Kullanıcı en yakın komşuluk değeri için bir eşik değeri, t , belirlenmelidir.

1. $i \leftarrow 1$ ve $k \leftarrow 1$ olarak belirlenir ve x_1 örneği C_1 kümesine atanır.
2. $i \leftarrow i + 1$ olarak belirlenir. Biraz önce kümelenen örneklerin arasından x_i 'nin en yakın komşusu bulunur. d_m ifadesi x_i ile en yakın komşusu arasındaki uzaklığı ifade etmektedir. En yakın komşunun m . kümede olduğu varsayılın.
3. Eğer $d_m \leq t$ ise, x_i örneği C_m kümesine atanır. Değilse, $k \leftarrow k + 1$ olarak belirlenir ve x_i örneği yeni oluşturulan C_k kümesine yerleştirilir.
4. Eğer tüm örnekler kümelere atanmışsa işlem bitirilir. Atama işlemi bitmemiş ise ikinci adıma dönülür.

Oluşturulan küme sayısı, K , t parametresinin bir fonksiyonudur ve t değeri arttıkça daha az küme oluşmaktadır.

2.2.3. Bilgisayarda programlama

Kümeler analizi hem SPSS programı kullanılarak doğrudan menüler yardımıyla hem de MATLAB programında uygun algoritma yazılarak gerçekleştirilmiştir.

2.2.3.1. SPSS ile kümeleme analizi

SPSS programı ile kümeler analizi gerçekleştirilirken Şekil 2.3'de verildiği gibi örneklerin bileşen dağılımları ve geldikleri yöreler veri olarak programa girildikten sonra;

Analyze>Classify>Hierarchical Cluster Analysis> menüleri üzerinden

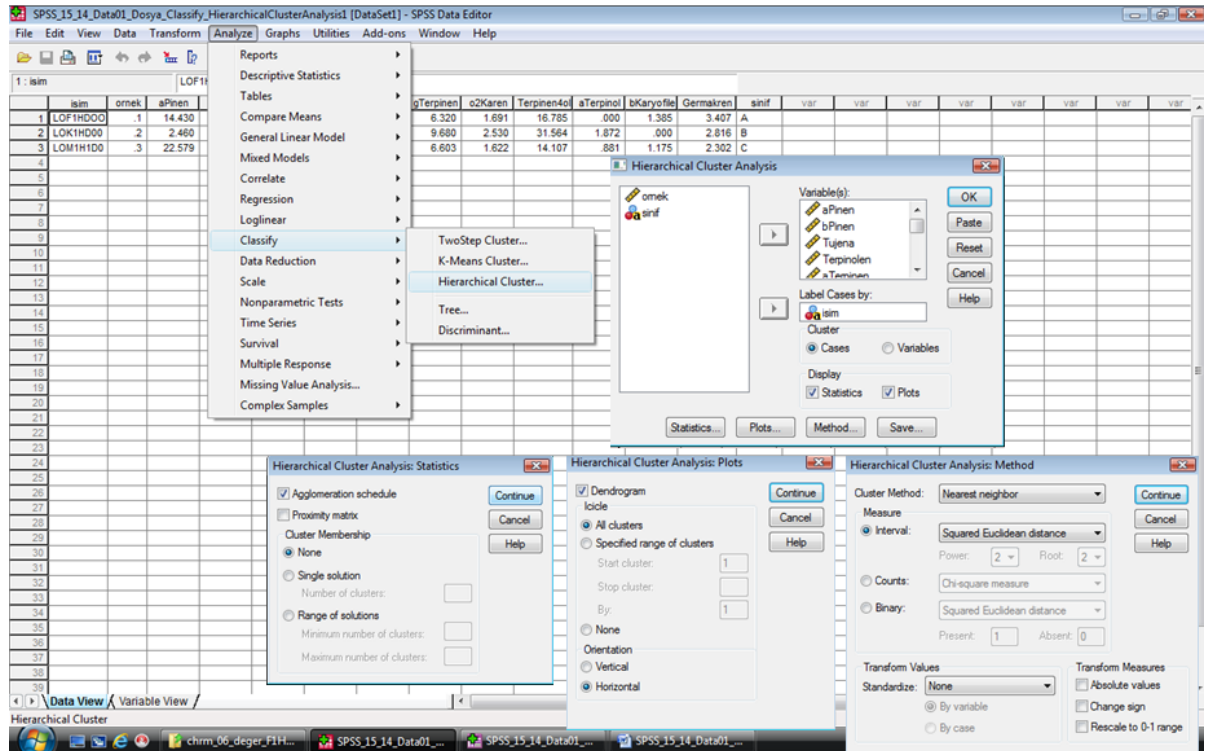
Cluster:Statistics-Plots-Method ayarları yapılmıştır. Sonuçlar da program çıktıları olarak alınmıştır.

2.2.3.2. MATLAB ile kümeleme analizi

MATLAB70 programı ile kümeler analizi gerçekleştirilirken Tablo 2.1’de verildiği gibi örneklerin bileşen dağılımları ve geldikleri yöreler veri olarak Excel Dosyası ile programda okutulmak üzere hazırlandıktan sonra, aşağıdaki algoritma kullanılmıştır.

Çizelge 2.1. Örneklerin bileşen dağılımları ve geldikleri yöreler

0	Isim	Ornek	a-Pinen	b-Pinen			Germakren D	Sinif
1	LOF1HD	0.1	14.4298	36.6995			3.40653	A
2	LOK1HD	0.2	2.45961	3.83932			2.81633	B
3	LOM1HD	0.3	22.5788	26.6416			2.30209	C



Şekil 2.3. Kümeler analizi için SPSS140 menüleri

```

'% CommandWindow'
'% MATLAB_704_Data01_Dosya_Classify_HierarchicalClusterAnalysis1.m*'
clc , clear , clf
'% Veri matrisi '
'MATLAB_704_Data01_Dosya_Classify_HierarchicalClusterAnalysis1.xls'
Orn_GC_MS_01=xlsread ...
('MATLAB_704_Data01_Dosya_Classify_HierarchicalClusterAnalysis1.xls')
'% Bileşenlerin Yeniden Düzenlenmesi'
'% Veri matrisi '
X=Orn_GC_MS_01(2:4,4:14);X

'% x veri matrisi için hesaplanan öklid uzaklıkları'
Y = pdist(X)
'% y öklid uzaklık vektörünün karesel formu'
squareform(Y)
'% Hesaplanan gruplararası öklid uzaklıkları'
Z = linkage(Y)

'% Z matrisi için dendrogramın çizilmesi'
dendrogram(Z) ,
xlabel('% Organik Bileşik Dağılımı'), ylabel('Dağılım Kumeleri')
title('MATLAB_704_Data01_Dosya_Classify_HierarchicalClusterAnalysis1')

```

Sonuçlar da, sayısal program çıktıları olarak MATLAB Command Window ara yüzeyinden, görsel olarak da çizdirilen Dendrogramlar MATLAB Figure penceresinden alınmıştır.

3. BULGULAR VE İRDELEME

Yapılan çalışmada; kabukları yerine değerlendirilmek amacıyla farklı coğrafi alanlardan aynı dönemlerde toplanan sığla ağacı yapraklarından alınan örneklerin, hidrodestilasyon, yöntemiyle ayrıştırılan uçucu yağlarında ki bileşenlerinin kalitatif ve kantitatif kromatografik analizleri GC-MS kullanılarak yapılan örneklerdeki uçucu yağların majör bileşenlerdeki değişimi tayin edilmiş olup, farklı yörelerde doğal olarak yetişen Günlük Ağacının eş zamanlı hasat edilen yapraklarının majör bileşenlerindeki değişimler kümelenmiş ve bu kümelerin geldikleri yöreye göre gruplanmasının uçucu yağ bileşimi ile ilgili bağlantısı saptanmıştır

3.1. Gaz Kromatografisi Kütle Spektroskopisi Analizleri

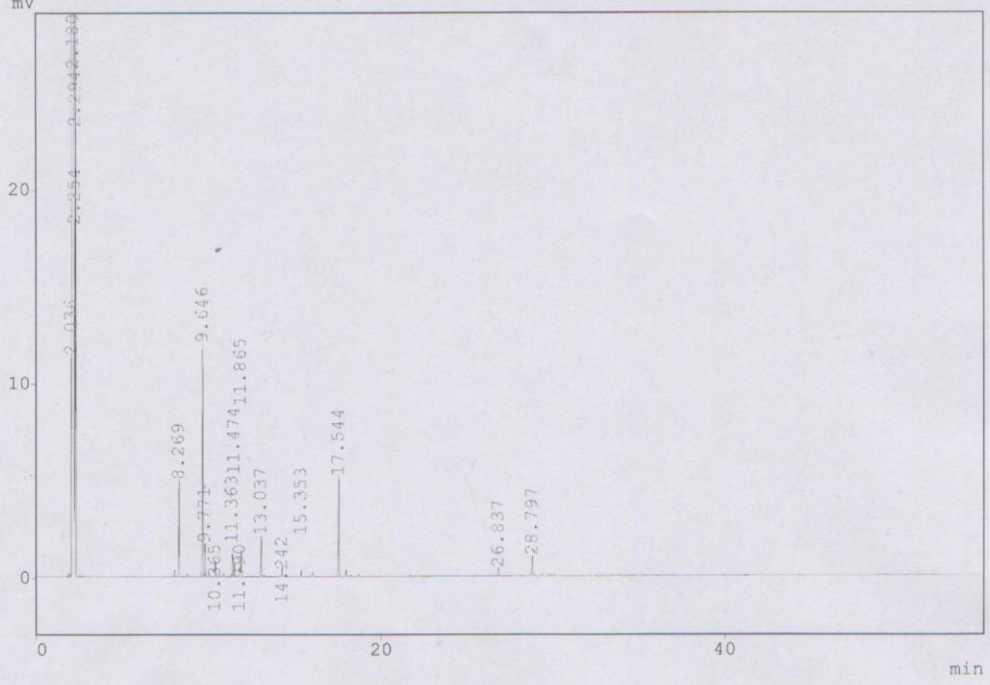
Bu çalışmada üç ayrı alandan toplanan örnekler Fethiye (LOF1HD00), Köyceğiz (LOK1HD00) ve Marmaris (LOM1HD00)'den 2011 Temmuz ayında ağaçların yapraklarından hasat edilmiştir. Hidrodestilasyon, yöntemiyle ayrıştırılan uçucu yağların miktar ve bileşenlerini belirlemek için yapılan gaz kromatografisi tayinleri sonucu alınan kromatogramlara ilişkin bir pik raporları Şekil 3.1.a, Şekil 3.1.b, Şekil 3.1.c'de verilmiştir. Analiz sonucu alınan karışımların bileşimindeki majör bileşenler MS destekli olarak incelenmiş ve ayrıntılı olarak saptanmıştır. Şekil 3.2'de üç örneğe ilişkin kütle spektrumları verilmiş olup, bu spektrumlar majör bileşenlerin daha etkin belirlenebilmesi için o bileşenlerin olduğu bölgeler daha büyütülerek takip edilmiştir. Major bileşenler için spektrumların ayrıntılı görünimleri Şekil 3.3'de sunulmuştur. Örneklerle göre majör bileşenlerin dağılımı ise Çizelge 3.1'de verilmiştir.

Çizelge 3.1. Örneklere göre majör bileşenlerin dağılımı

Sıra (Örnek)	1 (A)	2 (B)	3 (C)
İsim	LOF1H.D00	LOK1H.D00	LOM1H1.D00
a-Pinen	14.42980463	2.45961	22.5788
b-Pinen	36.69948003	3.83932	26.6416
Tujen-a	4.87859054	0	7.6542
Terpinolen	0	3.40209	0
a-Terpinen	3.445908425	1.83499	3.98308
g-Terpinen	6.320359433	9.6802	6.6033
2-Karen	1.691150487	2.5304	1.62207
Terpinen-4-ol	16.78529961	31.564	14.1071
a-Terpinol	0	1.87247	0.88089
b-Karyofilen	1.385228936	0	1.17489
Germakren D	3.406532384	2.81633	2.30209

CLASS-GC10 Ver.=2.01 SYS=1 Ch=1 REPORT.NO=4 DATA=LOF1H.D00 11/12/01 13:48:34
Sample : LOFthye1.ksm.h
ID : LOFthye1.ksm.h
Dilution Factor: 1
Type : Unknown
Detector : FID
Operator : GulsenTEL

*** Chromatogram *** Filename:LOF1H.C00



*** Peak Report ***

PKNO	TIME	AREA	HEIGHT	MK	IDNO	CONC	NAME
1	2.036	20916	11482	V		0.2429	
2	2.180	8291159	1103307	VE		96.2901	
3	2.254	33916	18345	V		0.3939	
4	2.294	165573	165456	V		1.9229	
5	8.269	14292	5027			0.1660	
6	9.646	36349	12075			0.4221	
7	9.771	4832	1722	V		0.0561	
8	10.365	2306	831			0.0268	
9	11.363	3413	1170			0.0396	
10	11.474	2175	745	V		0.0253	
11	11.790	1286	452			0.0149	
12	11.865	4049	1389	V		0.0470	
13	13.037	6260	2119			0.0727	
14	14.242	1675	556			0.0195	
15	15.353	1037	346			0.0120	
16	17.544	16625	5183			0.1931	
17	26.837	1372	414			0.0159	
18	28.797	3374	1033			0.0392	

8610609 1331651 100.0000

*** Calibration Result ***

** ID Table **

IDNO	NAME	TIME	CONC	RF1	RF2
------	------	------	------	-----	-----

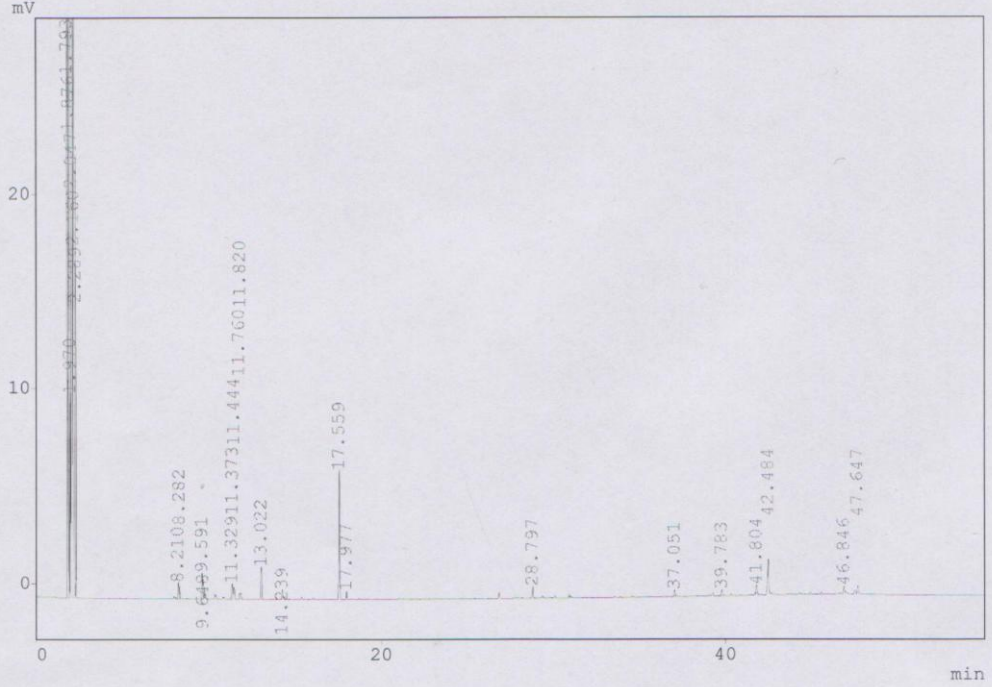
4 - 1/1

12/04/19 19:05:50

Şekil 3.1.a. GC tayinleri sonucu alınan kromatogramlara ilişkin pik raporu örneği (LOF1HD00)

CLASS-GC10 Ver.=2.01 SYS=1 Ch=1 REPORT.NO=2 DATA=LOK1H.D00 11/11/30 10:55:14
Sample : LOKycgz1.ksmHydro
ID : LOKycgz1.ksmHyd
Dilution Factor: 1
Type : Unknown
Detector : FID
Operator : GulenTEl

*** Chromatogram *** Filename:LOK1H.C00



*** Peak Report ***

PKNO	TIME	AREA	HEIGHT	MK	IDNO	CONC	NAME
1	1.793	148771	68867			2.1930	
2	1.876	900985	673634	V		13.2812	
3	1.970	20121	10509	V		0.2966	
4	2.047	3519286	1101614	VE		51.8769	
5	2.160	2099196	1100140	VE		30.9437	
6	2.289	23525	22727	V		0.3468	
7	8.210	2428	828			0.0358	
8	8.282	1772	680	V		0.0261	
9	9.591	3998	1351			0.0589	
10	9.648	2766	1046	V		0.0408	
11	11.329	2451	792			0.0361	
12	11.373	1322	622	V		0.0195	
13	11.444	2894	609	V		0.0427	
14	11.760	1036	289			0.0153	
15	11.820	1532	312	V		0.0226	
16	13.022	6974	1663			0.1028	
17	14.239	1823	493			0.0269	
18	17.559	22740	6648			0.3352	
19	17.977	1349	416			0.0199	
20	28.797	2029	605			0.0299	
21	37.051	1260	331			0.0186	
22	39.783	1058	291			0.0156	
23	41.804	1829	484	V		0.0270	
24	42.484	9036	1795			0.1332	
25	46.846	1945	380			0.0287	
26	47.647	1802	395	V		0.0266	

6783926 2997520 100.0000

*** Calibration Result ***

** ID Table **

IDNO NAME TIME CONC RF1 RF2

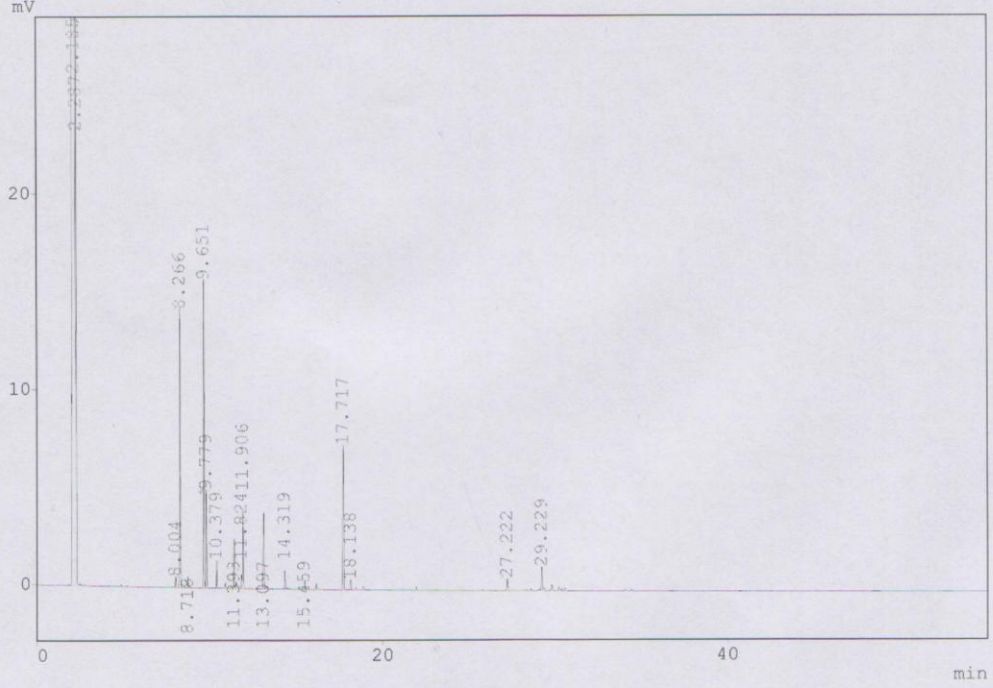
2 - 1/2

12/04/19 19:03:27

Şekil 3.1.b. GC tayinleri sonucu alınan kromatogramlara ilişkin pik raporu örneği
(LOK1HD00)

CLASS-GC10 Ver.=2.01 SYS=1 Ch=1 REPORT.NO=3 DATA=LOM1H1.D00 11/12/05 10:07:44
Sample : LOMrmrs1.ksm.h1
ID : LOMrmrs1.ksm.h1
Dilution Factor: 1
Type : Unknown
Detector : FID
Operator : GulsenTEL

*** Chromatogram *** Filename:LOM1H1.C00



*** Peak Report ***

PKNO	TIME	AREA	HEIGHT	MK	IDNO	CONC	NAME
1	2.185	9889363	1098982	E		95.4399	
2	2.287	294291	297999	V		2.8401	
3	8.004	1440	555			0.0139	
4	8.266	40242	14361			0.3884	
5	8.718	1522	562			0.0147	
6	9.651	47483	15808			0.4583	
7	9.779	13642	5045	V		0.1317	
8	10.379	4040	1435			0.0390	
9	11.393	7099	2565			0.0685	
10	11.824	2087	723			0.0201	
11	11.906	11678	3938	V		0.1127	
12	13.097	11769	3920			0.1136	
13	14.319	2891	958			0.0279	
14	15.459	1426	474			0.0138	
15	17.717	25143	7422			0.2427	
16	18.138	1570	512			0.0151	
17	27.222	2094	608			0.0202	
18	29.229	4103	1195			0.0396	

10361883 1457061 100.0000

*** Calibration Result ***

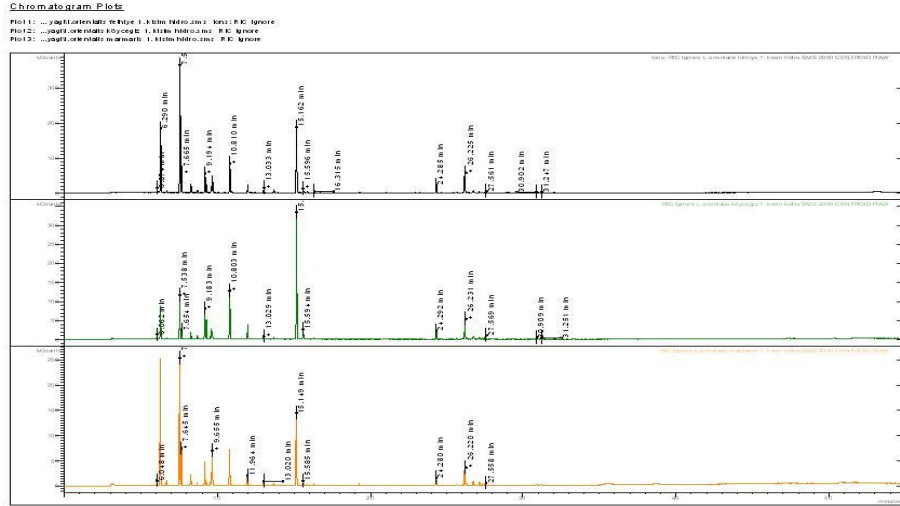
** ID Table **

IDNO	NAME	TIME	CONC	RF1	RF2
------	------	------	------	-----	-----

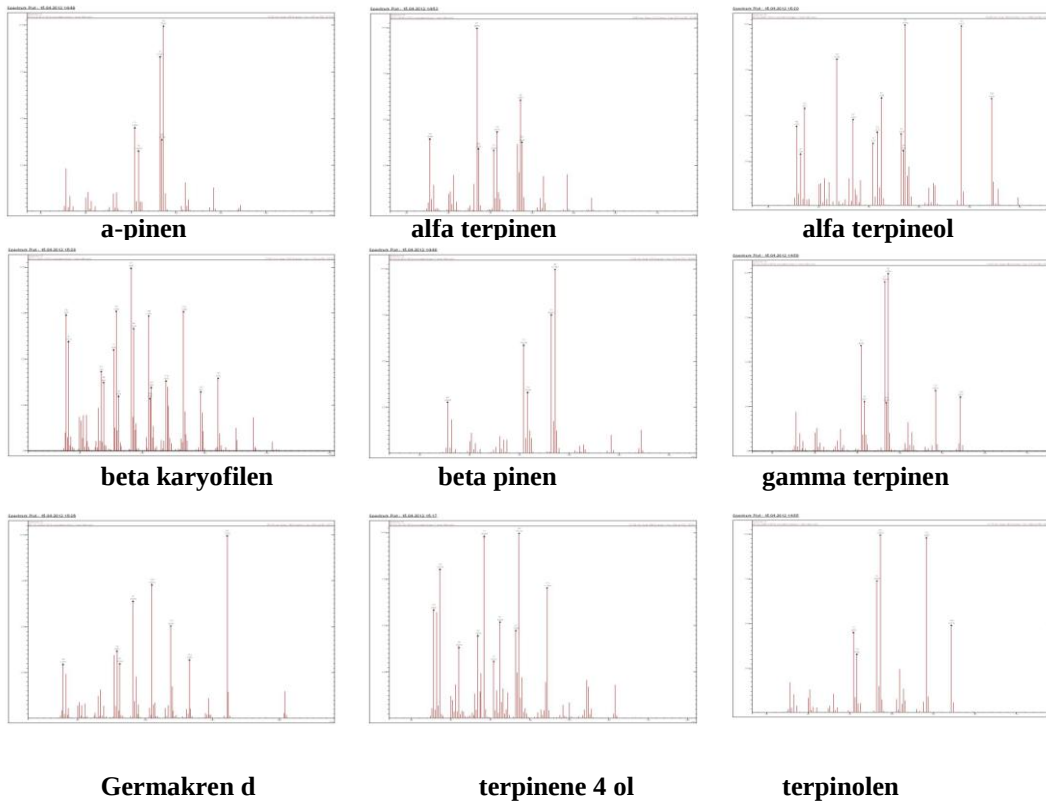
3 - 1/1

12/04/19 19:03:51

Şekil 3.1.c. GC tayinleri sonucu alınan kromatogramlara ilişkin pik raporu örneği
(LOM1HD00)



Şekil 3.2. Üç örneğin kütle spektrumları (LOF1HD00, LOK1HD00 ve LOM1HD00)



Şekil 3.3. Üç örneğin kütle spektrumları (LOF1HD00, LOK1HD00 ve LOM1HD00) üzerinde majör bileşenlerin ayrıntılı görünüşleri

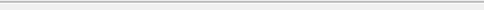
Üç farklı alandan toplanan örneklerin kromatografik analizleri incelendiğinde karşılaşılan 11 major bileşenin, karışımlarının yaklaşık % 70-80'lik büyük kısmını oluşturduğu anlaşılmaktadır. Çizelge 3.1'den görülebileceği gibi bu major bileşenler sırasıyla **a-Pinen, b- Pinen, Tujen-a, Terpinolen, a-Terpinen, g-Terpinen, 2-Karen, Terpinen-4-ol, a-Terpinol, b-Karyofilen ve Germakren D'**dir. Örneklerle asıl kimyasal ve fiziksel özelliklerini kazandıran major bileşiklerin varlığı olup, bunların bileşim oranları ise o bileşiklerin kümeleme analizi yardımıyla birbirlerine olan yakınlıklarını tanımlar ve böylece örneklerin hasat alanları veya hasat zamanları gibi parametreler çerçevesinde gruplaşmalarını ve böylece de coğrafi olarak yetiştirme bölgelerinin belirlenebilmesini, bunun sonucundaki ileri aşamalarda da yetiştirildikleri topraklar veya hasat zamanlarının optimum koşullarının saptanmasında etkin olarak rol alırlar. Bundan dolayı da çalışmanın devamında bu bileşenlere göre örneklerin yakınlıkları kümeleme teknikleri kullanılarak istatistiksel olarak değerlendirilmiştir. Bu değerlendirme hem SPSS14 bilgisayar paket programı hem de MATLAB da algoritma yazılımı ile gerçekleştirilmiştir.

3.2. İstatistiksel Değerlendirme

3.2.1. SPSS ile kümeleme analizi

Çizelge 3.2 örnek bileşimlerine ilişkin verilerin saklandığı SPSS14 veri dosyası olup, yapılan kümeleme analizi sonuçları ve dendrogram aşağıda verilmiştir.

Çizelge 3.2. SPSS14 veri dosyası

SPSS_15_14_Data01_Dosya_Classify_HierarchicalClusterAnalysis1 [DataSet1] - SPSS Data Editor														
File Edit View Data Transform Analyze Graphs Utilities Add-ons Window Help														
														
18 : GermakrenD														
	isim	ornek	aPinen	bPinen	Tujena	Terpinolen	aTerpinen	gTerpinen	o2Karen	Terpinen4ol	aTerpinol	bKaryofile	Germakren	sinif
1	LOF1HDOO	.1	14.430	37	5	0	3	6.320	1.691	16.785	.000	1.385	3.407	A
2	LOK1HD00	.2	2.460	4	0	3	2	9.680	2.530	31.564	1.872	.000	2.816	B
3	LOM1H1D0	.3	22.579	27	8	0	4	6.603	1.622	14.107	.881	1.175	2.302	C
4														

Cluster

[DataSet1]

C:\Users\Oguz\Desktop\S_C_Odabas\data1\chrn_06_deger_F1HDK1HDM1HD\SP
SS_15_14_Data01_Dosya_Classify_HierarchicalClusterAnalysis1.sav

Case Processing Summary^{a,b}

Cases					
Valid		Missing		Total	
N	Percent	N	Percent	N	Percent
3	100.0	0	.0	3	100.0

a. Squared Euclidean Distance used

b. Single Linkage

Single Linkage

Agglomeration Schedule

Stage	Cluster Combined		Coefficients	Stage Cluster First Appears		Next Stage
	Cluster 1	Cluster 2		Cluster 1	Cluster 2	
1	1	3	184.858	0	0	2
2	1	2	1317.164	1	0	0

Horizontal Icicle

Case	Number of clusters	
	1	2
2:LOK1HD00	X	X
	X	
3:LOM1H1D0	X	X
	X	X
1:LOF1HDOO	X	X

Dendrogram

***** HIERARCHICAL CLUSTER ANALYSIS *****

Dendrogram using Single Linkage

Rescaled Distance Cluster Combine

0	NaN	NaN	NaN	NaN	NaN	NaN
0.0000	NaN	0.0000	1.4430	3.6699	3.4065	4.8786
0.0000	NaN	0.0000	2.4596	3.8393	2.8163	0
0.0000	NaN	0.0000	2.2579	2.6642	2.3021	7.6542

.....

NaN	NaN	NaN	NaN	NaN	NaN	NaN
0	3.4459	6.3204	0.1691	1.6785	0	1.3852
3.4021	1.8350	9.6802	2.5304	0.3156	1.8725	0
0	3.9831	6.6033	1.6221	1.4107	0.0000	1.1749

% Bilesenlerin Yeniden Duzenlenmesi

X =

1.0e+014 *

1.443	3.669	4.878	0	3.445	6.320	0.169	1.678	0	1.385	3.406
0	9	6		9	4	1	5		2	5
2.459	3.839	0	3.402	1.835	9.680	2.530	0.315	1.872	0	2.816
6	3		1	0	2	4	6	5		3
2.257	2.664	7.654	0	3.983	6.603	1.622	1.410	0.000	1.174	2.302
9	2	2		1	3	1	7	0	9	1

% X veri matrisi için hesaplanan oklid uzakliklari

Y =

1.0e+014 *

7.9704	3.6325	9.6353
--------	--------	--------

% Y oklid uzaklik vektorunun karesel formu

1.0e+014 *

0	7.9704	3.632
7.9704	0	9.6353
3.6325	9.6353	0

% Hesaplanan gruplararasi oklid uzakliklari

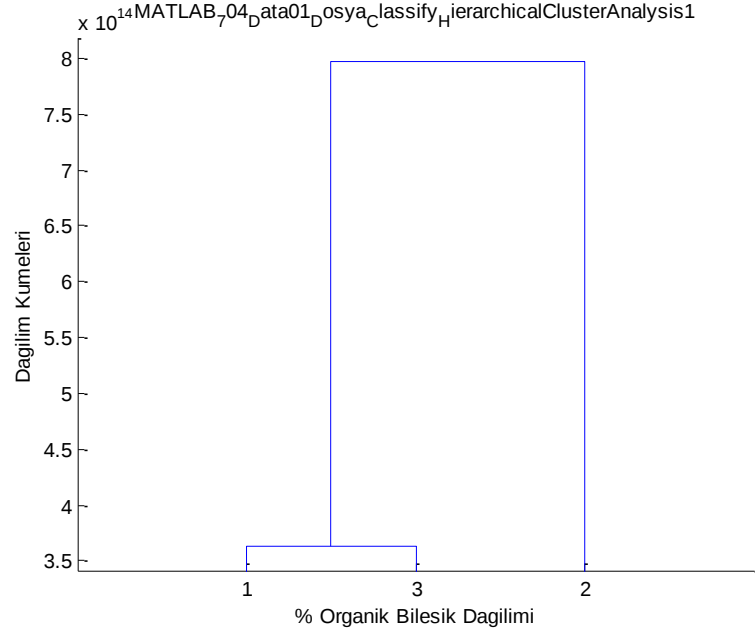
Z =

1.0e+014 *

0.0000	0.0000	3.6325
0.0000	0.0000	7.9704

% Z matrisi için dendrogramın çizilmesi

>>



Şekil 3.4. Üç örneğin (LOF1HD00 (1), LOK1HD00 (2) ve LOM1HD00 (3)) majör bileşenlerinin dağılımına göre yakınlık ilişkilerini veren kümeleme analizi

Her iki programın sonuçları aynı olup incelendiklerinde 1 ve 3 numaralı Fethiye ve Marmaris'ten gelen örneklerin çok yakın benzer özellikler taşıdığı, Köyceğiz'den gelenin ise diğerlerinden oldukça farklı olduğu açıkça görülebilmektedir.

4. SONUÇLAR VE ÖNERİLER

Veri tabanlarında farklı özelliklerdeki belli bir anlamı olmadan kayıtlı olan verilerin yorumlanarak anlamlı ve kullanışlı bilgiler haline getirilmesi veri madenciliğinin temelini oluşturmaktadır. Sınıflandırma ise veri madenciliğinin en çok kullanıldığı alandır ve sınıflandırmada yaklaşımlardan biri olan kümelemenin amacı ise heterojen olan ana kütleli homojen gruplara ayırmaktır. Kümeleme önceden tanımlanmış sınıfları esas alarak gruplamaz yani kümeleme gruplamada benzerlikleri kullanırken sınıflandırma önceden tanımlanmış sınıflar modelini temel alır. Yapılan herhangi bir araştırma sonrasında elde edilen verilerin analizinde ham veri çok büyük öneme sahiptir ve bu verilerin değerlendirilmelerine geçmeden önce, özellikleri dikkatle incelenmeli ve elde edilen veriler arasında birbirinden çok farklı değerler bulunuyorsa ya da değerler arası büyük dengesizlikler varsa, verilerin benzer birimlere çevrilmeleri veya ortalama değer etrafında toplanması gerekmektedir. Gelişen bilgisayar teknolojisi ve hazır istatistik paket programlarıyla bu konuda yapılacak olan sayısal işlemler oldukça kolaylaşmıştır. Bu amaca yönelik olarak da çok sayıda paket programda geliştirilmiştir ve bunlardan en çok bilinenleri ise SAS, SPSS, MATLAB, MINITAB, JMP, NTSYS, SYSTAT, MEGA, POPGEN ve MVSP' dir.

Bu çalışmanın amacı da; örnek obje olarak seçilen bir bitkinin yetiştirildiği ortam koşullarına veya hasat zamanına bağlı olarak değişen kimyasal özelliklerinin yapılacak analizlerle belirlenmesi ve bu özelliklere bağlı olarak, yetiştiği yörenin çevresel özelliklerinin veya mevsimsel olarak hasat zamanının saptanabilmesi için sınıflandırılmasında, kullanılacak olan kümeleme tekniklerinin incelenmesi ve bu bağlamda hazırlanan bu tez kapsamında da, kabukları yerine değerlendirilmek amacıyla farklı coğrafi alanlardan farklı zamanlarda toplanan sığla ağacı yapraklarından alınan örneklerin, hidrodestilasyon, yöntemiyle ayrıştırılan uçucu yağlarında ki bileşenlerinin kalitatif ve kantitatif kromatografik analizleri yapılarak örneklerdeki uçucu yağların majör bileşenlerdeki değişimi tayin edilmiş olup, farklı

yörelere doğal olarak yetişen Günlük Ağacının farklı zamanlarda hasat edilen yapraklarının majör bileşenlerindeki değişimler kümelenmiş ve bu kümelerin geldikleri yöreye göre gruplanması, toplanma zamanı ve ayırıştırma yöntemi gibi parametrelerin uçucu yağ bileşimi ile ilgili bağlantısı araştırılmıştır.

Çalışma iki bölümden oluşmaktadır. Birinci bölümde deneysel aşama, verilerin elde edilişi anlatılmaktadır. İkinci bölümde ise elde edilen deneysel verilerdeki majör bileşenlerin istatistiksel analizinin yapılmasıdır. Çalışmanın aşamaları aşağıda kısaca sunulmuştur.

1. Farklı yörelere ancak aynı dönemde toplanan ve doğal olarak yetişen Günlük ağacı yapraklarının majör bileşenlerindeki olabilecek değişimlere bakılması.
2. Elde edilen verilere göre majör bileşenler arasındaki değişimlerin örneklerin alındığı yörelere göre SPSS ve MATLAB programları yardımıyla kümeleme analizinin yapılması.

Araştırmada üç ayrı alandan toplanan örnekler Fethiye (LOF1HD00), Köyceğiz (LOK1HD00) ve Marmaris (LOM1HD00)'den 2011 Temmuz ayında ağaçların yapraklarından hasat edilmiştir. Hidrodestilasyon, yöntemiyle ayırıştırılan uçucu yağların miktar ve bileşenleri MS destekli olarak incelenmiş ve ayrıntılı olarak saptanmıştır. Üç farklı alandan toplanan örneklerin kromatografik analizleri incelendiğinde karşılaşılan 11 major bileşenin, karışımlarının yaklaşık % 70-80'lik büyük kısmını oluşturduğu anlaşılmaktadır. Bu major bileşenler sırasıyla **a-Pinen, b-Pinen, Tujen-a, Terpinolen, a-Terpinen, g-Terpinen, 2-Karen, Terpinen-4-ol, a-Terpinol, b-Karyofilen ve Germakren D'dir**. Örnekler asıl kimyasal ve fiziksel özelliklerini kazandıran major bileşiklerin varlığı olup, bunların bileşim oranları ise o bileşiklerin kümeleme analizi yardımıyla birbirlerine olan yakınlıklarını tanımlar ve böylece örneklerin hasat alanları veya hasat zamanları gibi parametreler çerçevesinde gruplaşmalarını ve böylece de coğrafi olarak yetiştirme bölgelerinin belirlenebilmesini, bunun sonucundaki ileri aşamalarda da yetiştirildikleri topraklar veya hasat zamanlarının optimum koşullarının saptanmasında etkin olarak rol alırlar.

Bundan dolayı da çalışmanın devamında bu bileşenlere göre örneklerin yakınlıkları kümeleme teknikleri kullanılarak istatistiksel olarak değerlendirilmiştir. Bu değerlendirme hem SPSS14 bilgisayar paket programı hem de MATLAB da

algoritma yazılımı ile gerçekleştirilmiştir. Her iki programın sonuçları beklenildiği gibi aynı olup incelendiklerinde 1 ve 3 nolu Fethiye ve Marmaris'ten gelen örneklerin çok yakın benzer özellikler taşıdığı, Köyceğiz'den gelenin ise diğerlerinden oldukça farklı olduğu açıkça görülebilmektedir. Araştırmanın devamında farklı zamanlarda hasat edilen ürünlerin, farklı yöntemlerle bileşimlerinin tayin edilmesi ve sonuçların kümelendirilerek hasat zamanı parametreleri ile ürünlerin ayrıştırılma yöntemlerinin optimizasyonu ve bu sonuçların bitki yetiştirilme koşulları ile ilişkilendirilmesinin hem bitkinin yetiştirilmesi hem de ürünlerin elde edilmesi aşamasında ekonomik açıdan önemli ölçüde katkılar sağlayacağı açıktır.

KAYNAKLAR

- Akpolat, O. (2010). *Matlab Uygulamaları ile Bilişim Teknolojileri*. Muğla Üniversitesi Yayınları, Muğla, 141-148s.
- Altuğ, T. ve Y. Elmacı, (1998). Gıdalarda Doğal Olarak Bulunan Lezzet Bileşenleri. *Gıda Kimyası*, Hacettepe Üniv. Yayınları, Ankara, 453-486.
- Anonim, Baharatların Sınıflandırılması, *Türkiye Gıdacılar Topluluğu*, (2003). Ankara.
- Arıcı, H., (2001). *İstatistik Yöntemler ve Uygulamaları*. Meteksan A.Ş. Ankara.
- Arifoğlu, U. (2005). *Matlab 7.04 Simulink ve Mühendislik Uygulamaları*, Alfa Ltd
- Berry, M.J.A., Linoff, G., (2004). *Data mining techniques: for marketing, sales, and customer relationship managemen*, Wiley.
- Berson, S., (1997). *Data Warehousing, Data Mining & OLA*, McGraw-Hill.
- Brereton R.G., (2003). *Chemometrics data analysis for the laboratory and chemical plant*, John Wiley & Sons, Ltd.
- Cao, F., (1998). *An Integrated Overview of Key Issues in Data Mining*, Yüksek Lisans Tezi, Computer Science, DalTech, Dalhousie University, Halifax, Nova Scotia.
- Ceylan, A., (1987). *Tıbbi Bitkiler 2 (Uçucu Yağ İçerenler)*, Ege Üniversitesi, Ziraat Fakültesi, Tarla Bitkileri Bölümü, 481:188, İzmir.
- Ceylan, A., (1997). *Tıbbi Bitkiler-II*, Ege Üniversitesi Ziraat Fakültesi Yayınları , İzmir.
- Çoban, T., (2009). *Farklı Yöntemlerle Elde Edilen Turunçgil Uçucu Yağ Komponentlerinin Karşılaştırılması*, Kimya, Adana.

- Demir, C., Tokatlı, F., Ertas, H., Özdemir, D., (2009). *Kemometri Yaz Okulu II Ders Notları*, I.Y.T.E. Fen Fak. Kim. Böl. İzmir.
- Fakhari, A. R., Salehi, P., Heydari, R., Ebrahimi, S. N., Haddad, P. R., (2005). *Hydrodistillation-Headspace Solvent Microextracti On, A New Method For Analysis Of The Essential Oil Components Of Lavandula angustifolia Mill.* J.of Chromatography A, 1098, 14-18.
- Fayyad, U.M., Piatetsky-Shapiro, G., Smyth, P., Uthurusamy, R.,(1996). *From data mining to knowledge discovery in database*, Advances in Knowledge Discovery and Data Mining, AAAI/MIT Press, Cambridge, Massachussets.
- Fayyad,U.; Shapiro,G.P.; Smyth,P.,(1996). *Knowledge Discovery and Data Mining: Towards a Unifying Framework*, Proceedings of the second Conference on Knowledge Discovery and Data Mining(KDD-96), Portland, Oregon, (Ağustos).
- Flynn J. R., (1999). The Discovery of IQ Gains Over Time, Searching for Justice, *American Psychologist*, 5-20.
- Gevrekçi Y., Ataç F. E., Takma Ç., Akbaş Y., (2011). Classification of West Anatolian Cities for Sheep Production, *Kafkas Univ. Vet. Fak. Derg.*, 17 (5): 755-760
- Grabmeier, J. ve Rudolph, A., (2002). Techniques of Cluster Algorithms in Data Mining, *In Data Mining and Knowledge Discovery*, No.4, s.340.
- Han, J. ve Kamber, M., (2001). *Data Mining Concepts and Techniques.*”, Morgan Kauffmann Publishers Inc.
- Han, J.; Kamber, M.; Tung, A. K. H. , (2001). Spatial Clustering Methods in Data Mining: A Survey, in H. Miller and J. Han (eds.), *Geographic Data Mining and Knowledge Discovery*, Taylor and Francis.
- He, X., Bigelow, K.A., Boggs, C.H., (1997). Cluster Analysis of Longline Sets and Fishing Strategies within the Hawaii-Based Fishery, *Fisheries Research*, No:31, 147-158p.

Heller, G., ve E.S. Venkatraman, (2004). *Statistical Methodology*, Nonparametric Test to Compare Survival Distributions With Covariate Adjustment, *Journal of the Royal Statistical Society*, 530–813.

Hernandez, O.M., Fraga, J.M.G., Jimenez, A.I., Arias, J.J., (2005). Characterization of Honey From The canary Islands: Determination of The Mineral content by Atomic Absorption Spectrophotometry, *Food Chemistry*, 449-458.

Huang, Tung-Chun, (2001). The Effects of Linkage Between Business and Human Resource Managment Strategies, *Personal Review*, No.2, s.11.

Huř, S., (1949). Sıęla Ağacının (Liquidambar Orientalis Mill) Ormancılık Bakımından Önemli ve Sıęla Yaęının Kimyasal Arařtırılması, *Orman Genel Müdürlüęü Yayınları*, Sayı:83

http://tr.wikipedia.org/wiki/Anadolu/Sıęla_aęacı, 12.08.2011

Jain, A.K., Murty, M.N.; Flynn, P.J., (1999). Data Clustering: A Review, *ACM Computing Surveys*, No. 3., (September)

Jain,A.K., Dubes,R.C., (1998). Algorithms For Clustering Data, *Prentice Hall*, Englewood Cliffs, New Jersey, 07632.

Karypis, G., Han,E.H.; Kumar,V. , (1999). CHAMELEON: A hierarchical clustering algorithm using dynamic modeling, *IEEE Computer*, 32(8).

Kaufman, L., Rousseuw, P., (1990). *Finding Groups Data : An Introduction to Cluster Analysis*, John Wiley and Sons Inc., New York.

Kaya, H., Köymen, K., (2008). *Veri Madencilięi Kavramı ve Uygulama Alanları, Doęu Bölgeleri Arařtırması*.

Kaya, Z. ve Alan, M., (2003). EUFORGEN Technical Guidelines for genetic conservation and use for oriental sweetgum (Liquidambar orientalis), *International Plant Genetic Resources Institute*, Rome. Italy.

King, B., (1967). Step-wise clustering procedures, *Journal of the American Statistical Association*, 69:86-101.

- Koedam, A., Scheffer, J.J.C., (1979). *Comparison of Isolation procedures for Essential Oils.II. Ajowan, Caraway, Coriander and Cumin*, Z. Lebensm. Forsch, 168,106.
- Koedam, A., Scheffer, J.J.C., (1980). Monoterpenes in the Volatile Leaf Oil of *Abiesx arnoldiana* Nitz., *J. Agric. Food Chem.*, 28, p.862.
- Koedam, A., Scheffer, J.J.C., (1981). Comparison of Isolation procedures for Essential Oils.IV. Leyland cypress. *Perfumer and Flavorist* 5.56.
- Kohavi,R.; Provost,F., (2001). *Applications of Data Mining to Electronic Commerce*, Kluwer Academic Publishers Inc.
- Korum, U., (1991). *Sosyal Bilimlerde İstatistik*, Turhan Kitabevi, Ankara.
- Lawrance, M. B., (1995). *The Isolation of Aromatic Materials from Natural Plants Products, A Manual On The Essential Oil Industry*, Anadolu Üniversitesi Basımevi, Eskişehir.
- Linskens, H.F., Jackson, J.F, (1997). Modern Methods of Plant Analysis, *Plant Volatile Analysis*, Springer, Germany.
- Ludwig, J.A. and Reynolds J. F.. (1988). *Statistical ecology: a primer of methods and computing*, Wiley Press, New York, New York. 337 p.
- Maes, F., (1998). *Segmentation and Registration of Multimodal Medical Images: from Theory, Implementation and Validation to a Useful Tool in Clinical Practice*, PhD thesis, Catholic University of Leuven, Belgium.
- Manly, Bryan F.J. , (1994). *Multivariate Statistical Methods. Second Edition*, Londra: Chapman – Hall
- McCune, B., Grace, J. B. ve Urban, D. L., (2002), *Analysis of Ecological Communities*, Gleneden Beach, Oregon: MjM Software Design.
- Mercer, N., (2005). Sociocultural discourse analysis: analysing classroom talk as a social mode of thinking, *Journal of Applied Linguistics*, 137-168.

- Moshkovich, H. M., Mechitov, A. I. ve Olson, D. L., (2002). Rule Induction in Data Mining: Effect of Ordinal Scales, *Expert Systems with Applications*, No: 4, s.303.
- Oğuzlar, A., (2003). Veri Ön İşleme, *Erciyes Üniversitesi İ.İ.B.F. Dergisi*, Sayı.21, (Temmuz-Aralık).
- Orekici, T.G., Çamdeviren H., Akkuş Z., (2005). *Sınıflama Ağaçları Yardımıyla Restless Legs Syndrome (RLS) Hastalarına Tanı Koyma*, Yüksek Lisans, Mersin Üniversitesi, Mersin.
- Özdamar, K., (1999). *Paket Programlar ile İstatistiksel Veri Analizi*, Eskişehir, Kaan Kitabevi.
- Özkan, Y., (2008). *Veri Madenciliği Yöntemleri*, Papatya Yayıncılık, Ankara.
- Özmen, Ş., (2003). Veri Madenciliği Süreci, *Veri Madenciliği ve Uygulama Alanları*, İstanbul Ticaret Üniversitesi, İstanbul, Mayıs, s.37.
- Pamukcuoğlu, A., (1964). Memleketimizin Liquidambar orientalis Orman Sahası. *Türk Biyoloji Dergisi*, 14(2).
- Pierce, N.E., (2006). Molecular phylogeny and systematics of the Pieridae (Lepidoptera: Papilionoidea): higher classification and biogeography. *Zoological Journal of the Linnaean Society*, 147: 239-275
- Polat, Ö. ve Ötleş, S., (1997). Esansiyel Yağların Ekstraksiyon Yöntemleri, *Ege Üniversitesi Mühendislik Fakültesi Dergisi*, Seri:B, Gıda Mühendisliği, 15(1-2), 155 – 161.
- Rangahau, M. K., (2001). Essential oils and their production. *Crop and Food Research*, Nr. 39, October.
- Reineccius, G., (2005). Flavor Chemistry and Technology, Taylor and Francis Group, A CRC Press Book, 2nd edition, 216 – 219.
- Sharma, S., (1996). *Applied Multivariate Techniques*, John Wiley and Sons, Inc., 493p. USA.

- Shearer, C., (2000). The CRISP-DM Model: ‘The New Blueprint for Data Mining’, *Journal of Data Warehousing*, No.4.
- Shinn, C., (1999). The right to the highest attainable standart of health: Public health’s opportunity to reframe a human rights debate in the USA, *Health Hum Rights*.
- Simoudis, E., (1996). Reality check for data mining, *Intelligent systems and their applications*, 11 (5):26-33.
- Sneath, P. H. A., Sokal, R. R., (1973). *Numerical Taxonomy*, Freeman, San Fransisco, California, 114-253.
- Şenocak, M., (1990). *Temel Biyoistatistik*, Çağlayan Kitabevi, İstanbul.
- Tanker, M., Tanker, N., (1990). *Farmakognozi*, Ankara Üniversitesi Eczacılık Fakültesi Yayını, Ankara.
- Tserper, G., Peristeraki, P., Po Tamas, G., Tsimnides, N., (1999). Species Distribution in The Southern Aegean Sea Based on Bottom-Trawl Surveys, *Aquat. Living Re sour.*, 167-175.
- Turanlı, M., GÜRİŞ, S., (2000). *Temel İstatistik*, DER yayınları, İstanbul.
- Tümen, G., Satıl, F., Yıldırım, O., (1992). *Trakya ve Orta Anadolu’da Yetişen Satureja Türlerinin Sistematik Revizyonu ve Uçucu Yağların Kimyasal Bileşimlerinin Mukayesesi*, Yüksek Lİsans, Balıkesir Üniversitesi, Balıkesir.
- Voncina, D.B., Kodba, Z.C., Novic, M., (2005). Multivari ate Data Analysis in Classification of Vegetable Oils Characterized by The Content of Fatty acids. *Chemometrics and Intelligent la boratory Systems*, 31-43p.
- Ward, J. H. (1963). Hierarchical grouping to optimize an objective function, *Journal of the American Statistical Association*, 58:236-24.
- Witten, I. H. and Frank, E. (1999). *Data Mining: Practical machine learning tools with Java implementations*, Morgan Kaufmann, San Francisco.

- Woodman, N., (2000). *Cryptotis merriami* Choate in Costa Rica: Syntopy with *Cryptotis nigrescens* (Allen) and Possible Character Displacement (Mammalia: Insectivora). *Caribbean Journal of Science*, No:3-4, 289-299.
- Yazgan, E., Kayaalp, G.T., (2002). Kümeleme (Cluster) Analizi Yöntemlerinin Karşılaştırmalı Olarak İncelenmesi ve Tarımsal Araştırmalarda Kullanılması *Çuk. Üniv. Ziraat Fak. Dergisi* 17 (3): 7- 14.

Ek 1. İstatistiğe İlişkin Bazı Temel Tanım Ve Kavramlar

İstatistikte bilinip denetlenen ya da bilinmeyen ve çeşitli etkenlerin ara etkileri gibi rastlantısal olarak farklılıklar gösterebilen birimler "değişken" adını alır. Eğer bunlar, boy, kilo gibi ölçülebilir nesnelerden oluşuyorsa 'nicel' özellik, değilse 'nitel' özellik (kan grubu, evli olmak gibi) adı verilir ve ancak tam sayılarla gösterilebilirler. Eğer bu değişkenlerin belirli değerleri, belirli olasılıkları alma kurallarına uyuyor ve bu olasılıklar belirli olasılık fonksiyonları ile ifade edilebiliyorsa, (aynı ilacı alan iki farklı kişinin kan örnek analizlerinde gözlenen değerlerin farklılığı) bu değişkenler şans değişkenleri olarak adlandırılır ve sürekli, kesikli ve kategorik olmak üzere üç gruba ayrılırlar. Aşağıda bu temel kavramlar üzerinde kısaca durulacaktır (Şenocak, 1990; Akpolat, 2010).

Ölçümlere gelince, bunlarda oransal ölçüm (ölçüm aralığı sonsuz küçük, sıralı ve başlangıç noktası özelliğine sahip- boy, yaş, ağırlık vs.), sınıfsal ölçüm (adsal veya nominal ölçüm-sınıfları altında/kan grubu), sıralayıcı ölçüm (nitel özellikler arasında mantıksal ilişkiler-zayıf/orta/şişman gibi) ve aralıklı ölçümler (0-4,4-9,9-15 yaş grupları) biçimindedir. İstatistiksel karar ise kısaca nedensellik ilişkisinin açıklama çabasıdır ve bu da aşağıdaki gibi özetlenebilir:

1. Araştırmacı değerlendirmeden önce ön kabulü belirler (normallik payı). Biyolojik bilimlerde bu genel olarak 0.95'tir. ($p > 0.05$), ancak bu araştırmanın özelliğine göre değişebilmektedir. Kalan %5'lik bölümün anormallik alanı olduğu düşünülürse, istatistiksel yargılamanın temeli, istatistiksel hesaplar sonucu bulunmuş olan bir (p) olasılığının durumuna bağlıdır.

2. Bu alanın dışında ($p < 0.05$) oluşumların belirtisi olan olasılıklar, kıyaslanan değerlerin 'anamlı olarak farklı' olduğunu vurgular ve %95 güven aralığı içinde hesaplamalar sonucu elde edilen olasılığın büyüklüğüne göre yapılacak yorumlar aşağıdaki gibi sıralanabilir.

$p > 0.05$	anamlı fark yok (güven aralığında)
$p < 0.05$	anamlı fark
$p < 0.01$	ileri düzeyde anamlı fark
$p < 0.001$	çok ileri düzeyde anamlı fark

İstatistiksel hesaplamalar, konunun özelliğine göre çeşitli testler kullanılarak yapılır ve bu hesaplamaların sonuçları bu testlerin özel simgelerinin (t , z , F , χ^2 ... vs) karşılığı olarak belirlenir.

3. Değerlendirmede anamlı bulunabilecek en düşük olasılık sınırı test çizelgelerinde aranır ve farklılıklar etkenlerin ve özelliklerin konumuna göre, gerçek veriler göz önünde bulundurularak yapılır. Her testin, araştırmacının deney düzeyine (veya diğer ölçüm yöntemlerine) bağlı olarak (deney sayısı, gözlem sayısı vs.) ve bu özelliklere belirlediği duruma ait serbestlik derecesi içerisinde bir olasılık düzeyi belirlenir. Örneğin; 14 serbestlik derecesinde t testinde yapılan çalışmada $t=2.17$ bulunmuş ise buna bağlı olasılık $p < 0.05$; $t=0.23$ bulunmuş ise $p > 0.05$ düzeylerindedir. Bu değerlere çeşitli kaynaklarda bulunan istatistik değerlendirme Çizelgelerinden

kolayca ulaşılabilir. Ancak her testin kullanımı ayrı ayrı ve dikkatlice öğrenilmelidir (Çizelge 2.1). Yapılan değerlendirme sonuçları listeler, Çizelgeler (frekans) ve grafikler (veya fonksiyonlar $y=f(x)$) ile sunulmalı ve sonuçlar mutlaka kullanılan testlerle birlikte elde edilen olasılık düzeylerine göre yorumlanmalıdır.

Çizelge Ek 1.1. İstatistiksel değerlendirme için uygun test seçim kılavuzu

	Nicel özellikler :Sürekli birimlerle belirlenen		Nitel özellikler:Adsal, sıralı aralıklı ölçek,Süreksiz adetlerle belirlenen
KONU	Parametrik	Non-parametrik	
Varyans Kıyaslama Toplam Örnek Grubu Kıyas.	F (testi) z, t	Medyan için İşaret testi	Adsal ölçek : χ^2 Diğer: KOLMOGOROV-SMIRNOV
İKİ n_1 ve $n_2 > 30$ n_1 ve/veya $n_2 < 31$ ÖRNEK varyans eş düzenli n_1 ve/veya $n_2 < 81$ GRUBU varyans farklı	t Ortak SD kullanarak t t ve SOA veya t'	MANN-WHITNEYU MANN-WHITNEYU	Denek sayılarının durumuna göre; Adsal ölçekte: χ^2 , Yates, x^2 , FISHER Sıralı-aralıklı ölçekte KOLMOGOROV-SMIRNOV Ayrıntılı bilgi gerekli
EŞLENDİRİLMİŞ DİZİ	t (eşli diziler için)	WILCOXON	MC_NEMAR χ^2 Ayrıntılı konular: Tanı testleri
İKİDEN FAZLA GRUP KIYASLA.	Varyans Çözümleme		χ^2 (uygun şekilde)
BAĞINTI	Korelasyon (r) (Regresyon)	Spearman (r _s) (sadece korelasyon için)	χ^2 , cc, Q

Çizelge Ek 1.1'den görüleceği gibi nicel ve nitel özelliklerin değerlendirilmesinden önce aşağıda kısaca açıklanan bazı temel teknik terimlerin anlaşılması yararlı olacaktır. Belirli bir değişken bakımından yapılan gözlemler sonucu elde edilen verilerin anlamlı bir şekilde özetlenme türlerinden biri frekans dağılımları (Çizelgeleri) ve frekans grafikleridir. Örneğin; 20 kişilik bir toplulukta yapılan çürük diş sayısı değerleri 2,0,1,3,2,1,0,3,2,1,1,2,2,4,1,1,2,3,0 ve 1 olsun. Aşağıda verilen çürük diş sayılarına ve frekans (sıklık) dağılımına göre Şekil E1.1 de ki çürük diş frekans grafiğini çizilmiştir.

<u>Çürük diş Sayısı (x_i)</u>	<u>Frekans (a_i)</u>
0	3
1	7
2	6
3	.3
4	1
<u>≥4</u>	<u>0</u>
(n:Denek=20)	Toplam=20 ($\sum a_i = n$)

Temel tanımlar

Populasyon: Araştırma sınırları içindeki tüm insanlar grubu

Örnek: Araştırma sınırları içinde seçilen insanlar grubu ve,

Aritmetik ortalama :
$$\sum_{i=1}^n \frac{x_i}{n} = \bar{x} \quad (E1.1.)$$

Geometrik ortalama :
$$\left(\prod_{i=1}^n x_i \right)^{1/n} = \bar{x}_G \quad (E1.2.)$$

Harmonik ortalama :
$$\frac{1}{\left(\frac{1}{n} \sum_{i=1}^n \frac{1}{x_i} \right)} = \bar{x}_H \quad (E1.3.)$$

Ağırlıklı ortalama :
$$\frac{\sum_{i=1}^n a_i x_i}{\sum_{i=1}^n a_i} = \bar{x}_A \quad (E1.4.)$$

Medyan (Ortanca) :
$$3 \ 7 \ 6 \ 3 \ 1 \ 0 \quad 3+3+1+0=7 \quad (E1.5.)$$

$$7=7 \quad (E1.6.)$$

Tepe değeri :
$$3 \ 7 \ 6 \ 3 \ 1 \ 0 \quad (E1.7.)$$

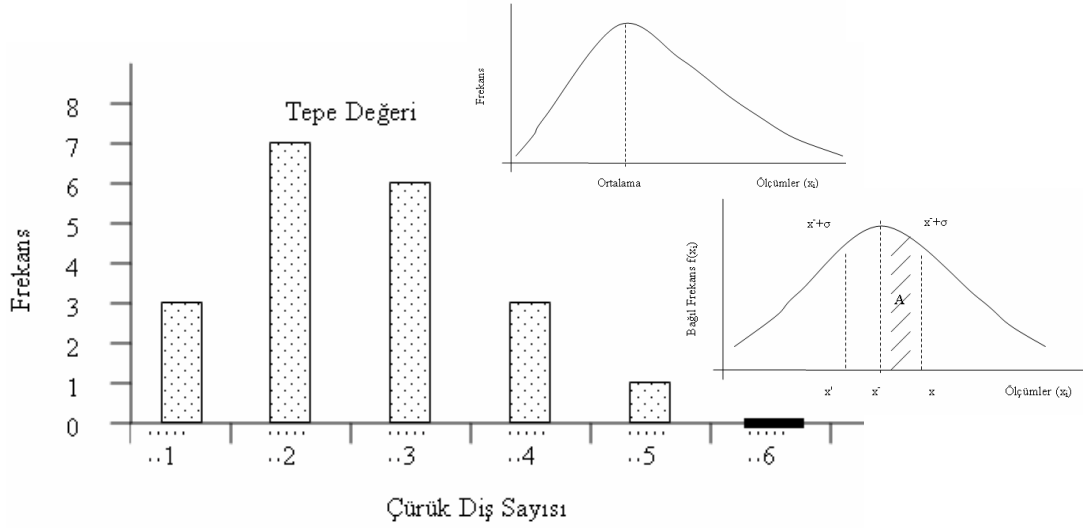
($x_{\max}$ - $x_{\min}$) Değişim aralığı:
$$0 \ 1 \ 2 \ 3 \ 4 >4 \quad 4-0=4 \quad (E1.8.)$$

Varyans :
$$\frac{1}{n} \sum_{i=1}^n \left(x_i - \bar{x} \right)^2 \quad (E1.9.)$$

(σ) Standart sapma :
$$+ \sqrt{\frac{1}{n} \sum_{i=1}^n \left(x_i - \bar{x} \right)^2} \text{ veya} \quad (E1.10.)$$

(Örnek grubu için
$$+ \sqrt{\frac{1}{(n-1)} \sum_{i=1}^n \left(x_i - \bar{x} \right)^2} \quad (E1.11.)$$

olarak yazılır.



Şekil Ek 1.1. Çürük diş ölçümsel ve teorik frekans (sıklık) grafiği

Deneysel ya da ölçümsel alınan verilerin frekans dağılımı ve bu verilerin bir olasılık eğrisini oluşturacak şekilde (en yaygın olanı normal dağılımdır) düzenlenmesi sırasıyla Şekil Ek 1.1 de sunulmuştur. Bu matematiksel olarak, sürekli bir x_i şans değişkeni için yaklaşım yaptığı kabul edilen bir $f(x)$ olasılık fonksiyonu ile ifade edilir ve bağıl frekans olarak adlandırılır. Burada normal dağılım gösterdiği kabul edilen populasyon ortalaması $\bar{x}$ olmak üzere bağıl frekansı veren $f(x)$ fonksiyonu

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\bar{x})^2}{2\sigma^2}} \quad (-\infty < x < +\infty) \quad (E1.12.)$$

İstatistiksel yorumlar ve yapılan istatistiksel karşılaştırmalar için daha önce verilen kılavuz Çizelgesundan da anlaşılacağı gibi karşılaştırma amacına uygun olarak seçilebilecek muhtelif testler vardır. Ancak bu işlemlerden sonra yapılan deneylerin (ya da deney sonuçlarının) istatistiksel olarak ne düzeyde anlam ifade ettiği anlaşılabilir ve bu sonuçlar arasındaki uyum oranını belirlemek için regresyon korelasyon işlemine başvurulur. Bu model testlerden en basiti olanı t (student-t) testi olup, burada yalnızca örnek grubu kıyaslaması için kullanılan bu test üzerinde durulacaktır. 30 denemeden küçük gruplarda uygulanabilen bu testte t dağılışı normal dağılışı gibi ortalaması 0 olan simetrik dağılımdır. Tek parametresi serbestlik derecesi (n-1) olan ve $n < 30$ için varyans $(\frac{1}{n} \sum (x - \bar{x})^2)$, ve $n > 30$ için standart sapması

$\sqrt{\frac{\sum (x - \bar{x})^2}{n-1}}$ olarak kullanılan bu testin kullanım Çizelgesunda serbestlik derecesi ve istatistiksel dağılım fonksiyonu $f(t)$ bulunmaktadır. Her testte olduğu gibi burada da bir bilimsel hipotezin kurulup test edilmesinde başlıca 4 eleman vardır. Bunlar:

- a) sıfır hipotezi (H_0)
- b) test istatistiği (H)
- c) ret etme bölgesi
- d) alternatif hipotez (H_1 ve/veya H_2)

Burada bu durum kısaca örneklenirse:

Hipotez H_0 $x=x_0$ Serbestlik derecesi: $n-1$ (1 örnekli test)

Karşı hipotez H_1 $x>x_0$ Serbestlik derecesi: $n-1$

$$z = \frac{\bar{x} - x_0}{\frac{\sigma}{\sqrt{n}}} \quad n < 30 \quad t = \frac{\bar{x} - x_0}{\frac{\sigma}{\sqrt{n}}} \quad (\text{Küçük örnek grupları için})$$

Student t dağılışına ilişkin tek yönlü olarak verilen kritik değerler Çizelge E1.2’de sunulmuştur.

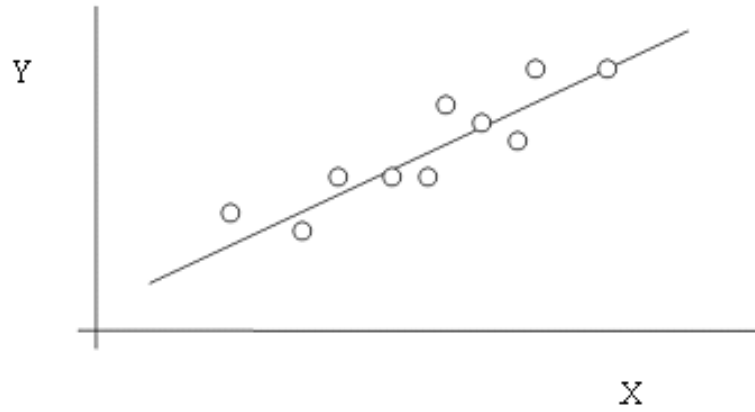
İstatistiksel olarak değerlendirilen değişkenlerden biri diğerlerine bağlı olarak $z=f(x,y,...t)$ gibi bir fonksiyonla ifade edilebiliyorsa bu değişken bağımsızdır ve bağımlı-bağımsız değişkenler arasındaki oluşturulan fonksiyonun bu noktalarla uyumu korelasyon olarak adlandırılır. Bu noktalara en küçük hata ile uygun bir fonksiyon modelinin türetilmesi de regresyon işlemi olarak adlandırılır. En basit model doğrusal olup, bu modeldeki fonksiyon $y=\mathbf{Ax}+\mathbf{B}$ şeklindedir ve uyum (korelasyon) katsayısı aşağıdaki bağıntı ile ifade edilir.

Çizelge Ek 1.2. Student-t dağılışına ilişkin kritik değerler (Tek yönlü)

Serbestlik Derecesi	Olasılık												
	0.450	0.400	0.350	.300	0.250	0.200	0.150	0.100	0.050	0.025	0.010	0.005	0.0005
1	0.158	0.325	0.510	0.727	1.000	1.375	1.963	3.078	6.314	12.705	31.821	63.65	636.19; .
2	0.142	0.289	0.445	0.617	0.816	1.061	1.380	1.880	2.920	4.303	6.955	9.925	31.5980;
3	0.137	0.277	0.424	0.584	0.765	0.978	1.250	1.638	2.353	3.182	4.541	5.841	12.924; .
4	0.134	0.271	0.414	0.569	0.741	0.941	1.190	1.533	2.132	2.776	3.747	4.004	8.610; .
5	0.132	0.267	0.408	0.559	0.727	0.920	1.156	1.476	2.015	2.571	3.385	4.032	6.869; .
6	0.131	0.265	0.404	0.553	0.718	0.906	1.134	1.440	2.447	2.571	3.143	3.707	5.959; .
7	0.130	0.263	0.402	0.549	0.711	0.895	1.119	1.415	1.895	2.365	2.998	3.499	5.408; .
8	0.130	0.262	0.399	0.546	0.706	0.889	1.108	1.397	1.860	2.306	2.896	3.355	5.041; .
9	0.129	0.261	0.398	0.543	0.703	0.883	1.100	1.388	1.833	2.262	2.821	3.250	4.781; .
10	0.129	0.260	0.397	0.542	0.700	0.879	1.093	1.372	1.812	2.228	2.764	3.169	4.587; .
11	0.129	0.260	0.396	0.540	0.697	0.875	1.088	1.363	1.796	2.201	2.718	3.106	4.437; .
12	0.128	0.259	0.395	0.539	0.695	0.872	1.083	1.356	1.782	2.179	2.681	3.055	4.318; .
13	0.128	0.259	0.394	0.538	0.694	0.870	1.079	1.350	1.771	2.160	2.650	3.012	4.221; .
14	0.128	0.258	0.393	0.537	0.692	0.868	1.076	1.345	1.761	2.145	2.624	2.977	4.140; .
15	0.128	0.258	0.393	0.536	0.691	0.866	1.074	1.341	1.753	2.131	2.602	2.947	4.073; .
16	0.128	0.258	0.392	0.535	0.690	0.865	1.071	1.337	1.746	2.120	2.588	2.921	4.015; .
17	0.128	0.257	0.392	0.534	0.689	0.863	1.069	1.333	1.740	2.110	2.567	2.898	3.965; .
18	0.127	0.257	0.392	0.534	0.688	0.862	1.067	1.330	1.734	2.101	2.552	2.878	3.923; .
19	0.127	0.257	0.391	0.533	0.688	0.861	1.066	1.328	1.729	2.093	2.539	2.861	3.882; .
20	0.127	0.257	0.391	0.533	0.687	0.860	1.064	1.325	1.725	2.086	2.528	2.845	3.850; .
21	0.127	0.257	0.391	0.532	0.686	0.859	1.063	1.323	1.721	2.080	2.518	2.831	3.819; .
22	0.127	0.256	0.390	0.532	0.685	0.858	1.061	1.321	1.717	2.074	2.508	2.819	3.792; .
23	0.127	0.256	0.390	0.532	0.685	0.858	1.060	1.319	1.714	2.069	2.500	2.807	3.767; .
24	0.127	0.256	0.390	0.531	0.685	0.857	1.059	1.318	1.711	2.064	2.492	2.797	3.745; .
25	0.127	0.256	0.390	0.531	0.684	0.856	1.058	1.315	1.708	2.060	2.485	2.787	3.725; .
26	0.127	0.258	0.390	0.531	0.684	0.856	1.058	1.315	1.706	2.056	2.479	2.779	3.707; .
27	0.127	0.258	0.389	0.531	0.684	0.855	1.057	1.314	1.703	2.052	2.473	2.771	3.690; .
28	0.127	0.258	0.389	0.530	0.683	0.855	1.056	1.313	1.701	2.048	2.467	2.763	3.674; .
29	0.127	0.258	0.389	0.530	0.683	0.854	1.055	1.311	1.699	2.045	2.462	2.758	3.659; .
30	0.127	0.258	0.389	0.530	0.683	0.854	1.055	1.310	1.697	2.042	2.457	2.750	3.648; .
40	0.126	0.255	0.388	0.529	0.681	0.851	1.050	1.303	1.684	2.021	2.423	2.704	3.551; .
60	0.126	0.254	0.387	0.527	0.679	0.848	1.045	1.296	1.671	2.000	2.390	2.660	3.460; .
120	0.126	0.254	0.386	0.525	0.677	0.845	1.041	1.289	1.658	1.980	2.358	2.617	3.372; .
240	0.126	0.254	0.386	0.524	0.674	0.842	1.036	1.282	1.645	1.960	2.326	2.576	3.291; .

$$R = \frac{\sum xy - \frac{(\sum x)(\sum y)}{n}}{\sqrt{\left[\sum x^2 - \frac{(\sum x)^2}{n} \right] \left[\sum y^2 - \frac{(\sum y)^2}{n} \right]}} \quad (E1.13.)$$

Böyle bir fonksiyonun grafiği Şekil 2.2 de şematik olarak gösterilmiş olup tüm noktalar fonksiyonla tam uyumlu ise yani fonksiyon çizgisinin tam üzerinde bulunuyorsa korelasyon katsayısı 1 dir.



Şekil Ek 1.2. Doğrusal modelli bir regresyon korelasyon eğrisi

Kimyasal denemelerde özellikle aygıtlı analizde farklı derişim aralıklarında standartlara ve örneklere ilişkin verilerin değerdendirilebilmesi için kalibrasyon grafiğinin oluşturulması regresyon korelasyon analizi ile yapılır. Gerek fen gerekse sosyal bilimlerde bilimsel araştırmalarda elde edilen verilerin çözümlenmesinde ve bulguların yorumlanmasında çeşitli istatistik yöntemler yoğun olarak kullanılmaktadır. Özellikle 20. yüzyılın son çeyreğinden itibaren bir yandan istatistik yöntem ve tekniklerinin her alanda yaygınlaşmasına, diğer yandan bu alanda çalışma yapanların önceliklerinin değışmesine neden olan iki önemli gelişme yaşanmıştır (Korum, 1991). Bunlardan birincisi bilgisayar teknolojisinde meydana gelen ilerlemeler ve ikinci ise bilginin üretilmesinde ve iletiminde bilgisayar ve iletişim teknolojilerinin sağladıkları olanaklardır.

Ek 2. Veri Madenciliği

Veri madenciliği son yıllarda kullanımının yaygınlaşmasının nedenleri aşağıdaki kategorilerde toplanabilir (Fayyad vd., 1996; Cao, 1998):

- Veri Artışı: Teknolojinin ve bilimin ilerlemesiyle gereksinim duyulan ve saklanan veri boyutu her geçen gün artmaktadır. Veri boyutunun arttığı çeşitli alanlar şöyledir:

- Görüntü, sıcaklık, uzaklık gibi bilgilerin elde edildiği sensör veritabanları,
- Barkod bilgileri,
- Hükümet kayıtları,
- Biyoloji, tıp ve genetik veritabanları,
- Uzaysal araştırma veritabanları,
- Finansal veri tabanları,
- Coğrafi bilgi sistemleri,
- Meteoroloji ve atmosfer veritabanları,
- Sosyal bilimler ve davranış bilimleri,
- Web mining (İnternet ve adres bilgileri),
- E-ticaret,
- Metin madenciliği,
- Çoklu ortam madenciliği,
- Veri görselleştirme,
- İstatistik,
- Veri tabanı yönetimi.

Veri tabanlarının büyümesiyle veri madenciliğine duyulan gereksinim de artmaktadır. Verilerin, veri madenciliği yapılmadan sadece saklanması bilimsel yorum getirme, geleceğe yönelik tahminlerde bulunma, sınıflandırma gibi konularda yeterli olmamaktadır.

- Veri Ambarları: Çeşitli veri depolarındaki verilerin ortak bir formatta ve belli anahtar ve alan tanımlamalarıyla bir araya getirildiği veri ambarlarının kullanımı günden güne artması veri madenciliği uygulamalarını daha cazip hale getirmektedir.

- Bilgisayar Donanımlarındaki Gelişmeler: Veri madenciliği algoritmaları, büyük miktardaki verilerin taşınmasını gerektirmektedir. Ayrıca başarılı bir veri işleme için birçok hesaplama yapılması gerekebilir. Bilgisayar yazılım ve donanımlarının (bellek, sabit disk, işlemci gücü ve I/O bant genişliği, vb..) maliyetlerindeki hızlı düşüş mevcut veri madenciliği metotlarının daha kolay gerçekleşmesini ve yeni veri madenciliği metotlarının geliştirilmesini sağlamaktadır.

- Pazarlama Alanındaki Rekabet: İş dünyasındaki gelişmeler pazarlama alanında artan bir rekabet oluşturmuştur. Veri madenciliği, bu rekabette başarılı olabilmek için

önemli bir silahtır. Bilgi zengini şirketler, veri tabanları üzerinde veri madenciliği uygulayıp piyasa hareketlerini görerek stratejilerini belirlemektedir. Böylece rekabette öne geçebilmektedirler

- Veri Madenciliği Programları: Teknolojinin ilerlemesi ve veri madenciliğinin gelişmesiyle veri madenciliği uygulama programlarının da ticari piyasası oluşmuştur. Böylece veri madenciliği toplum tarafından kabul edilirlilik kazanmıştır.

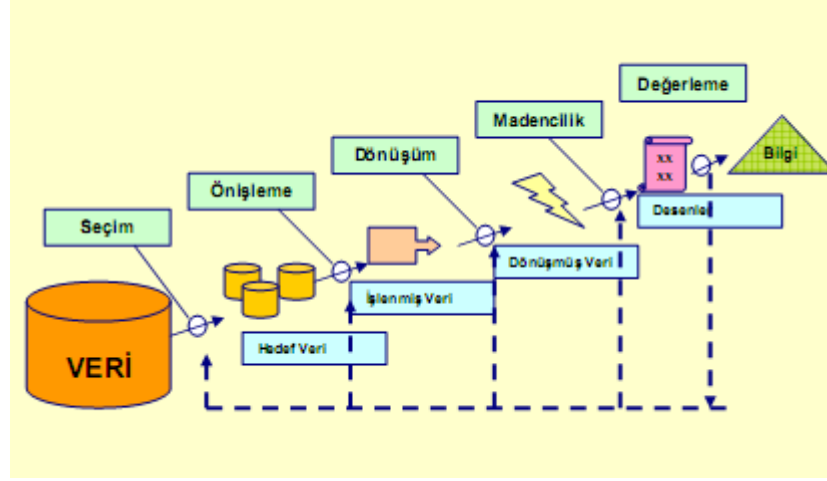
Etkili bir veri madenciliği işleminin gerçekleştirilmesi için uygulanan bilgi keşfi sisteminin özellikleri ve veri madenciliği teknikleri geliştirilirken karşılaşılabilecek güçlüklerin incelenmesi gerekmektedir (Kohavi, Provost, 2001; Han, Kamber, Tung, 2001). Bunlar:

- Farklı veri setlerinin kullanımı:
- Veri madenciliği algoritmasının etkililik ve ölçeklendiribilirliği:
- Kullanışlılık:
- Çoklu ayırma (abstraction) seviyelerinde etkileşimli madencilik bilgisi:
- Farklı veri kaynaklarının kullanımı:
- Gizlilik ve veri güvenliğinin korunması:

Veri madenciliğinin esas amaçlarına gelince, Berry ve Linoff veri madenciliğinin amaçlarını aşağıdaki gibi özetlemişlerdir (Berry ve Linoff , 2004) :

- Sınıflandırma (Classification)
- Tahmin (Estimation)
- Öntahmin (Prediction)
- Benzer gruplama (Affinity grouping)
- Kümeleme (Clustering)
- Tanımlama (Description)

Veri madenciliği süreci genellikle uygulamadaki iki temel soruna cevap bulmayı amaçlamaktadır. Bunlardan birincisi mevcut veritabanından tahminler yapmak (tahmin edici model), ikincisi ise veriden elde edilecek davranışları betimlemektir (tanımlayıcı bilgi) (Moshkovich vd., 2002). Tahmin edici modeller, verideki eğilimleri, örüntüleri ve gizli ilişkileri bulmaya yardımcı olurken; tanımlayıcı bilgi ilginç örüntü ve ilişkileri bulmak için veriyi, görselleştirme ve keşfetmeye odaklanmaktadır (Melody ve Ajit, 2001). Veri madenciliğin gerçekleştirilmesi ise, çeşitli karar verme işlemlerinin etkileşimli ve tekrarlayan aşamalar sonucunda olmaktadır. Bu aşamalar, Brachman ve Anand tarafından Şekil Ek 2.1'deki gibi gösterilmektedir (Fayyad vd., 1996);



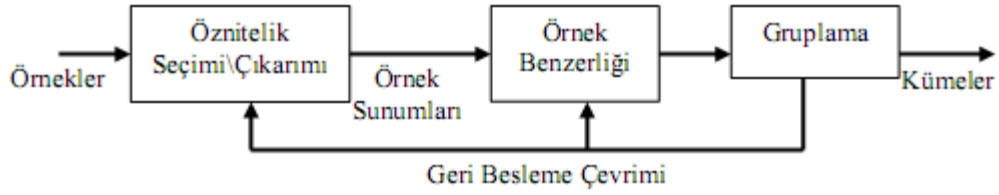
Şekil Ek 2.1. Veri madenciliği aşamaları

Ek 3. Kümeleme Analizi Ve Teknikleri

Ek 1. Kümeleme Analizi: Klasik bir örnek kümeleme işleminde takip edilmesi gereken adımlar şunlardır (Brereton, 2003; Arifoğlu, 2005; Özkan, 2008; Demir vd., 2009; Akpolat, 2010);

- 1) Örneklerin sunulması,
- 2) Örneklerin uzaklık ölçümlerinin veri tabanına uygun olarak tanımlanması,
- 3) Kümeleme veya gruptama,
- 4) Veri ayıklama (gerekli olduğu durumlarda yapılmaktadır),
- 5) Çıktının değerlendirilmesi (gerekli olduğu durumlarda yapılmaktadır).

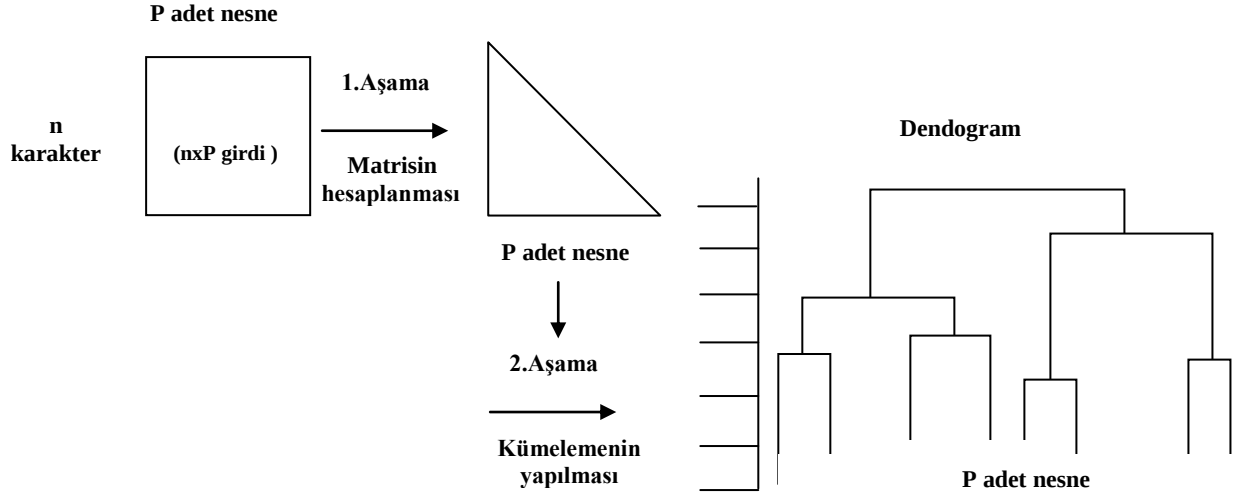
Yukarıda bulunan ilk üç adım Şekil Ek 3.1’de görülmektedir. Geri besleme, kümeleme sonucunda elde edilen çıktının, örnek uzaklık ölçümlerine ve özniteliklerin çıkarılmasına etki etmektedir.



Şekil Ek 3.1. Kümeleme İşleminin Adımları

Birbirine benzer nesneleri aynı kümeye, benzemeyenleri de farklı kümelere yerleştiren bu yöntemde, önce noktaların yani nesnelerin birbirine uzaklıkları hesaplanır ve bu uzaklıklara dayanarak kümeler genellikle hiyerarşik bir düzende oluşturulur. Kümeleme analizinin bütünleyici algoritmasına göre kümeleme işleminin aşamaları aşağıdaki şekilde (Şekil Ek 3.2) görülmektedir.

2. Kümeleme Teknikleri: Kümeleme teknikleri; uzaklık matrisini kullanarak nesneleri veya değişkenleri kendi içinde homojen ve kendi aralarında heterojen gruplar oluşturmaya imkan sağlamaktadır. Kümeleme analizi için birçok algoritma öne sürülmüştür. Ancak literatür de bu algoritmalar iki başlık altında toplanmıştır. Bunlardan ilki, dendogram oluşturan hiyerarşik kümeleme teknikleri, diğeri ise hiyerarşik olmayan kümeleme tekniği olarak adlandırılmaktadır. Her iki teknikte de ortak amaç kümeler arasındaki farklılıkları ve kümeler içi benzerlikleri en yüksek düzeye çıkarmaktır. Yani, küme içi homojenlik artırılırken kümeler arası homojenlik ise azaltılmaktadır. Hangi tekniğin kullanılacağı küme sayısına bağlı olmakla birlikte her iki tekniğin beraber kullanılması çok daha yararlıdır.

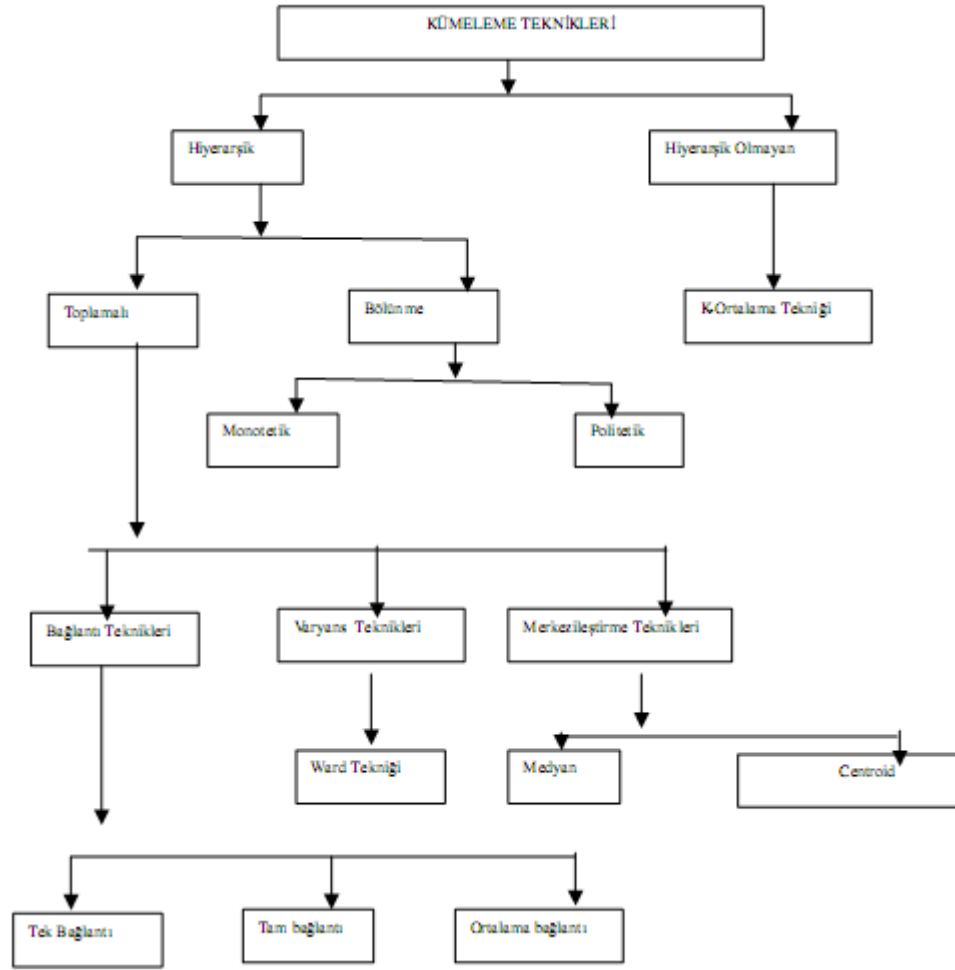


Şekil Ek 3.2. Bütünleyici algoritmaya göre kümeleme işlemi

Böylece hem sonuçları hem de iki tekniğin hangisinin daha uygun sonuçlar verdiğini karşılaştırmak mümkün olmaktadır. Kümeleme tekniklerinin genel yapısı Şekil Ek 3.3’de verilmiştir (Grabmeier ve Rudolph, 2002).

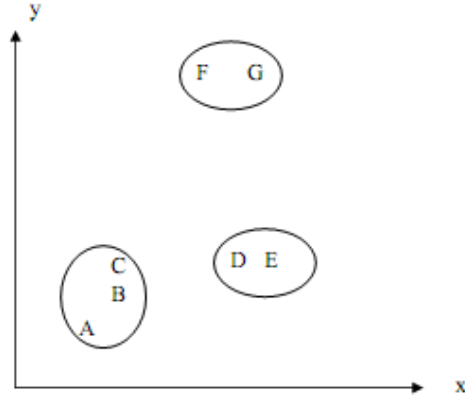
Ek 2.1. Hiyerarşik Kümeleme

Hiyerarşik kümeleme, dendrogram olarak bilinen kümelerin ağaç yapısını veya diğer bir deyişle bir küme hiyerarşisi ortaya çıkarır. Her küme düğümü, ortak ebeveynlere sahip noktaların çocuk ve kardeş küme bölünmelerini içerir. Böyle bir yaklaşım, farklı seviyelerdeki düğümlerin dikkatle incelenmesine olanak sağlar. Hiyerarşik kümeleme metotları toplanmış (agglomerative) ve bölen (divisive) metotları olarak sınıflandırılabilir (Jain vd, 1988, Kaufman vd, 1990). Toplanmış kümeleme tek noktadan (singleton) oluşan kümelerden başlar ve bu kümeleri tekrarlı olarak en uygun kümeleri Şekil-E3.4’de olduğu gibi daha fazla kümede birleştirir. Bölen kümeleme ise bütün veri noktalarını içeren tek bir kümeden başlar ve bu kümeleri en uygun biçimde tekrarlı olarak parçalara ayırır. Bu süreç bitirme kriteri sağlanana kadar devam eder.

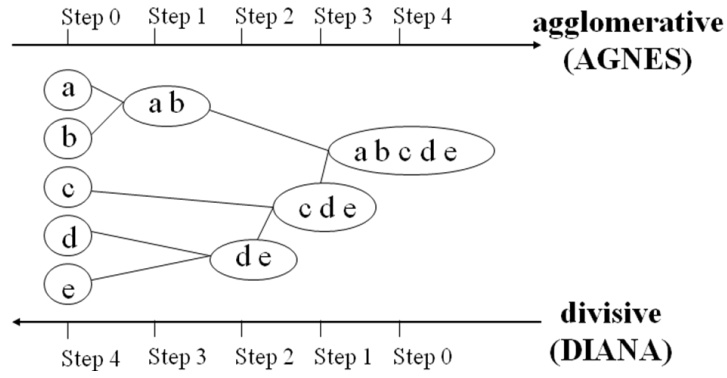


Şekil Ek 3.3. Kümeleme Tekniklerinin Genel Sınıflandırılması

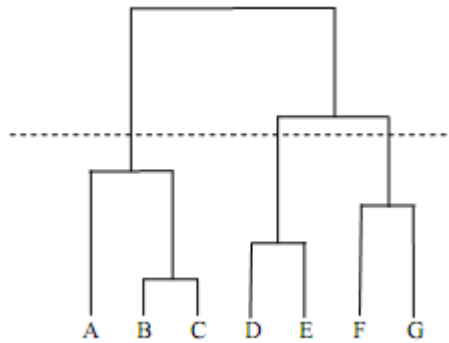
Şekil Ek 3.4'deki birleştirme ağacından da görüleceği üzere ilk adımda B-C, ikinci adımda D-E, üçüncü adımda F-G, dördüncü adımda A-BC birleştirilmiştir. Sonraki adımda DE-FG ile ve en son adımda da ABC-DEFG birleştirilerek bütün kayıtları içeren en büyük küme oluşturulmuştur. Adını bu hiyerarşik birleştirme ağacından alan hiyerarşik yöntemlerin sonlanması, daha önceden belirlenen küme sayısına bağlıdır. Örneğin bu birleştirme ağacında 3 küme oluşacak şekilde bir kümeleme gerçekleştirilmiştir.



Şekil Ek 3.4.a. Hiyerarşik yapıda kümeleri iki boyutlu düzlemde gösterme



Şekil Ek 3.4.b. Hiyerarşik yapıda kümeleri birleştirme ya da ayrıştırma



Şekil Ek 3.4.c. Hiyerarşik yapıda kümeleri birleştirme ağacı (dengdongram)

Hiyerarşik yöntemlerin en güzel yanlarından bir tanesi de bir kez bu birleştirme ağacının oluşturulmasından sonra ağacın istenilen seviyelerden (istenilen sayıda küme sağlayacak şekilde) ayrılabilmesidir (Jain ve Dubes, 1988). Çoğu hiyerarşik

kümeleme yöntemi, tek-bağ (Sneath ve Sokal, 1973), tüm-bağ (King, 1967) ve en küçük-varyans (Ward, 1963) algoritmalarının değişik türevleridir. Bunlardan tek-bağ ve tüm-bağ algoritmaları en yaygın kullanılan olanlarıdır. Bu iki algoritma arasındaki fark, birleştirme işleminde iki küme için hesapladığı benzerlik (uzaklık) çıkarma yönteminden ileri gelmektedir. En basit uzaklık tanımı Öklidyen Uzaklık (Eucliden Distance) olarak tanımlanmıştır.

Ek 2.2. Verilerin Analize Hazırlanması

Kemometrik analizlerde ham veri büyük öneme sahiptir. Yapılan bir deneysel çalışma sonrasında elde edilen verilerin kemometrik hesaplamalarına geçmeden önce verilerin özellikleri dikkatle incelenmelidir. Elde edilen veriler arasında birbirinden çok farklı değerler bulunuyorsa ya da değerler arası büyük dengesizlikler varsa, verilerin benzer birimlere çevrilmeleri veya ortalama değer etrafında toplanması gerekebilir. Bunun gibi durumlarda verilere merkezileştirme, standartlaştırma ve normalizasyon gibi verilerin analize hazırlanmasına yönelik işlemler uygulanır. Bunun için öncelikle hesaplamalarda gerekli olan veri ve fark matrisinin tanımlanması aşağıdaki gibi yapılmalıdır (Kaya ve Köymen, 2008). Veri matrisinde n veri sayısı, p nitelik sayısı ve fark matrisinde de d(i,j) iki veri arasındaki uzaklık olmak üzere;

$$\begin{bmatrix} x_{11} & \dots & x_{1f} & \dots & x_{1p} \\ \dots & \dots & \dots & \dots & \dots \\ x_{i1} & \dots & x_{if} & \dots & x_{ip} \\ \dots & \dots & \dots & \dots & \dots \\ x_{n1} & \dots & x_{nf} & \dots & x_{np} \end{bmatrix} \quad \text{Veri Matrisi } i=1 \dots n, f=1 \dots p$$

$$\begin{bmatrix} 0 & & & & & & \\ d(2,1) & 0 & & & & & \\ d(3,1) & d(3,2) & 0 & & & & \\ . & . & . & . & & & \\ . & . & . & . & . & & \\ . & . & . & . & . & . & \\ d(n,1) & d(n,2) & \dots & \dots & \dots & \dots & 0 \end{bmatrix} \quad \text{Fark Matrisi}$$

Ham verilerin sayısal ölçek değerleri birbirinden çok farklı ise yukarıda oluşturulan matrisler aşağıdaki yaklaştırma metotlarından biriyle yeniden hesaplanarak oluşturulur. Bunlar kısaca;

- **Merkezileştirme**=Her bir kolonun ortalamasının kolon elemanlarından çıkarılmasıyla matris verisinin merkezileştirme işlemi yapılmış olmaktadır.

$$x_{ij}^m = x_{ij} - x_j$$

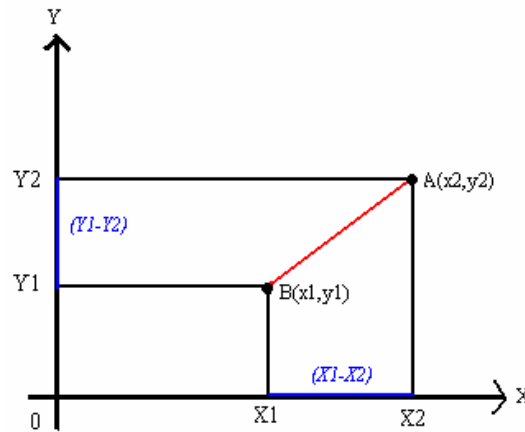
- **Standartlaştırma**=Öncelikle her bir kolonun ortalama ve standart sapması hesaplanır. Daha sonra kolondaki her bir elemandan kolon ortalaması çıkarılır ve kolon standart sapmasına bölünür. Böylece matris verisi standartlaştırılmış olur.

$$x_{ij}^s = \frac{x_{ij} - \bar{x}_j}{\sqrt{\frac{\sum_{i=1}^I (x_{ij} - \bar{x}_j)^2}{N - 1}}}$$

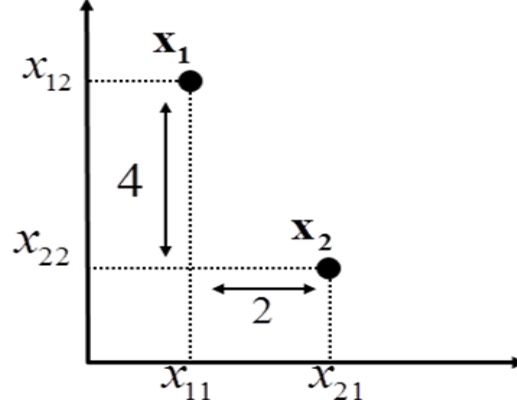
- **Normalizasyon**=Matris üzerinde önce satır elemanları toplanır ve toplam değerler her satır için ayrı ayrı hesaplanır. Sonra her bir satır toplamı her satırın kendi elemanlarına bölünür ve normalizasyon yapılmış olur.

$$x_{ij}^N = \frac{x_{ij}}{\sum_{i=1}^I x_i}$$

Değişkenler arası benzerlik ya da farklılıklara dayanır. Yani kümeleme analizinde bir birime ait değişkenlerin birbirleriyle olan uzaklıkları hesaplanır. Kısaca uzaklık matrisi ya da benzerlik matrisinden yararlanıyoruz. Koordinat sisteminde iki nokta arasındaki uzaklığın gösterimi ise Şekil Ek 3.5’de verilmiştir. Fark matrisinin oluşturulması sırasında noktalar arasındaki uzaklık (ya da yakınlık) hesaplaması için Öklit, Manhattan ve Supremum gibi farklı tanımlar yapılmış olup, bunlardan en basiti olan Şekil E3.6’da gösterilen Öklit fark tanımı aşağıdaki gibi yazılmıştır. Bu uzaklık ölçüsü değişkenlerin aynı ağırlıkla ölçeklenmemiş olması durumunda kullanılır.



Şekil Ek 3.5. Koordinat sisteminde iki nokta arasındaki uzaklık



Şekil Ek 3.6. Öklit uzaklık ölçütü

$$d^*(x_i, x_j) = \left[\sum_{k=1}^p w_k^2 (x_{ik} - x_{jk})^2 \right]^{1/2}$$

Formülde w_k , k. değişkenin standart sapma değerinin (s_k) ya da değişim aralığı değerinin tersidir. w_k 'nın s_k değerinin tersi olması durumunda elde edilen uzaklığa Karl-Pearson uzaklığı da denilmektedir. $d_{\text{Öklit}}(x_1, x_2) = \sqrt{4^2 + 2^2} = 4.472$ olarak hesaplanmıştır. Kümeler arası uzaklık da farklı yöntemlerle hesaplanmaktadır. Bunlardan en basiti Ortalama Link (Average Linkage) aşağıdaki gibi yazılmıştır.

$$d_{ki} = \frac{d_{Ai} + d_{Bi}}{2}$$

2.3. Örnek Çözümleme

Kalsiyum ve fosfat içeriğine göre 6 kan örneği arasında benzerlik ilişkisinin incelendiği örneğin sayısal çözümlemesi aşağıda verilmiştir.

Örnek	Kalsiyum (mg/100ml)	Fosfat (mg/100ml)
1	8.0	5.5
2.	8.25	5.75
3	8.7	6.3
4	10.0	3.0
5	10.25	4.0
6	9.75	3.5

$$d(i,k) = \left(\sum_{j=1}^d |x_{ij} - x_{kj}|^2 \right)^{1/2} = \sqrt{(\mathbf{x}_i - \mathbf{x}_k)^T (\mathbf{x}_i - \mathbf{x}_k)}$$

$$d_{12} = \left[(8-8.25)^2 + (5.5-5.75)^2 \right]^{1/2} = 0.354$$

$$d_{14} = \left[(8-10)^2 + (5.5-3)^2 \right]^{1/2} = 3.201$$

Örnek	1	2	3	4	5	6
1	0					
2	<u>0.354</u>	0				
3	1.063	0.711	0			
4	3.201	3.260	3.347	0		
5	2.704	2.658	2.774	1.031	0	
6	2.658	2.704	2.990	0.559	0.707	0

$$d_{5*4} = \frac{d_{54} + d_{56}}{2} = \frac{1.031 + 0.707}{2} = 0.869$$

$$d_{4*3} = \frac{d_{43} + d_{63}}{2} = \frac{3.347 + 2.990}{2} = 3.269$$

1.İndirgenmiş matris

Örnek	1*	3	4	5	6
1*	0				
3	0.887	0			
4	3.231	3.347	0		
5	2.681	2.774	1.031	0	
6	2.681	2.990	<u>0.559</u>	0.707	0

2.İndirgenmiş matris

Örnek	1*	3	4*	5
1*	0			
3	0.887	0		
4*	2.956	3.269	0	
5	2.681	2.774	0.869	0

3.İndirgenmiş matris

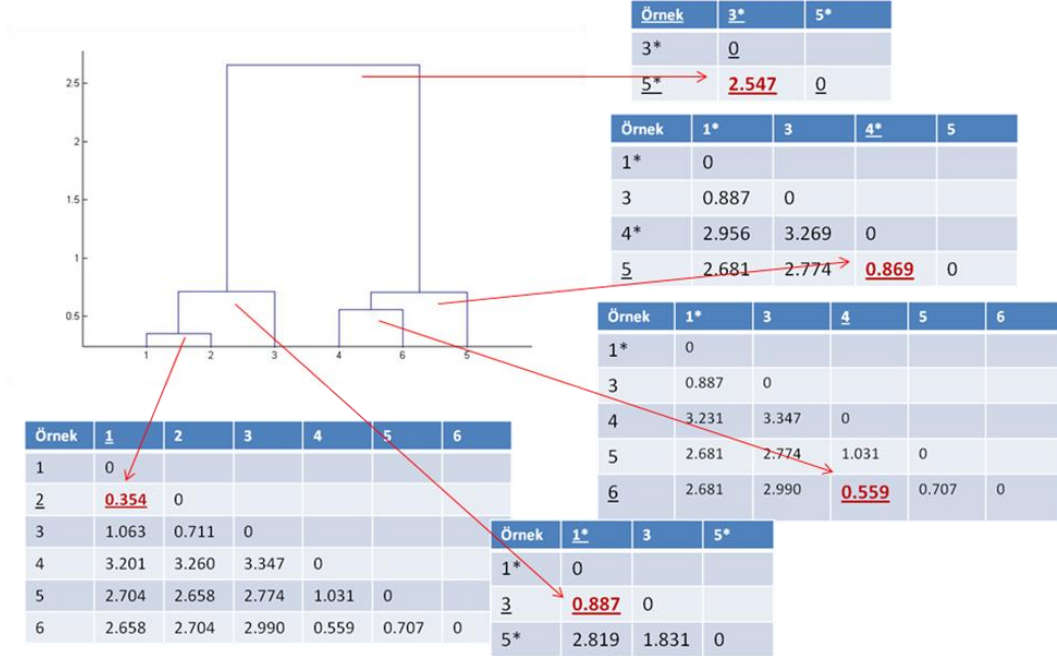
Örnek	1*	3	5*
1*	0		
3	<u>0.887</u>	0	
5*	2.819	0.831	0

4.İndirgenmiş matris

Örnek	3*	5*
3*	0	
5*	<u>2.547</u>	0

İndirgenmiş matris hesaplamaları sonucu çizilen dengdonram aşağıda ayrıntılı olarak gösterilmiştir.

Hesaplamalar sonucu oluşan dendrogram



Ek 2.4. Kümeleme Analiz Yönteminin Belirlenmesi

Kümeleme yöntemleri, birim ya da değişkenleri uygun gruplara ayırırken grupları belirlemede izledikleri yaklaşımlara göre;

1. Aşamalı Kümeleme Yöntemleri (Hierarchical Cluster Analysis Methods)
2. Aşamalı Olmayan Kümeleme Yöntemleri (Nonhierarchical Cluster Analysis Methods)

biçiminde iki temel gruba ayrılmaktadır.

Aşamalı kümeleme yöntemlerinde birimlerin birbirleri ile birleştirilmesinde değişik yaklaşımlar uygulanmaktadır. Bu yöntemlerden sıklıkla kullanılan ve bilinenler;

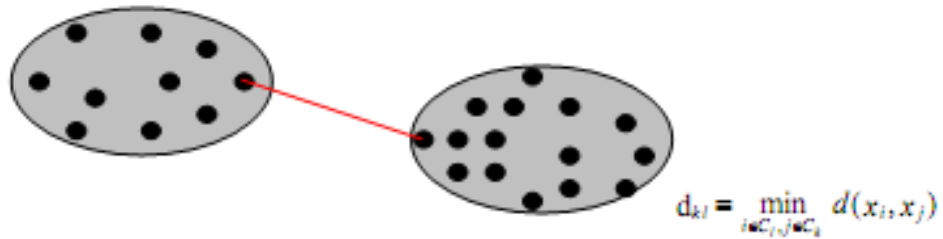
1. Tek Bağlantı Kümeleme Yöntemi,
2. Ortalama Bağlantı Kümeleme Yöntemi,
3. Tam Bağlantı Kümeleme Yöntemi,
4. Mc Quitty Bağlantı Kümeleme Yöntemi,
5. Küresel Ortalama Bağlantı Kümeleme Yöntemi,
6. Ortanca Bağlantı Kümeleme Yöntemi,
7. Ward Bağlantı Kümeleme Yöntemi'dir

Aşamalı olmayan kümeleme yöntemleri arasında en çok kullanılan yöntemler;

1. k-Ortalama Yöntemi,
2. En çok Olabilirlik Yöntemi'dir.

Tek bağ kümeleme En Yakın Komşuluk (Nearest Neighbor-Single Linkage-SLINK)

Aşamalı Kümeleme Yöntemi olarak *Öklit uzaklığının* hesaplanmasını kullanan Nearest Neighbor (En Yakın Komşu) yöntemi en basit yöntem olarak tek-bağ algoritmasında iki küme arasındaki uzaklık, her iki küme arasında yer alan kayıtlardan birbirlerine en yakın olanların uzaklığı olarak değerlendirilmektedir. Bu işlem Şekil E3.7'de ayrı ayrı gösterilmiştir:



Şekil Ek 3.7. Tek-Bağ kümelemede iki küme arası uzaklık

Ek 4 ve Ek 5'de ayrıntılı olarak tanıtılan MATLAB ve SPSS programının uygulamasına yönelik olarak aşağıda iki çözümleme sunulmuştur.

Ek 3. Kümeleme analizi için örnek MATLAB uygulaması

Önceki bölümde sunulan örnek çözümlemenin MATLAB algoritması ve program çıktıları aşağıda verilmiştir.

% Hiyerarsik Kumeleme Metotlari Algoritma Dosyasi

clc , clear , clf

'% Veri matrisi '

'% Olcum(mg/100ml) Kalsiyum=101 Fosfat=102'

```
Orn_Kal_Fos=[0 101 102; ...  
             1 8 5.5 ; ...  
             2 8.25 5.75; ...  
             3 8.7 6.3 ; ...  
             4 10 3 ; ...
```

```
5 10.25 4 ; ...  
6 9.75 3.5];
```

```
'% Bileşenlerin Yeniden Duzenlenmesi'
```

```
'% Veri matrisi '
```

```
X=Orn_Kal_Fos(2:7,2:3)
```

```
'% x veri matrisi için hesaplanan oklid uzakliklari'
```

```
Y = pdist(X)
```

```
'% y öklid uzaklık vektörünün karesel formu'
```

```
squareform(Y)
```

```
'% Hesaplanan gruplararası öklid uzaklıkları'
```

```
Z = linkage(Y)
```

```
'% Z matrisi için dendogramın çizilmesi'
```

```
dendrogram(Z) , xlabel('Birim'), ylabel('Z')
```

```
% Hiyerarsik Kumeleme Metotlari CommandWindowi Cozum Arayuzey
```

```
ans =
```

```
% Veri matrisi
```

```
% Olcum(mg/100ml) Kalsiyum=101 Fosfat=102
```

```
% Bileşenlerin Yeniden Duzenlenmesi
```

```
% Veri matrisi
```

```
X =
```

```
8.0000 5.5000
```

```
8.2500 5.7500
```

```
8.7000 6.3000
```

```
10.0000 3.0000
```

```
10.2500 4.0000
```

9.7500 3.5000

ans =

% x veri matrisi için hesaplanan oklid uzakliklari

Y =

Columns 1 through 8

0.3536 1.0630 3.2016 2.7042 2.6575 0.7106 3.2596 2.6575

Columns 9 through 15

2.7042 3.5468 2.7735 2.9904 1.0308 0.5590 0.7071

% y öklid uzaklık vektörünün karesel formu

0 0.3536 1.0630 3.2016 2.7042 2.6575

0.3536 0 0.7106 3.2596 2.6575 2.7042

1.0630 0.7106 0 3.5468 2.7735 2.9904

3.2016 3.2596 3.5468 0 1.0308 0.5590

2.7042 2.6575 2.7735 1.0308 0 0.7071

2.6575 2.7042 2.9904 0.5590 0.7071 0

ans =

% Hesaplanan gruplararası öklid uzaklıkları

Z =

1.0000 2.0000 0.3536

4.0000 6.0000 0.5590

5.0000 8.0000 0.7071

3.0000 7.0000 0.7106

9.0000 10.0000 2.6575

ans =

% Z matrisi için dendogramın çizilmesi

>>

Ek 4. Kümeleme analizi için SPSS menü seçenekleri

Bu bölümde ise SPSS çözümlemesi için programın menü seçenekleri aşağıda verilmiştir.

Kazi.sav [DataSet1] - SPSS Data Editor

	Name	Type	Width	Decimals	Label	Values	Missing	Columns	Align	Measure
1	isim	String	8	0		None	None	6	Left	Nominal
2	ornek	Numeric	8	1		None	None	8	Right	Scale
3	Ti	Numeric	8	3		None	None	8	Right	Scale
4	Sr	Numeric	8	0		None	None	8	Right	Scale
5	Ba	Numeric	8	0		None	None	8	Right	Scale
6	Mn	Numeric	8	0		None	None	8	Right	Scale
7	Cr	Numeric	8	0		None	None	8	Right	Scale
8	Ca	Numeric	8	3		None	None	8	Right	Scale
9	Al	Numeric	8	3		None	None	8	Right	Scale
10	Fe	Numeric	8	3		None	None	8	Right	Scale
11	Mg	Numeric	8	3		None	None	8	Right	Scale
12	Na	Numeric	8	3		None	None	8	Right	Scale
13	K	Numeric	8	3		None	None	8	Right	Scale
14	sinif	String	8	0		None	None	8	Left	Nominal
15										

SPSS Processor is ready

Kazi.sav [DataSet1] - SPSS Data Editor

Analyze > Classify > Hierarchical Cluster...

	isim	ornek		in	Cr	Ca	Al	Fe	Mg
1	A1	,1		642	60	1,640	8,342	3,542	,458
2	A2	,2		792	64	2,017	8,592	3,696	,509
3	A3	,3		588	48	1,587	7,886	3,221	,540
4	A4	,4		526	62	1,032	8,547	3,455	,546
5	A5	,5		775	184	1,334	11,229	4,637	,395
6	E1	,6		1377	90	1,370	10,334	4,543	,408
7	E2	,7		820	86	1,206	9,537	4,099	,427
8	E3	,8					10,139	4,490	,502
9	E4	,9					10,501	4,641	,548
10	L1	1,0					9,802	4,280	,738
11	L2	1,1					8,960	3,828	,535
12	L3	1,2					8,163	3,265	,521
13	L4	1,3					8,465	3,437	,572
14	L5	1,4							
15	C1	1,5		415	193	1,226	7,207	3,102	,539
16	C2	1,6		1308	65	,907	8,401	3,743	,784
17	C3	1,7		921	69	1,164	7,926	3,431	,636
18	C4	1,8	,252	1674	63	,927	7,980	3,748	,549
19	C5	1,9	,261	750	53	1,171	8,070	3,536	,599
20	G8	2,0	,397	849	61	1,311	7,819	3,770	,668
				939	61	1,260	8,694	4,146	,656

SPSS Analyze Classify Hierarchical Cluster

Hierarchical Cluster Analysis

Variable(s):

Label Cases by:

Cluster: ☒ Cases ☐ Variables

Display: ☒ Statistics ☒ Plots

Statistics... Plots... Method... Save...

Hierarchical Cluster Analysis

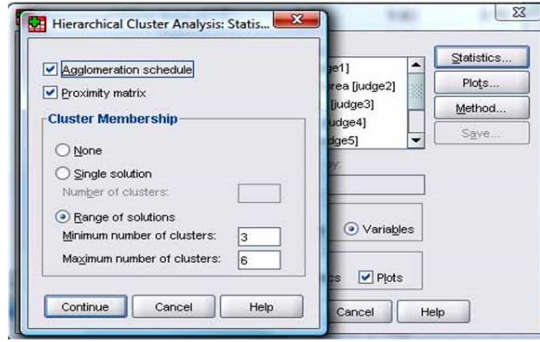
Variable(s):

Label Cases by:

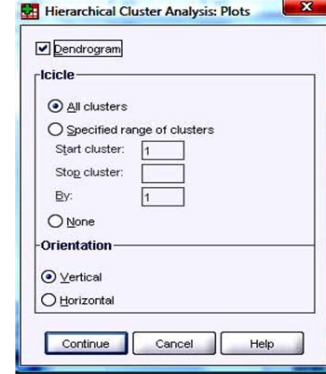
Cluster: ☒ Cases ☐ Variables

Display: ☒ Statistics ☒ Plots

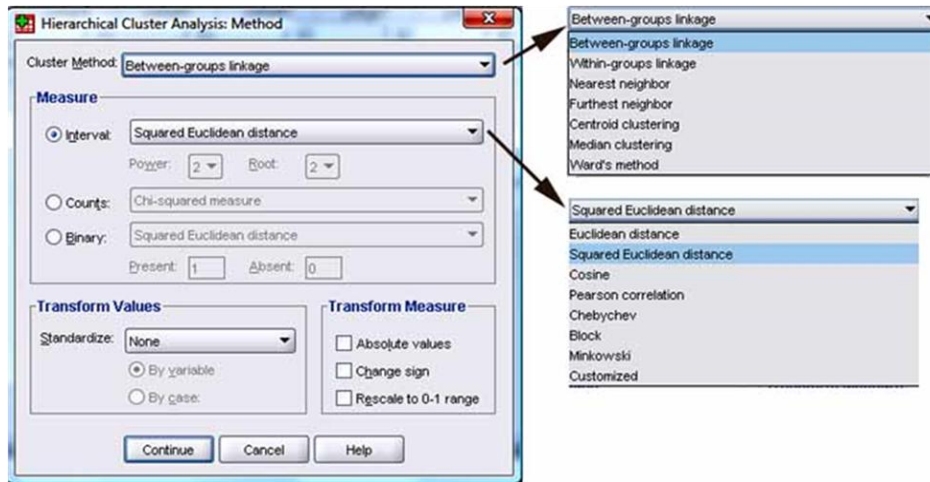
Statistics... Plots... Method... Save...



Plots

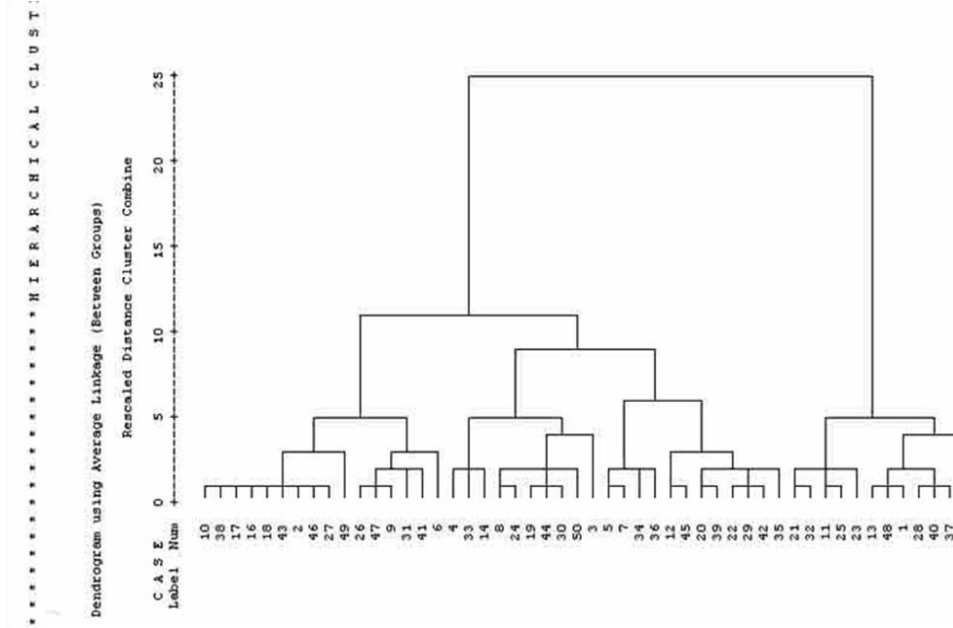


Method



Şekil Ek 3.8. Kümeleme analizi için SPSS menü seçenekleri

Dendrogram



Şekil Ek 3.9. Kümeleme analizi sonucu SPSS te oluşturulan dendrogram

Ek 4. MATLAB İle Programlama

MATLAB paket programı kullanıcıya program parçacıklarını yazabilmesi ve düzenleyebilmesi için tıpkı Windows-Notepad ya da Ofis-Word editörlerine benzer olarak M-File adıyla dosya sayfaları oluşturma olanağı tanımaktadır. M-File editörünü daha iyi tanıyabilmek için kitabın bundan sonraki kısımlarında açıklamaların büyük bir kısmı Word sayfası yerine bu editörde yazılmış olarak sunulacaktır (Akpolat, 2010).

% MATLAB m-file

% -----

% MATLAB Programı

% MATrix LABoratory sözcüklerinin kısaltılmasıyla adlandırılmış olan MATLAB programlama dilinin 2002 Ağustos ayı itibarıyla yayınlanan MATLAB 7.0 versiyonu C dilinde yazılmış olup sayısal analiz, matris ve dizi işlemleri, sinyal işleme, algoritma geliştirme, C, C++, Java ve Internet ile programlama ve grafiksel kullanıcı ara yüzü (Graphical User Interface – GUI) formu program yazma gibi sayısal işlemleri, kolay kullanımlı bir grafik ara yüzü üzerinden, diğer programlama dillerindeki geleneksel kodlamaya karşı matematiksel denklem yazma kolaylığını da sağlamaktadır. Kullanım kolaylığı, işletim sistemi uyumluluğu, hazır fonksiyonları, donanımın kısıtlamasız görüntüleme özellikleri, grafiksel kullanıcı ara yüzü ile teknik veri işleme kolaylıkları gibi üstünlükleri olan bu programın bir alt toolbox'ı olan simulink ise, dinamik sistem modellerin kurulması, benzeşimi ve çözümlemesinde kullanılan sanal bir laboratuvar seti gibi düşünülebilir.

% MATLAB Sistemi

% Altı ana kısma ayrılabilir.

Geliştirme ortamı (Arayüz)

Matematiksel fonksiyon kütüphanesi

Matlab dili Grafik İdaresi,(Handle Graphics)

Uygulamalı program arayüzü (API)

MATLAB derleyicisi (Compiler)

% Aşağıda kurulumdan sonra Bilgisayar ekranında karşılaşılabilecek MATLAB 7.0' in bazı ara yüzeyleri görülmektedir.

% FigEk 4.1: MATLAB Arayüzü

% FigEk 4.2: M-File Arayüzü

% FigEk 4.3: Figure Arayüzü

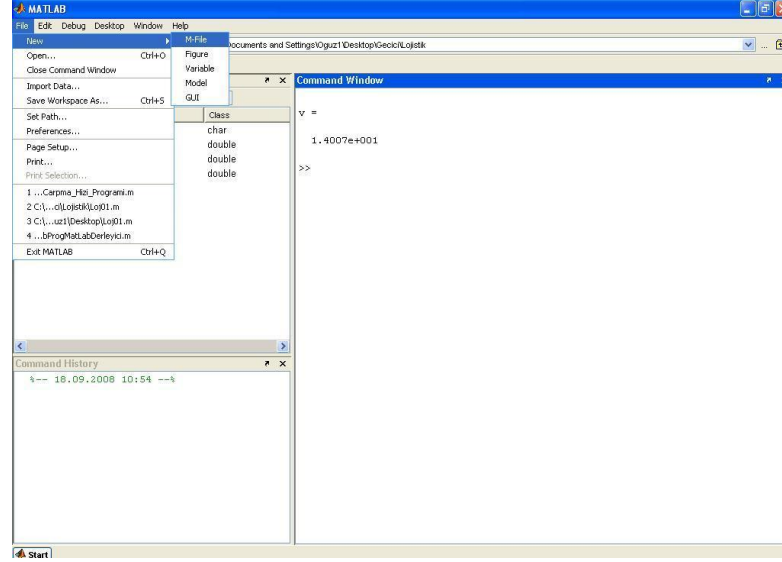
% FigEk 4.4: Help Arayüzü

% -----

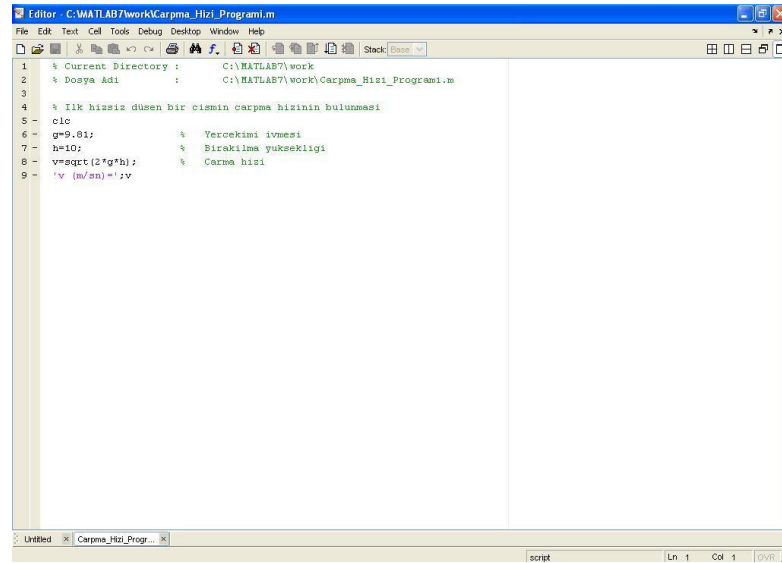
MATLAB Dilinin Temel Özellikleri

Aşağıda MATLAB programlama dilinin temel özellikleri kısaca özetlenmiştir.

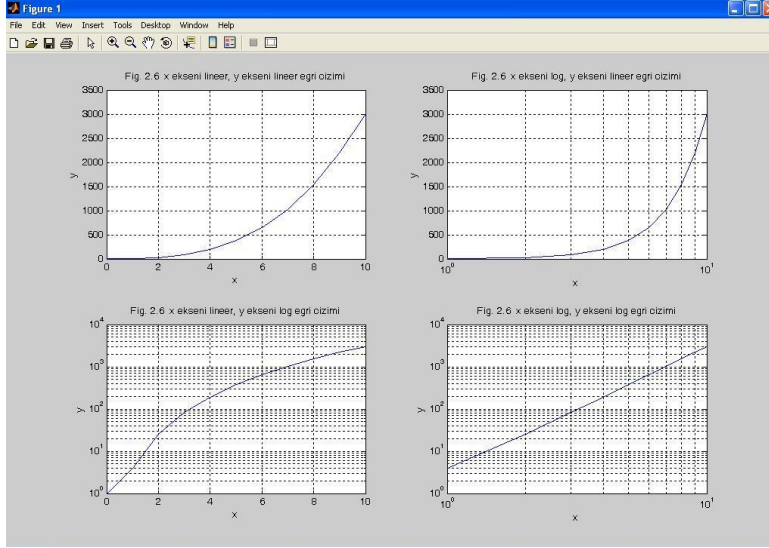
% MATLAB m-file



Şekil Ek 4.1. MATLAB Arayüzü [MATLAB 7.0'dan alınmıştır].



Şekil Ek 4.2. M-File Arayüzü [MATLAB 7.0'dan alınmıştır].



Şekil Ek 4.3. Figure Arayüzü [MATLAB 7.0'dan alınmıştır]

% -----

% 1. MATrix LABoratory

% 1985-C.B. MOLAR

% Yazılım: C programlama dili

% Yazılım programları:

% Structring Programing (SP): Yapısal programlama (Basic, Fortran , ...)

% Object Oriented Programing (OLE): Nesneye Yönelik (Visual Basic 6.0, C, C++,MAT LAB)

% 1970: Programlama yalnızca komutlar (ve/veya fonksiyonlar) aracılığı (DOS komutlari) ile yapılırken,

% 1990: Görselliğin ön plana çıkmasıyla ekran nesneleri(Windows nesneleri) de kullanılmaya başlanmıştır. Böylece nesne özellikleri ve nesne-komut ilişkisi ile bu özelliklerin değiştirilmesi ve de komutlar yardımıyla gerekli düzenlemelerin (hesaplamaların) yapılması gündeme gelmiştir. Bunu takiben Object Linking Embedding (OLE) Nesneyi bağlama yerleştirme yöntemleri geliştirilmiştir. Böylelikle Windows nesnelerini kullanan programlar arasında proje (program) alışverişleri mümkün olabilmektedir. Visual Basic 6.0, C, C++,.....,MAT LAB, Word, Excel, ...yapıları arasında paylaşımlar söz konusudur. Biri diğerine bağlanarak çalıştırılabilmektedir. Bir Word ya da Excel sayfası MAT LAB tarafından okunup işlenebilmektedir.

% 2. MATLAB Kullanım Alanları:

% Matematik/Hesaplama/Algoritma geliştirme

% Modelleme/Benzetim/Prototipleme

% Verilerin Analizi/İncelenmesi/Görüntülenmesi

% Bilimsel ve Mühendislik alanında Grafik İşlemleri

% Grafiksel Kullanıcı Ara yüzü nüde içeren Uygulama Geliştirme
% MATLAB, temel veri elemanı için boyutlandırma gerektirmeyen
% bir dizim (array) olan etkileşimli sistemdir.

% 3. MATLAB Sistemi:

% 5 Ana kısımdan oluşur.

% i. MATLAB Dili:

% Şartlı deyimler, fonksiyonlar, veri yapıları, giriş/çıkış ve nesneye yönelik programlama özellikleri içeren yüksek seviyeli bir matris dizilim dilidir.

% Toolbox 6 adet klasör içinde organize edilmiştir.

% ops: İşletmenler ve özel karakterler

% lang: Programlama dili yapıları

% strfun: Karakter dizileri

% iofun: Dosya giriş/çıkış

% timefun: Zaman ve tarihler

% datatypes: Veri türleri ve yapıları

% ii. Çalışma Ortamı:

% Programcıya çalışma imkanı sağlayan araç ve gereçler takımıdır. Çalışma

% ortamındaki değişkenleri kontrol etmeyi, verileri çalışma ortamının dışına

% veya içine aktarmayı, uygulamaları geliştirmeyi, (M-dosyalar), hataları

% izlemeyi sağlar general adli bir klasörde toplanmıştır.

% iii. Grafik Sistemi:

% (Handle Graphics): 2- ya da 3- boyutlu veri görüntüleme, işleme, animasyon

% ve grafik sunumu için yüksek seviyeli komutları içerir. Ayrıca kullanıcı

% ara birimi (GUI) oluşturmak kadar, grafiklerin görüntüsünü düzenlemeye

% kadar düşük seviyeli komutları da içerir.

% Toolbox içerisinde 5

% klasörde yer alır.

% graph2d: 2-boyutlu grafikler

% graph3d: 3-boyutlu grafikler

% specgraph: Özel grafikler

% graphics: Grafik yönetimi

% uitools: Grafik kullanıcı ara birimi

% iv. Matematiksel Fonksiyon Kütüphanesi

% sum, sin, ... gibi elemensal fonksiyonlardan ve karmaşık aritmetikten,

% matris işlemleri, Bessel fonksiyonları ve Fourier dönüşümleri gibi

% hesaplama algoritmaları içerir. Toolbox 8 klasör içinde düzenlenmiştir.

% elmat: Elemansal matrisler ve matris işlemleri

% elfun: Elemansal matematik fonksiyonları

% specfun: Özel matematik fonksiyonları

% matfun: Matris fonksiyonları-sayısal lineer cebir

% datafun: Veri analizi ve Fourier dönüşümleri

% polyfun: Ara değer (interpolasyon) ve polinomlar

% funfun: Fonksiyon fonksiyonları ve dif. denk. (ODE) çözüm fonksiyonları

% sparfun: Seyrek (sparse) matrisler

% v. Uygulama Ara Birimi (API):

% MATLAB ile etkileşimli çalışan C ve Fortran programları yazılması için hazırlanmıştır.

% MATLAB rutinlerini çağırma (dinamik ilişki), MATLAB i bir hesaplama motoru gibi çağırma

% ve MATLAB dosyalarını okuma ve yazma için gerekli araçlar.

% MATLAB'a İlişkin Kısa Erişim Kılavuzu

Aşağıda MATLAB programlama dilinin temel özellikleri ve komutları kısaca özetlenmiştir.

% MATLAB m-file

% -----

% MATLAB Programlama Dilinden Seçilmiş Bazı Komutlar

% ÇALIŞMA ORTAMI KOMUTLARI

% help:	<i>Fonksiyon yardım komutu</i>
% demo, intro:	<i>Demo ve tanıtım komutu</i>
% whos, what:	<i>Kayıtlı değişkenler ve türleri</i>
% pwd, cd:	<i>Klasör gösterme, değiştirme komutu</i>
% size:	<i>Matris uzunluk komutu</i>
% length:	<i>Matris uzunluk boyutu</i>
% tic, toc:	<i>Zaman başlatma ve görüntüleme komutu</i>
% quit, exit:	<i>Matlab'den çıkış komutu</i>
% clear:	<i>Çalışma ortamındaki tüm değişkenleri silme komutu</i>
% clear x:	<i>Çalışma ortamındaki x değişkenini silme komutu</i>
% save :	<i>Çalışma ortamını matlab.mat olarak kaydetme komutu</i>
% load ad :	<i>Çalışma ortamına ad.mat yükleme komutu</i>
% print	<i>Çalışma alanını veya sekli yazdırma komutu</i>

% SEMBOLLER

% :	<i>Windows, Unix, dos ... Başlatma Komutu</i>
%,	<i>Açıklama getirme ifadesi</i>
% []:	<i>Matris girme ifadesi</i>
% ():	<i>İndis ve değer girme ifadesi</i>
% = :	<i>Eşitlik ifadesi</i>
% :	<i>Kolon ifadesi</i>

% MATEMATİKSEL OPERATÖRLER

% +, -, *, /, ^, .*, /., ^ sqrt: *Toplama, Çıkarma, Çarpma, Bölme, Üst alma*
Elemanter Çarpma, bölme ve üst alma Karekök alma

% MANTIKSAL OPERATÖRLER

% >, =:	<i>Büyüktür, büyük eşittir ifadesi</i>
% <, <=:	<i>Küçüktür, küçük eşittir</i>
% ==:	<i>Koşullu eşittir (Estir)</i>
% ~=:	<i>Döngülerde eşittir ve eşit değildir ifadesi</i>
% & ~:	<i>Ve, veya, değil ifadeleri</i>

% SABİT İFADELER

% pi:	<i>3.14.....</i>
% inf:	<i>sonsuz</i>

% TRİGONOMETRİK, LOGARİTMİK VE KOMPLEKS FONKSİYONLAR

% cos, sin, tan, cot:	<i>Trigonometrik ifadeler</i>
% exp, log:	<i>e ve ln anlamlı logaritma ifadeleri</i>
% log10:	<i>10 tabanında logaritma</i>
% ab:	<i>Mutlak değer</i>
% fix:	<i>Sıfır yuvarlama komutu</i>

% KOORDİNAT DÖNÜŞTÜRME FONKSİYONLARI

% cart2sp:	<i>Kartezyenden küresele</i>
------------	------------------------------

% TEMEL MATRİS İŞLEMLERİ

% ones, zeros:	<i>Sıfırlar ve birler matrisi</i>
% eye:	<i>Birim köşegen matrisi</i>
% rand, randn:	<i>Rasgele matris ve rasgele pozitif ve negatif</i>
%	<i>değerli matris oluşturma ifadeleri</i>

% MATEMATİKSEL İŞLEMLER

% sum, cumsum:	<i>Toplama ve kümülatif toplama</i>
% diff:	<i>Çıkarma</i>
% mean:	<i>Ortalama ifadesi</i>
% median:	<i>Orta değeri gösterme komutu</i>
% min, max:	<i>Minimum ve maximum ifadeleri</i>
% sort:	<i>Matris elemanlarını küçükten büyüğe sıralama</i>
% factorial:	<i>Faktoriyel alma</i>

% LİNEER CEBİR

%inv:	<i>Tersini alma</i>
% eig:	<i>Özdeğer bulma</i>
% det :	<i>Determinant alma</i>
% ' :	<i>Transpoze alma</i>
% root:	<i>Kök bulma</i>

GRAFİK

% plot, plot3	<i>2 ve 3 boyutlu grafik çizdirme komutu</i>
% histbar, bar:	<i>Histogram ve bar grafiği çizdirme komutu</i>
% surf, mesh:	<i>3 boyut tarama komutu</i>
% colormap:	<i>Renk haritası komutu</i>
% colrbar:	<i>Renk bari komutu</i>
% axis, grid:	<i>Eksenler ve tarama komutları</i>
% view:	<i>Görüntüleme komutu</i>
% xlabel, ylabel	
% zlabel, title:	<i>Grafik yazı girdi komutları</i>
% text, gtex:t	<i>Yazı ekleme komutu</i>
% hold on/off:	<i>Grafiği saklar, işlemi iptal etme komutu</i>

% EGRİ UYDURMA

% polyfit:	<i>Polinom uydurma komutu</i>
% polyval :	<i>Polinom değerleri bulma komutu</i>

% PROGRAMLAMA

% function [...] = isim (...):	
%	<i>FonksiyonM-Fonksiyon dosyası</i>
% nargin:	<i>Fonksiyon girdi sayısı komutu</i>
% nargsout:	<i>Fonksiyon çıktı sayısı komutu</i>

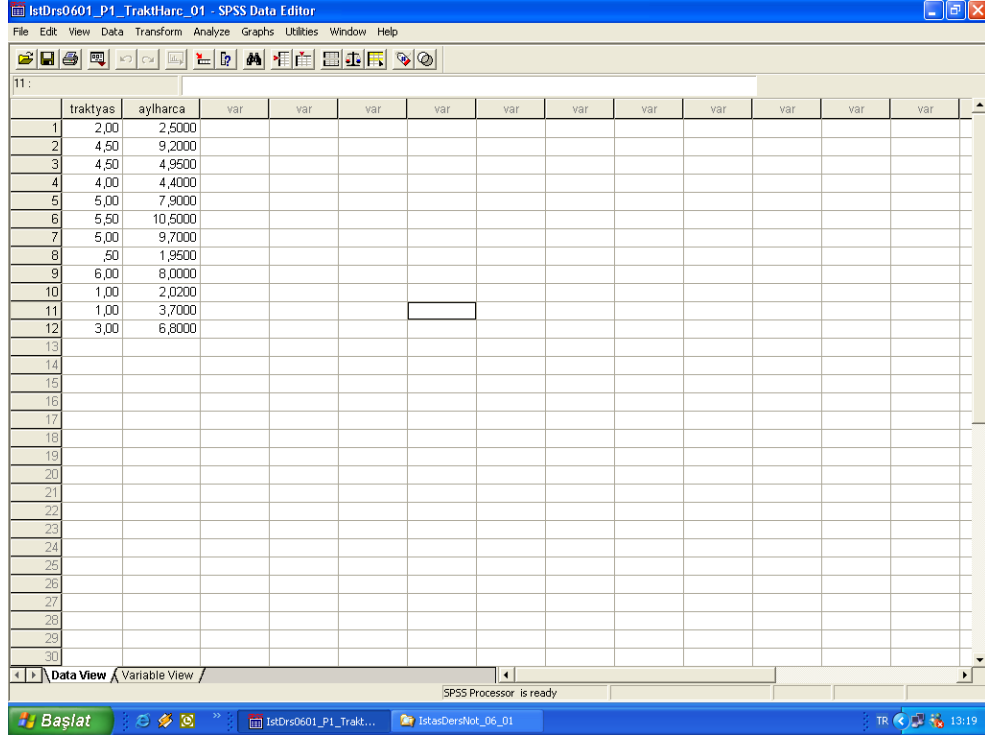
<code>% for ... end:</code>	<i>For döngüsü</i>
<code>% if ... else ...:</code>	<i>f döngüsü</i>
<code>% elseif ... end :</code>	<i>If döngüsü</i>
<code>% while ... end :</code>	<i>While döngüsü</i>
<code>% input:</code>	<i>Girdi alma komutu</i>
<code>% disp:</code>	<i>Karakter (string), değişken ve ekran çıktısı alma komutu</i>
<code>% fprintf:</code>	<i>İşlemsel ifade yapan ekran çıktısı alma komutu</i>

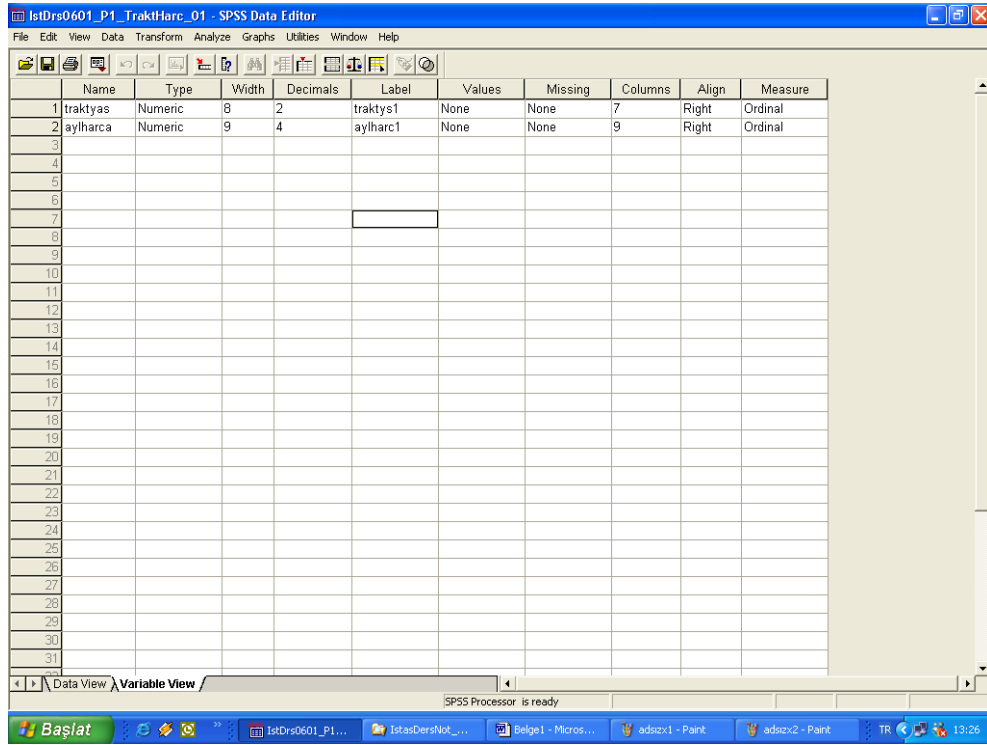
% KARAKTER DİZİSİ FONKSİYONLARI

<code>% char(S1, S2, ...):</code>	<i>Karakter ya da hücre dizilerinden</i>
<code>%</code>	<i>karakter dizileri oluşturma komutu</i>
<code>% double(S:</code>	<i>Karakter ASCII formuna dönüştürme komutu</i>
<code>% upper(S):</code>	<i>Büyük harfe çevirme komutu</i>
<code>% lower(S):</code>	<i>Küçük harfe çevirme komutu</i>
<code>% num2str:</code>	<i>Sayıyı karaktere çevirme komutu</i>
<code>% int2st :</code>	<i>Tam sayıyı karaktere çevirme komutu</i>
<code>% mat2str:</code>	<i>Matris elemanlarını karaktere çevirme komutu</i>

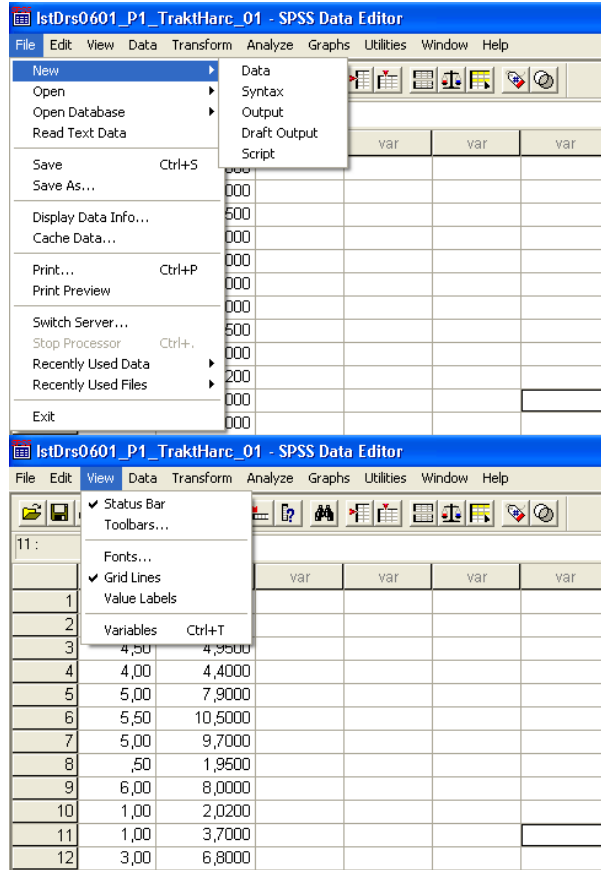
Ek 5. SPSS İle Programlama

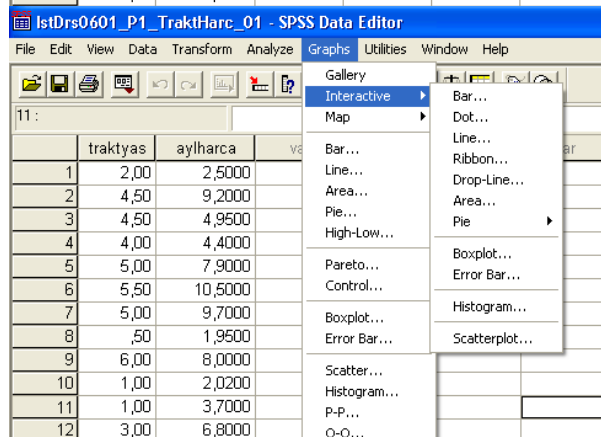
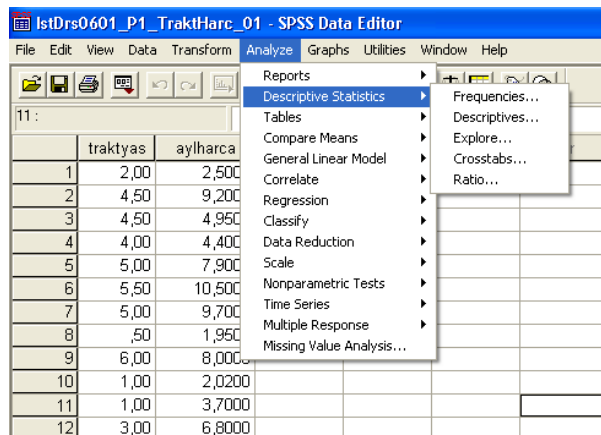
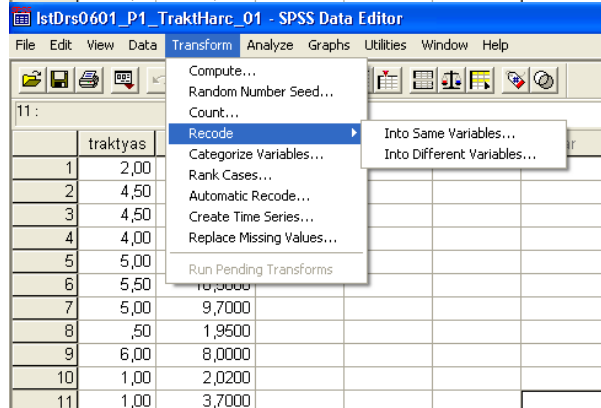
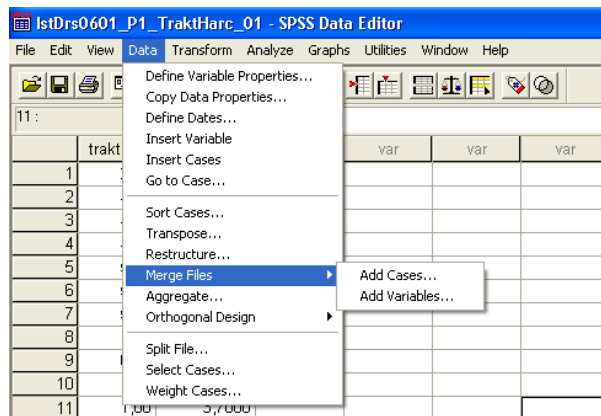
SPSS paket programı kullanıcıya istatistiksel problemlerin çözümünde buna yönelik olarak hazırlanmış olan menü'leri yardımıyla büyük kolaylıklar taşımaktadır. Aşağıdaki şekillerde bu program menü düzenlemeleri kısaca gösterilmiştir (Akpolat, 2008).

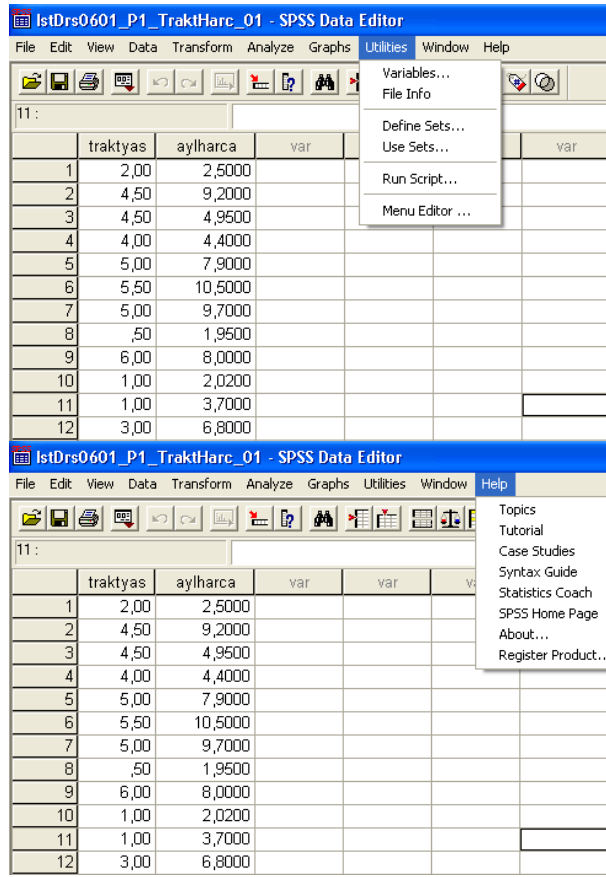




Şekil Ek 5.1. Verilerin istatistiksel değerlendirilmesi için SPSS11.5 te yazılmış kayıt dosyası.







Şekil Ek 5.2. Verilerin istatistiksel değerlendirilmesi için SPSS11.5 ta var olan menü alternatifleri

ÖZGEÇMİŞ

Sinem Çağlar ODABAŞ

Adres: Ilıca mah. Karel sok. no:15/8 B blok Narlıdere/İZMİR

Cep Telefonu: 05378259642

E-mail: caglarodabas@gmail.com

Kişisel Bilgiler

Doğum Tarihi:01.05.1985

Doğum Yeri: İzmir

Eğitim

2004 - 2009	<u>Muğla Üniversitesi</u> Fen Edebiyat Fakültesi Kimya Bölümü
1999 – 2003	<u>Yunus Emre Anadolu Lisesi (İzmir)</u> Lise
1996 – 1999	<u>Misakı Milli İlköğretim Okulu</u> Orta öğretim
1991 – 1996	<u>Balçova İlkokulu</u> İlkokul

İş ve Staj Deneyimi

2006	<u>AFZ TEKSTİL (DENİZLİ)</u> 6 Haftalık yaz pratiği (Tesis İşletmesi ve Laboratuvar)
------	--

Yabancı Dil Bilgisi

<u>Almanca</u> (Orta Seviyede-ÜDS)	<u>İngilizce</u> (Orta Seviyede)
--	--

Bilgisayar Bilgisi

- **İşletim Sistemleri:** Windows 95/98, Windows XP

- **Paket Programlar:** Microsoft Word, Excel, Access, Powerpoint, Matlab, Makromedia Flash, SPSS, Dreamweaver

Katılan Eğitimler/Seminerler

2009	<u>9. Dinamik Sistemler Çalıştayı</u> •Diferensiyel denklemler •Fark denklemleri •Zaman skalası, •Kısmi türevli denklemler, •Sayısal yöntemler, •İlgili uygulamalar.
2009	<u>II. KEMOMETRİ YAZ OKULU</u> • Temel İstatistik • İstatistiksel Proses Kontrol • Metot Validasyonu • Deneysel Tasarım ve Optimizasyon • Sınıflandırma ve Kümeleme Teknikleri • Çok Değişkenli Kalibrasyon Metotları
2010	<u>AKADEMİK BİLİŞİM (MUĞLA ÜNİVERSİTESİ)</u>
2011	<u>AKADEMİK BİLİŞİM (İNÖNÜ ÜNİVERSİTESİ)</u>
2011	<u>TEMEL KEMOMETRİ EĞİTİMİ (AKDENİZ ÜNİ.)</u>

Yayınlar

Odabaş, S. Ç., Özevci, G. , Coşkun, B., Ayhan, F., Ayhan, H., Akpolat, O., (2011), Fotoçarazbağlı Hema Hidrojellerin Dinamik Şişme Davranışının Modellenmesi, Akademik Bilişim 2011, İnönü Üniversitesi, Malatya.