

**KÜMELEME ANALİZİ YÖNTEMLERİNİN HAYVANCILIK  
VERİLERİNDE KARŞILAŞTIRILMALI OLARAK  
İNCELENMESİ**

**Mehmet DİNLER**

**Yüksek Lisans Tezi**

**Zootekni Anabilim Dalı**

**Danışman: Yrd. Doç. Dr. Şenol ÇELİK**

**2014**

**Her hakkı saklıdır**

T.C.  
BİNGÖL ÜNİVERSİTESİ  
FEN BİLİMLERİ ENSTİTÜSÜ

**KÜMELEME ANALİZİ YÖNTEMLERİNİN  
HAYVANCILIK VERİLERİNDE KARŞILAŞTIRILMALI  
OLARAK İNCELENMESİ**

**YÜKSEK LİSANS TEZİ**

**Mehmet DİNLER**

**Enstitü Anabilim Dalı : ZOOTEKNİ**

**Tez Danışmanı : Yrd. Doç. Dr. Şenol ÇELİK**

**Mayıs 2014**

T.C.  
BİNGÖL ÜNİVERSİTESİ  
FEN BİLİMLERİ ENSTİTÜSÜ

**KÜMELEME ANALİZİ YÖNTEMLERİNİN HAYVANCILIK  
VERİLERİNDE KARŞILAŞTIRILMALI OLARAK İNCELENMESİ**

**YÜKSEK LİSANS TEZİ**

**Mehmet DİNLER**

**Enstitü Anabilim Dalı : ZOOTEKNİ**

**Bu tez 26.05.2014 tarihinde aşağıdaki jüri tarafından oy birliği ile kabul edilmiştir.**

**Prof. Dr.  
Turgay ŞENGÜL  
Jüri Başkanı**

**Yrd. Doç. Dr.  
Şenol ÇELİK  
Üye**

**Doç. Dr.  
Ramazan MERAL  
Üye**

**Yukarıdaki sonucu onaylarım**

**Doç. Dr. İbrahim Y. ERDOĞAN  
Enstitü Müdürü**

## ÖNSÖZ

Tez çalışmaları süresince yardımlarını ve bilgi birikimini esirgemeyen, çalışmaların tamamlanabilmesi için gerekli desteği veren değerli hocam Yrd. Doç. Dr. Şenol ÇELİK'e teşekkür ederim.

Bilgi ve tecrübeleriyle beni yönlendiren, yardımlarını hiçbir zaman esirgemeyen, Yrd. Doç. Dr. Zeki DOĞAN hocama teşekkür ederim.

Prof. Dr. Hüseyin NURSOY, Yrd. Doç. Dr. Hakan İNCİ, Yrd. Doç. Dr. Remzi EKİNCİ hocalarıma, göstermiş oldukları yakın ilgi ve vermiş olduğu destek ve emeklerden dolayı teşekkürlerimi sunuyorum.

Bilimsel hazırlık döneminde desteklerinden dolayı Bingöl Üniversitesi Ziraat Fakültesi Dekanı Prof. Dr. Turgay ŞENGÜL hocama teşekkür ederim.

Son olarak bende büyük emekleri olan, benim için hiçbir fedakârlıktan kaçınmayan ve dualarını esirgemeyen anne ve babama, tezin hazırlanması sırasında gösterdikleri sabır, fedakârlık ve desteklerinden dolayı eşim Sakine DİNLER, çocuklarım Yavuz Berk DİNLER ve Muhammed Tarık DİNLER'e teşekkürü bir borç bilirim.

**Mehmet DİNLER**

**Bingöl 2014**

## İÇİNDEKİLER

ÖNSÖZ.....	ii
İÇİNDEKİLER.....	iii
SİMGELER VE KISALTMALAR LİSTESİ.....	vi
ŞEKİLLER LİSTESİ.....	vii
TABLOLAR LİSTESİ.....	viii
ÖZET.....	x
ABSTRACT.....	xi
GİRİŞ.....	1
ÖNCEKİ ÇALIŞMALAR.....	4
KÜMELEME ANALİZİ.....	10
3.1. Kümeleme Analizinin Tanımı.....	10
3.2. Kümeleme Analizinde Benzerlik ve Yakınlık Ölçüleri.....	12
3.2.1. Öklid (Euclidean) Uzaklığı .....	13
3.2.2. Karesel Öklid (Squared Euclidean) Uzaklığı .....	13
3.2.3. Manhattan ( City-Block) Uzaklığı .....	14
3.2.4. Pearson Uzaklığı .....	14
3.2.5. Mahalanobis Uzaklığı .....	15
3.2.6. Minkowski Uzaklığı .....	15
3.2.7. Karesel Pearson Uzaklığı .....	15
3.2.8. Hotelling T <sup>2</sup> Uzaklığı .....	16
3.2.8. Canberra Ölçütü.....	16
3.3. Kümeleme Yöntemleri.....	16
3.3.1. Hiyerarşik Kümeleme Yöntemleri.....	17
3.3.1.1. Tek Bağlantı Kümeleme Yöntemi.....	18

3.3.1.2. Tam Bağlantı Kümeleme Yöntemi.....	18
3.3.1.3. Ortalama Bağlantı Kümeleme Yöntemi.....	19
3.3.1.4. McQuitty Bağlantı Kümeleme Yöntemi.....	20
3.3.1.5. Merkezi Bağlantı Kümeleme Yöntemi.....	20
3.3.1.6. Medyan Bağlantı Kümeleme Yöntemi.....	21
3.3.1.7. Ward Bağlantı Kümeleme Yöntemi.....	22
3.3.2. Hiyerarşik Olmayan Kümeleme Yöntemleri.....	22
3.3.2.1. k- ortalamalar Kümeleme Yöntemi.....	23
3.4. Küme Sayısının Belirlenmesi.....	24
KÜMELEME ANALİZİ YÖNTEMLERİNİN KARŞILAŞTIRILMASI.....	26
4.1. Verilerin Özellikleri.....	26
4.2. Analizde Kullanılan Kümeleme Yöntemleri ve Programlar.....	28
4.3. Uzaklık Ölçülerinin Karşılaştırılması .....	28
4.3.1. Tek BKY ile Uzaklık Ölçülerinin Karşılaştırılması.....	29
4.3.2. Tam BKY ile Uzaklık Ölçülerinin Karşılaştırılması.....	29
4.3.3. Ortalama BKY ile Uzaklık Ölçülerinin Karşılaştırılması.....	30
4.3.4. Merkezi BKY ile Uzaklık Ölçülerinin Karşılaştırılması.....	31
4.3.5. Medyan BKY ile Uzaklık Ölçülerinin Karşılaştırılması.....	32
4.3.6. McQuitty BKY ile Uzaklık Ölçülerinin Karşılaştırılması.....	33
4.3.7. Ward BKY ile Uzaklık Ölçülerinin Karşılaştırılması.....	34
4.4. Kümeleme Yöntemlerinin Karşılaştırılması .....	36
4.4.1. Öklid Uzaklığına Göre Yöntemlerin Karşılaştırılması.....	36
4.4.2. Karesel Öklid Uzaklığına Göre Yöntemlerin Karşılaştırılması.....	37
4.4.3. Manhattan Uzaklığına Göre Yöntemlerin Karşılaştırılması.....	37
4.4.4. Pearson Uzaklığına Göre Yöntemlerin Karşılaştırılması.....	38
4.4.5. Karesel Pearson Uzaklığına Göre Yöntemlerin Karşılaştırılması.....	39
4.5. Kümeleme Yöntemleri ve Uzaklık Ölçülerine Göre Ortak Kümeler.....	40
4.6. Değişkenlerin Kümeleme Yöntemlerine Etkisi .....	41
4.6.1. Öklid, K. Öklid ve Manhattan Uzaklıklarına Göre Etkisi.....	42
4.6.2. Pearson ve K. Pearson Uzaklıklarına Göre Etkisi.....	43
4.7. İllerin Birbirine Uzaklık ve Yakınlıkları .....	44

SONUÇLAR VE ÖNERİLER.....	46
KAYNAKLAR.....	48
EKLER.....	51
ÖZGEÇMİŞ.....	76

## SİMGELER VE KISALTMALAR LİSTESİ

TBKY	: Tam bağlantı kümeleme yöntemi
OBKY	: Ortalama bağlantı kümeleme yöntemi
MBKY	: Merkezi bağlantı kümeleme yöntemi
MDBKY	: Medyan bağlantı kümeleme yöntemi
WBKY	: Ward bağlantı kümeleme yöntemi
MQBKY	: McQuitty bağlantı kümeleme yöntemi
TKBKY	: Tek bağlantı kümeleme yöntemi
TUİK	: Türkiye İstatistik Kurumu



## ŞEKİLLER LİSTESİ

Ek Şekil 1.	Ortalama bağlantı kümeleme yöntemi dendogramı (öklid) .....	51
Ek Şekil 2.	Merkezi bağlantı kümeleme yöntemi dendogramı (öklid) .....	52
Ek Şekil 3.	Tam bağlantı kümeleme yöntemi dendogramı (öklid).....	53
Ek Şekil 4.	McQuitty bağlantı kümeleme yöntemi dendogramı (öklid).....	54
Ek Şekil 5.	Medyan bağlantı kümeleme yöntemi dendogramı (öklid).....	55
Ek Şekil 6.	Tek bağlantı kümeleme yöntemi dendogramı (öklid).....	56
Ek Şekil 7.	Ward bağlantı kümeleme yöntemi dendogramı (öklid) .....	57
Ek Şekil 8.	Ortalama bağlantı kümeleme yöntemi dendogramı (pearson).....	58
Ek Şekil 9.	Merkezi bağlantı kümeleme yöntemi dendogramı (pearson).....	59
Ek Şekil 10.	Tam bağlantı kümeleme yöntemi dendogramı(pearson).....	60
Ek Şekil 11.	McQuitty bağlantı kümeleme yöntemi dendogramı(pearson).....	61
Ek Şekil 12.	Medyan bağlantı kümeleme yöntemi dendogramı (pearson).....	62
Ek Şekil 13.	Tek bağlantı kümeleme yöntemi dendogramı (pearson).....	63
Ek Şekil 14.	Ward bağlantı kümeleme yöntemi dendogramı (pearson).....	64

## TABLolar LİSTESİ

Tablo 3.1. Veri matrisi.....	11
Tablo 3.2. Benzerlik/farklılık matrisi .....	12
Tablo 4.1. Tek bağlantı kümeleme yöntemi ile oluşan kümeler .....	29
Tablo 4.2. Tam bağlantı kümeleme yöntemi ile oluşan kümeler.....	30
Tablo 4.3. Ortalama bağlantı kümeleme yöntemi ile oluşan kümeler.....	31
Tablo 4.4. Merkezi bağlantı kümeleme yöntemi ile oluşan kümeler.....	32
Tablo 4.5. Medyan bağlantı kümeleme yöntemi ile oluşan kümeler.....	33
Tablo 4.6. McQuitty bağlantı kümeleme yöntemi ile oluşan kümeler.....	34
Tablo 4.7. Ward bağlantı kümeleme yöntemi ile oluşan kümeler.....	35
Tablo 4.8. Öklid uzaklığına göre kümeleme yöntemlerinin karşılaştırılması.....	36
Tablo 4.9. K. öklid uzaklığına göre kümeleme yöntemlerinin karşılaştırılması.....	37
Tablo 4.10. Manhattan uzaklığına göre kümeleme yöntemlerinin karşılaştırılması..	38
Tablo 4.11. Pearson uzaklığına göre kümeleme yöntemlerinin karşılaştırılması.....	39
Tablo 4.12. K. Pearson uzaklığına göre kümeleme yöntemlerinin karşılaştırılması..	40
Tablo 4.13. Öklid, karesel öklid, manhattan ortak kümeleri.....	41
Tablo 4.14. Pearson ve karesel pearson uzaklık ölçüleri ortak kümeleri.....	41
Tablo 4.15. İllerin değişken büyüklüklerine göre sıralanışı 1. kısım.....	42
Tablo 4.16. İllerin değişken büyüklüklerine göre sıralanışı 2. kısım.....	42
Tablo 4.17. İllerin birbirine uzaklık ve yakınlıkları.....	45
Ek Tablo 1. Yerli koyunculuk istatistikleri .....	65
Ek Tablo 2. Merinos koyunculuk istatistikleri .....	66
Ek Tablo 3. Kıl keçisi istatistikleri .....	67
Ek Tablo 4. Tiftik keçisi istatistikleri .....	68
Ek Tablo 5. Kültür sığırı istatistikleri .....	69
Ek Tablo 6. Melez sığır istatistikleri .....	70
Ek Tablo 7. Yerli sığır istatistikleri .....	71
Ek Tablo 8. Manda istatistikleri .....	72

Ek Tablo 9.	Et ve Yumurta Tavuđu İstatistikleri .....	73
Ek Tablo 10.	Kaz, Ördek ve Hindi İstatistikleri .....	74
Ek Tablo 11.	Arıcılık İstatistikleri .....	75

# KÜMELEME ANALİZİ YÖNTEMLERİNİN HAYVANCILIK VERİLERİNDE KARŞILAŞTIRILMALI OLARAK İNCELENMESİ

## ÖZET

Bu çalışmada; Türkiye’de 81 ilin hayvancılık potansiyeli yedi farklı kümeleme analizi yöntemi ve beş farklı uzaklık ölçüsü ile karşılaştırmalı olarak incelenmiştir. Türkiye İstatistik Kurumu internet sitesinde hayvancılık istatistikleri bölümünden dinamik sorgulama ile elde edilen 2012 yılına ait büyükbaş hayvan, küçükbaş hayvan, kümes hayvancılığı ve arıcılık faaliyetlerine ait veriler kullanılmıştır.

Veri olarak, büyükbaş hayvancılık ile ilgili yerli yetişkin sığır sayısı, genç-yavru yerli sığır sayısı, sağılan yerli sığır sayısı, yerli sığırdan elde edilen süt miktarı, yetişkin kültür sığır sayısı, genç-yavru kültür sığır sayısı, sağılan kültür sığır sayısı, kültür sığırdan elde edilen süt miktarı, yetişkin melez sığır sayısı, genç-yavru melez sığır sayısı, sağılan melez sığır sayısı, melez sığırdan elde edilen süt miktarı, yetişkin manda sayısı, genç-yavru manda sayısı, sağılan manda sayısı, mandadan elde edilen süt miktarı verileri kullanılmıştır. Küçükbaş hayvancılık ile ilgili yerli koyun sayısı, genç-yavru yerli koyun sayısı, sağılan yerli koyun sayısı, yerli koyundan elde edilen süt ve yapağı miktarı, merinos koyun sayısı, genç-yavru merinos koyun sayısı, sağılan merinos koyun sayısı, merinos koyundan elde edilen süt ve yapağı miktarı, kıl keçisi sayısı, genç-yavru kıl keçisi sayısı, sağılan kıl keçisi sayısı, kıl keçisinden elde edilen süt ve tiftik miktarı, tiftik keçisi sayısı, genç-yavru tiftik keçisi sayısı, sağılan tiftik keçisi sayısı, tiftik keçisinden elde edilen süt ve tiftik miktarı verileri kullanılmıştır. Kümes hayvancılığı ile ilgili yumurta tavuğu, et tavuğu, hindi, ördek ve kaz sayısı verileri kullanılmıştır. Arıcılık ile ilgili kovan sayısı, bal ve balmumu üretim miktarı verileri kullanılmıştır.

Verilerden elde edilen 44 adet değişken, hiyerarşik kümeleme analizi yöntemlerinden ortalama (average), merkezi (centroid), tam (complete), mcquitty, ortanca (median), tek (single) ve ward bağlantı kümeleme yöntemleri ile analiz edilmiştir. Uzaklık ölçüsü olarak öklid (euclidean), karesel öklid (squared euclidean), pearson, karesel pearson ve manhattan uzaklık ölçüleri kullanılmıştır. 5,6,7,8,9 ve 10’lu kümeler bütün kümeleme yöntemleri ve uzaklık ölçüleri için ayrı ayrı oluşturulmuştur. Kümeleme yöntemlerinden ward bağlantı kümeleme yöntemi dışındaki yöntemlerin benzer kümeler oluşturduğu gözlenmiştir. Uzaklık ölçülerinden öklid, karesel öklid ve manhattan uzaklıkları benzer kümeler oluştururken pearson ve karesel pearson kendi içinde benzer diğer uzaklık ölçülerinden farklı kümeler oluşturduğu gözlenmiştir.

**Anahtar Kelimeler:** Kümeleme analizi, hiyerarşik kümeleme yöntemleri, Türkiye’de hayvancılık.

# THE COMPARATIVE INVESTIGATION OF CLUSTER ANALYSIS METHODS IN LIVESTOCK DATA

## ABSTRACT

In this study; Turkey animal husbandry potential of the 81 provinces were examined and compared with five different distance measures and seven different clustering analysis methods. Turkey Statistical Institute section on the website livestock statistics for the year 2012 obtained by dynamic querying of cattle, small ruminants, poultry and beekeeping activities data were used.

In cattle raising-related data, numbers of domestic adult cattle, young-pup domestic cattle, milking domestic cattle, young-pup culture, milking culture cattle, adult hybrids cattle, young-pup hybrid cattle, milking hybrid cattle, adult buffalo, young-pup buffalo, milking buffalo and the amount of milk obtained from domestic cattle, adult culture cattle, crossbred cattle, buffalo were used. In Small Ruminant livestock-related data counts of domestic sheep, young-pup domestic sheep, milking domestic sheep, merino sheep, young-pup merino, milking merino, goat, teen-fry goat, milking goats, Angora goat, teen-fry angora goat, milking the angora goat and milk and wool-hair-mohair quantity obtained from domestic sheep, merino sheep, goat and Angora goat were used. In Poultry-related data, numbers of egg-chicken, meat-chicken, turkey, duck and goose were used. About beekeeping hives, honey and wax production quantity data were used.

44 units variable obtained from data, hierarchical cluster analysis of the mean (average), center (centroid), full (complete), mcquitty, median, single and ward link aggregation methods were analyzed. Measure of distance "x" as euclid (euclidean), squared Euclidean, pearson, squared pearson and manhattan quadratic distance measures were used. Of 5, 6, 7, 8, 9 and 10 clusters for all clustering methods and distance measures have been created separately. Except the ward link aggregation method, similar clusters have been observed in the other clustering methods. Of the euclidean distance measure, similar to the squared euclidian distance and manhattan creating clusters pearson, squared pearson and quadratic distance measure in itself different from other similar form clusters were observed.

**Keywords:** Cluster analysis, hierarchical cluster analysis methods, animal husbandry in Turkey.

## 1. GİRİŞ

Kümeleme analizi grupları kesin olarak bilinmeyen, birimleri, değişkenleri birbiriyle benzer alt kümelere (grup, sınıf) ayırmaya yardımcı olan çok değişkenli istatistiksel analiz yöntemlerinden biridir. Kümeleme analizinin temel amacı birimleri sahip oldukları karakteristik özellikleri temel alarak gruplandırmaktır. Kümeleme analizi son yıllarda gündemde olan analiz yöntemlerinden biridir. Bu yöntem özellikle bilim ve iş alanında birçok durumda uygulanabilen, en kolay yorumlanabilen ve en etkili olan yöntem olma özelliğini taşır. Bu nedenle hemen hemen tüm bilim alanlarında bu yöntemden yararlanılmaktadır (Özdamar 2004).

Genetik, istatistik, biyoloji, tıp, ziraat, veterinerlik, mühendislik, iktisadi ve idari bilimler, teknik bilimler, sosyal bilimler, veri madenciliği vb. alanlarda yapılan bilimsel çalışmalarda, özellikle verilen sınıflandırılmasında çok tercih edilen bir istatistiksel analiz yöntemi olarak karşımıza çıkmaktadır.

Kümeleme analizinde, heterojen yapıdaki verinin küme sayısı ve küme yapısı araştırılır. Kümelemede amaç, her bir küme içerisindeki gözlemlerin ya da nesnelere birbirlerine benzer ve kümelerin de birbirinden farklı olacak şekilde en uygun gruplama yapısını bulmaktır. Kümeleme analizi temel olarak ayrıştırma analizinden farklıdır. Ayrıştırma analizinde gözlemler daha önceden tanımlanmış ve sayısı bilinen gruplara parçalanırken, kümeleme analizinde ne grup sayısı hakkında ne de grupların yapısı hakkında bir ön bilgi bulunmaz. Gözlemlerin kümelere gruplanması için geliştirilen bazı yöntemlerde kümeleme, tüm gözlem çiftleri arasındaki benzerliklerin bulunmasıyla başlar. Bazı durumlarda benzerlikler, uzaklık ölçümlerine dayalı olarak bulunur. Diğer kümeleme yöntemlerinde, küme merkezlerinin seçimi veya küme içi ve kümeler arası değişimin karşılaştırılması yapılır. Değişkenlerin de kümelenebilirliği mümkündür (Galimberti and Soffritti 2007).

Kümeleme analizinde uygulama aşamaları dört basamaktan oluşmaktadır. İlk basamakta popülasyondan herhangi bir sınıflandırılmaya tabi tutulmamış gözlemler birimler ve değişkenlerden oluşan veri matrisi (Birim x Değişken ) haline getirilir. İkinci basamakta birimlerin ve değişkenlerin benzerliklerini veya farklılıklarını ortaya koyan öklid, manhattan, pearson vb. benzerlik/uzaklık ölçülerinden biri kullanılarak birimlerin birbirlerine olan uzaklık ve yakınlıklarını belirleyen matris (benzerlik/farklılık matrisi) elde edilir. Üçüncü basamakta tercih edilecek uygun hiyerarşik veya hiyerarşik olmayan kümeleme yöntemlerinden biriyle benzerlik/farklılık matrisindeki birimler/değişkenler kümelerine ayrılır. Dördüncü basamakta tercih edilen uygun ölçü ve yöntemlerle oluşturulan kümeler yorumlanır.

Kümeleme analizinin en kritik konusu küme sayısına karar vermektir. Araştırmacının küme sayısına karar vermede öznelliği minimize etmesi gerekmektedir. Küme sayısının doğru olarak seçilmiş olması oluşturulacak olan kümelerin kalitesini belirlemektedir. Uygun küme sayısının belirlenmesi amacıyla birçok yöntem geliştirilmiştir. Ancak günümüzde yayınlanan birçok bilimsel makalede küme sayısının belirlenmesinde net bir yöntem yoktur (Atbaş 2008).

Küme sayısının belirlenmesinde farklı yöntemlerden faydalanılsa da bu konuda araştırmacının bilgi düzeyi, mesleki tecrübesi ve sonuçların anlamlı olup olmaması en önemli etkidir. Kümeleme analizinde kullanılan pek çok uzaklık ölçüsü ve bu ölçüler üzerine kurulmuş pek çok yöntemle bağlı olarak sonuçlar çok farklı çıkabilmekte ve araştırmacı kararsızlığa düşebilmektedir. Bu nedenle son yıllarda çok kullanılan bir yol kümeleme analizinde temel bileşenlerden yararlanılmasıdır. Bu yolla (çok sayıda değişken olması durumunda) hem değişken sayısı azaltılmakta hem de özellikle ilk iki temel bileşen üzerindeki gözlem sonuçlarının (skor) çiziminden ayrıntılı bilgi çıkarmak mümkün olabilmektedir. Bu çizim araştırmacıya kümeleme analizi sonuçlarının doğruluğu ve yorumu konusunda fikir vermektedir (Tatlıdil 1996).

Tez çalışmasında, kümeleme analizi yöntemleri, uzaklık/yakınlık ölçüleri karşılaştırmalı olarak incelenerek, yöntemlerin zayıf ve güçlü yanları araştırılmıştır. Ülkemizde yapılan büyükbaş hayvancılık, küçükbaş hayvancılık, kümes hayvancılığı ve arıcılık ile ilgili Türkiye İstatistik Kurumunun internet sitesinden elde edilen 81 ilimize ait veri setine

kümeleme analizi yöntemleri uygulanarak çeşitli sayıda kümeler elde edilecektir. Elde edilen kümelerle kümeleme yöntemleri kıyaslanacaktır.

Bu çalışmada hazırlanan tez beş bölümden oluşmaktadır. Giriş bölümünde; kümeleme analizini önemi, gerekliliği, kümeleme analizinin tanımı, kümeleme analizi yöntemlerinin ve uzaklık/yakınlık ölçülerinin kullanılması, yapılan çalışmanın gerekçesi hakkında bilgiler verilmiştir. Bölüm 2’de; kümeleme analizi ilgili tanım ve kavramlar, kümeleme analizinin kullanım alanları, kümeleme analizi yöntemleri uzaklık/yakınlık ölçüleri incelenecektir. Bölüm 3’te kümeleme analizi ile ilgili literatür taramasına yer verilecektir. Kümeleme analizi ile ilgili önceki yapılan çalışmalar özetlenecektir. Bölüm 4’te, Türkiye’deki 81 ilin hayvancılık potansiyeli kümeleme analizi yöntemleri ile karşılaştırılacaktır. Yöntemlerin üstün ve zayıf yönleri kıyaslanacaktır. Bu çalışmanın son kısmı olan 5. bölümde ise elde edilen kümelerin anlamı, kümeleme analizi yöntemleri karşılaştırılacaktır.



## 2. ÖNCEKİ ÇALIŞMALAR

Özbeyaz vd (1999), Türkiye'nin farklı işletmeleri ve değişik bölgelerinde yetiştirilen yerli ve kültür ırkı olmak üzere toplam 820 baş sığira ait kan örneklerinden elde edilen veriler kümeleme analizi ile incelemiştir. Irkların sınıflandırılmasında kümeleme analizinden etkin bir şekilde faydalanılabileceğini söylemiştir.

Gürcan vd (2002), Karacabey harasında yetiştirilen Alman et merinosu ve Karacabey merinosu koyunlarının canlı ağırlık, vücut ölçüleri ve yapağı inceliği yönünden kümeleme analizi yöntemleriyle incelemiştir. Kümeleme analizinde kümeleme yöntemi olarak hiyerarşik kümeleme yöntemlerinden tek bağlantı kümeleme yöntemini ve uzaklık ölçüsü olarak öklid uzaklık ölçüsünü kullanmışlardır. Karacabey merinosu genotipinin bu merinosun elde edilmesinde baba ırkı olarak kullanılan Alman et merinosu genotipine benzer hale geldiğini söylemişlerdir.

Doğan (2002), Çifteler Tarım İşletmesinde yetiştirilen ve 1980-1998 yılları arasında satışa sunulan 535'i erkek, 392'si dişi olan toplam 927 Arap taylarına ait yaş, cidago yüksekliği, göğüs çevresi ve incik çevresine ait verileri hiyerarşik olmayan kümeleme yöntemlerinden k-ortala yöntemi ile analiz etmiştir. İki özellik bakımından tercih edilen bir kümenin üçüncü bir özellik bakımından tercih edilememesi ya da her bir değişken için farklı kümelerin seçilmesinin gerektiği durumlarda karar vericinin bir karar vermesi gerekebileceğini söylemiştir. Diğer seleksiyon metotlarına göre zaman, işgücü, maliyet, güvenilirlik, objektiflik vb. yönlerden avantajlı olmasından dolayı hayvan ıslahında özellikle de seleksiyon yaparken bir metot olarak kümeleme analizinin kullanılmasının uygun olacağını söylemiştir.

Urfalıoğlu vd (2004), Türkiye'deki 81 ilin sosyo-ekonomik yapısı 19 değişken kriter göz önünde bulundurularak kümeleme analizi ile gelişmişlik sınıflaması yapmışlardır.

Kümeleme analizinde hiyerarşik kümeleme yöntemlerinden tek bağlantı yöntemini, uzaklık ölçüsü olarak öklid uzaklık ölçüsünü kullanmışlardır.

Karabulut vd (2004), Türkiye'nin 81 iline ait 54 sosyo-ekonomik değişken (DPT) yardımıyla, aynı yapıyı gösteren homojen il gruplarının belirlenmesine çalışmışlardır. Bunun için hiyerarşik kümeleme analizi yöntemlerinden ortalama bağlantı kümeleme yöntemini, mesafe olarak kareli öklid ve pearson ölçülerini kullanmışlardır. Bu yöntemle illerin ayırımına ve homojen yapı göstermelerine neden olan değişkenler yardımıyla illerin oluşturduğu farklı sosyo- ekonomik bölgeleri belirlemişlerdir.

Çelik vd (2006), Lalahan Hayvancılık Araştırma Enstitüsünde yetiştirilen Ankara (Tiftik) keçisinin 1989-1997 arasında 9 yıllık kayıtlardan elde edilen tiftik özelliklerini (incelik, uzunluk, ondulasyon, elastikiyet, mukavemet) kümeleme analizi yöntemlerinden tek bağlantı yöntemi ile incelemişlerdir. Sonuç olarak çeşitli fenotipik değerler kullanılarak ele alınan Ankara keçisi grubunun 4 ile 7 arasında kümeye ayrıldığını tespit etmişlerdir. Araştırmada elde edilen dendogramlardan dışı Ankara keçilerinin erkeklere göre daha homojen bir yapıya sahip olduklarını görmüşlerdir.

Özgen vd (2008), üniversite öğrencilerinin coğrafya dersine yönelik tutumlarının kümelene eğilimlerini belirlemeye çalışmışlardır. Araştırmaya Siirt Üniversitesi, Eğitim Fakültesi, İlköğretim Sosyal Bilgiler ve Sınıf Öğretmenliği programlarında okuyan 314 (220 erkek, 94 kız) öğrenci almışlardır. Verilerin analizinde aşamalı kümeleme yöntemlerinden ward kümeleme yöntemini, uzaklık ölçüsü olarak karesel öklid uzaklığını seçmişlerdir. Ward kümeleme yönteminin anlamlı kümeler oluşturduğunu söylemişlerdir.

Öz vd (2009), beşeri sermayenin bileşenleri olan eğitim ve sağlık ile işgücü piyasalarına ilişkin göstergeler açısından Türkiye ve AB üyelerinin bir karşılaştırmasını yapmaktadırlar. Eğitime (6), sağlık (7) ve işgücü piyasalarına (10) ilişkin toplam 23 değişken analizde kullanmışlardır. Benzer ülke gruplarını belirleyebilmek amacıyla hiyerarşik kümeleme analizi yöntemlerinden ortalama bağlantı (average linkage) yöntemine başvurmuşlardır. Birimlerin sınıflandırılmasında ise kareli öklid uzaklık (Squared euclidean distance) benzerlik ölçüsü olarak almışlardır.

Kılıç vd (2010), Tarım İşletmeleri Genel Müdürlüğüne bağlı Gökhöyük Tarım İşletmesinde yetiştirilen 100 baş Karayaka ve 100 baş Bafra koyununun vücut ölçülerine ait yaşa göre düzeltilmiş verilerine hiyerarşik olmayan kümeleme yöntemlerinden bulanık kümeleme analizi yöntemini uygulamışlardır. Türkiye’de hayvancılık alanında yapılan araştırmalarda hemen hemen hiç uygulanmamış olan bulanık kümeleme yöntemi ile klasik kümeleme yöntemlerinden farklı olarak bireylerin küme üyeliklerindeki kararsızlık (bulanıklık) ortaya konulduğu ve küme üyelik oranları belirlenebildiği için konuyla ilgili bundan sonra yapılacak araştırmalarda hem birey hem de popülasyonun durumu hakkında daha ayrıntılı bilgiler elde edilebileceğini ifade etmişlerdir.

Atalay vd (2010), 1997-2006 yıllarında şehir dışında meydana gelen trafik kazası verilerini kullanarak illerin kümelemesini yapmışlardır. Kümeleme analizi için hem geleneksel k-ortalımalı hem de bulanık c-ortalımalı kümeleme yöntemlerini kullanmışlardır. Bulanık bulanık c-ortalımlar yönteminin daha karalı sonuçlar ürettiğini gözlemlemişlerdir.

Yılcı (2010), Türkiye’deki 81 il, 11 sosyo ekonomik değişken ile bölünmeli kümeleme yöntemlerinden bulanık kümeleme yöntemini kullanarak analiz etmiştir. Ayrıca, karşılaştırma yapmak amacıyla hiyerarşik olmayan kümeleme yöntemlerinden k-ortalımlar yöntemi de kullanmıştır. Bulanık kümeleme analizinin sonuçları, illeri iki farklı kümeye ayırmıştır. Elde ettiği sonuçları incelediğinde, bulanık kümeleme analizinin homojen birimlerin kümelenmesinde gücünün azaldığı, k-ortalımlar yönteminin ise önsel olarak belirlenen küme sayısına karşı duyarlı olduğunu söylemiştir.

Yılmaz vd (2011), yeni bir halı temizleyicisini pazarlayan bir işletmenin tüketici satın alma tercihleri üzerinde etkili 5 faktöre göre oluşturulan 22 profilden yararlanarak kümeleme analizi tekniği uygulanmışlardır. Bu amaçla hiyerarşik kümeleme analizi yöntemlerinden tek bağlantı yöntemini kullanılmışlardır.. Analizde mesafe olarak kullanılan karesel öklid ve pearson yakınlık ölçütlerine göre de birbirlerine en çok benzeyen profiller ve en az benzerlik gösteren profilleri saptamışlardır. Pazarlama alanında işletmelerin dikkate alması gereken en önemli ölçütlerin saptanmasında kümeleme analizi tekniğinin uygun bir yöntem olarak kullanılabileceğini söylemişlerdir.

Gevrekçi vd (2011), Türkiye İstatistik Kurumundan elde edilen Batı Anadolu'daki 11 ila ait koyunculuk verilerini kümeleme analizi ile analiz ederek illeri koyunculuk yapısına göre sınıflandırmıştır. Kümeleme analizinde uzaklık olarak öklid uzaklığını, kümeleme yöntemi olarak ise hiyerarşik kümeleme yöntemlerinden ward yöntemini kullanmıştır. Analiz sonucunda Batı Anadolu bölgesindeki illerin koyunculuk açısından 4 farklı kümede toplandığını tespit etmiştir.

Ada (2011), Türkiye'nin sürdürülebilir kalkınma açısından Avrupa Birliği üye ülkeleri karşısındaki konumu, kümeleme analizi yöntemleri çerçevesinde 15 sürdürülebilir kalkınma değişkenini kullanarak belirlemeye çalışmıştır. Hiyerarşik kümeleme yöntemlerinden ward yöntemini ve hiyerarşik olmayan kümeleme yöntemlerinden k-ortalama yöntemini kullanmıştır. Uzaklık matrisinin belirlenmesinde öklid uzaklığını kullanmıştır. Değişkenlere ait analiz sonuçlarına göre, Türkiye'nin AB ortalamasında bir sürdürülebilir kalkınma düzeyine sahip olduğu ve AB ülkelerinden büyük farklılıklar göstermediğini söylemiştir.

Banoğlu vd (2011), hiyerarşik kümeleme analizi yöntemlerinden ward yöntemi ve karesel öklid uzaklığı ile lise öğrencilerinin öğrenme ortamı algılarını inceleyerek, ortaya çıkan küme yapısına bağlı olarak liselerin öğrenme ortamı profilleri oluşturmuşlardır. Araştırma evreni olarak İstanbul ilindeki devlet liseleri belirlenmiş ve ilçelere göre tabakalama örnekleme alınarak, 12 ilçedeki 22 okuldan seçkisiz örneklemeyle 985 öğrenciye ulaşılmıştır. Kümeleme analizi sonucunda ortaya öğrenci algı kümeleri 4 tip öğrenme ortamı profilini belirlemişlerdir.

Sangün vd (2011), İskenderun körfezinde trol balıkçılığı ile yakalanan 68 türe hiyerarşik kümeleme yöntemleri uygulamış ve %90'lık benzerlik düzeyinde hemen hemen tüm yöntemlerde av kompozisyonunun çoğunluğunu oluşturan türlerin çimçim karidesi, lokum-zurna balığı, barbun, paşa barbunu ve yengeç olduğu belirlemiştir. Tüm yöntemler için oluşturulan ağaç grafiklerinde ward bağlantı kümeleme yöntemine ait ağaç grafiğinin en sade ve anlaşılır olduğunu söylemekte olup, bu tip çalışmalarda bu yöntemin uygun olacağını önermişlerdir.

Fırat vd (2012), Türkiye genelinde yeterli ölçüm uzunluğuna sahip 188 adet yağış gözlem istasyonuna ait verileri kullanarak yaptıkları sınıflandırma sonucunda küme sayısını 7 olarak belirlemişlerdir. Kümeleme analizinde hiyerarşik kümeleme yöntemlerinden ward yöntemi ile hiyerarşik olmayan kümeleme yöntemi k-ortalamlar yöntemi ile elde edilen sonuçlar karşılaştırmışlardır. Yıllık toplam yağışların sınıflandırılmasında k-ortalamlar yöntemi ile elde edilen sonuçların kabul edilebilir seviyede olduğunu söylemişlerdir.

Turhan vd (2012), Karadeniz Teknik Üniversitesi Tıp Fakültesi Farabi Hastanesi haziran 2011- haziran 2012 döneminde kan örnekleri tekrarlanmış olan 9777 yatan hastadan elde edilen son iki potasyum (K) ve hemoliz indeksi (HI) ölçümlerinden elde edilen verileri hiyerarşik olmayan kümeleme yöntemlerinden k- ortalama kümeleme yöntemi ve öklid uzaklık ölçüsünü kullanarak kümeleme yapmışlardır. Kümeleme yöntemi ile elde edilen sonuçların, mevcut verilerden daha kolay yararlanmayı sağladığını gözlemlemişlerdir. Kümeleme yönteminin, potasyum ve hemolizden etkilenen diğer testler için de düzeltme faktörü geliştirilmesinde ön çalışma olarak kullanılabileceğini söylemişlerdir.

Öztürk vd (2012), Şanlıurfa TİGEM Ceylanpınar Tarım İşletmesinden elde edilen 24 farklı erkek antep fıstığı tipi üzerine, kümeleme analizi yöntemleri uygulanarak en iyi küme yapısını ortaya koyan yöntem tespit etmeye çalışmışlardır. Çalışma sürecinde verileri tavsiye edilen herhangi bir kümeleme yöntemi ile değerlendirmenin çok zor olduğunu tespit etmişlerdir. Küme oluşumlarının veri setindeki değişkenlerin etkinliğine göre farklılık arz ettiğini ortaya koymuşlardır.

Dinler vd (2013), Doğu Anadolu Bölgesi illerindeki koyun yetiştiriciliğine ait verileri hiyerarşik kümeleme yöntemlerinden tek bağlantı, tam bağlantı ve ward bağlantı yöntemi ve uzaklık olarak öklid uzaklığını kullanarak analiz etmişlerdir. Tam ve ward bağlantı yöntemleri ile elde edilen kümelerin benzer olduğunu tespit etmişlerdir.

Yavuz vd (2013), biyobenzin üretimi için elverişli olan illerin belirlenebilmesi amacıyla il bazında Türkiye İstatistik Kurumu 2010 yılı genel tarım sayımı sonuçlarından 81 ile ait verim miktarları göz önünde bulundurularak elde edilen üretilen şeker pancarı, patates ve mısır bitkileri verileri Hiyerarşik kümeleme yöntemlerinden tek bağlantı ve tam bağlantı kümeleme yöntemi ile Hiyerarşik olmayan Kümeleme yöntemlerinden k-ortalama

yöntemi ile analiz etmişlerdir. K-ortalama yöntemi ile oluşturulan kümelerin diğer yöntemlere göre daha anlamlı kümeler oluşturduğunu gözlemlemişlerdir.

Çelik (2013), Türkiye İstatistik Kurumu internet sitesinden elde ettiği 2010 yılına ait, 81 ilimizin 10 farklı sağlık göstergesi verilerini, Hiyerarşik kümeleme analizi yöntemlerinden tek bağlantı kümeleme yöntemi, ward kümeleme yöntemi ve hiyerarşik olmayan kümeleme yöntemlerinden k- ortalama kümeleme yöntemleriyle analiz etmiştir. Analizde uzaklık ölçüsü olarak karesel öklid uzaklığını kullanmıştır. Yapılan kümeleme analizleri neticesinde illerin sağlık göstergelerine göre sınıflandırmasını yapmıştır. Sağlık hizmetleri yönünden illerin genel durumunu tespitini yapmış olup alınması gereken tedbirlerin neler olduğunu ortaya koymuştur.

### **3. KÜMELEME ANALİZİ**

Bu bölümde kümeleme analizinin tanımı yapılarak, benzerlik ve yakınlık ölçüleri, kümeleme yöntemleri ifade edilecek ve küme sayısının belirlenmesi yapılacaktır.

#### **3.1. Kümeleme Analizinin Tanımı**

Kümeleme analizi, bir araştırmada incelenen birimleri aralarındaki benzerliklerine göre belirli gruplar içinde toplayarak sınıflandırma yapmayı, birimlerin ortak özelliklerini ortaya koymayı ve bu sınıflar ile ilgili genel tanımlar yapmayı sağlar (Kaufman and Rousseeuw 1990).

Kümeleme analizi için bir başka tanım da şu biçimde yapılmaktadır. “ Kümeleme analizi, temel amacı nesnelere (birimleri) sahip oldukları olan çok değişkenli teknikler grubudur. Kümeleme analizi, nesnelere küme içerisinde çok benzer biçimde, kümeler arasında farklı olacak biçimde kümeler. Kümeleme işlemi başarılı olursa, bir geometrik çizim yapıldığında nesnelere küme içerisinde birbirine çok yakın, kümeler ise birbirinden uzak olacaktır (Hair et al. 1995).

Kümeleme analizinin genel amacı, gruplanmamış verileri benzerliklerine göre gruplamak ve araştırmacıya uygun, işe yarar özetleyici bilgiler elde etmede yardımcı olmaktır. Bireylerin sınıflanması, ait oldukları grupların belirlenmesi ile uğraşan, çok değişkenli analizlerle sıkı ilişkisi olan bir istatistiksel analiz yöntemidir (Tatlıdil 1996).

Kümeleme analizi ilk kez 1939 yılında Tryon tarafından kullanılmıştır. 1960'lı yıllardan sonra kullanımı yaygınlaşmıştır. 1963 yılında Robert Sokal ve Peter Sneath'in yazdığı “Sayısal Sınıflandırma İlminin Temelleri” adlı kitap bu alanda önemli bir adım olmuştur (Anderberg 1973).

Kümeleme analizi, temel olarak dört değişik amaca yönelik işlev yerine getirir.

1. n sayıda birimi, nesneyi, oluşumu p değişkene göre saptanan özelliklerine göre olabildiğince kendi içinde türdeş ve kendi aralarında farklı alt gruplara ayırmak,
2. p sayıda değişkeni, n sayıda birimde saptanan değerlere göre ortak özellikleri açıkladığı varsayılan alt kümelere ayırmak ve ortak faktör yapıları ortaya koymak
3. Hem birimleri hem de değişkenleri birlikte ele alarak ortak n birimi p değişkene göre ortak özellikli alt kümelere ayırmak,
4. Birimleri, p değişkene göre saptanan değerlere göre, izledikleri biyolojik ve tipolojik sınıflamayı ortaya koymak (Özdamar 2004).

Kümeleme analizinin uygulama aşamaları aşağıdaki gibi verilebilir.

Veri Matrisinin Belirlenmesi (Adım 1): Birim ya da değişkenlerin doğal gruplamaları hakkında kesin bilgilerin bulunmadığı popülasyonlardan alınan n sayıda birimin p sayıda değişkenine ilişkin gözlemlerin elde edilerek tablo 3.1 şeklinde matris haline getirilir.

Tablo 3.1. Veri Matrisi

	1.Değişken	2. Değişken	3. Değişken	....	p. Değişken
1.Birim	$X_{11}$	$X_{12}$	$X_{13}$	....	$X_{1p}$
2.Birim	$X_{21}$	$X_{22}$	$X_{23}$	....	$X_{2p}$
3.Birim	$X_{31}$	$X_{32}$	$X_{33}$	....	$X_{3p}$
				....	
n.Birim	$X_{n1}$	$X_{n2}$	$X_{n3}$	....	$X_{np}$

Benzerlik ya da Farklılık Matrisinin Belirlenmesi (Adım 2): Birimlerin/değişkenlerin birbirleri ile olan benzerliklerini ya da farklılıklarını gösteren uygun bir benzerlik ölçüsü ile birimlerin/değişkenlerin birbirlerine uzaklıkları hesaplanır ve bu uzaklıklara göre tablo 3.2 şeklinde benzerlik/farklılık matrisi oluşturulur. Çok sık kullanılan uzaklık ölçüleri öklid, karesel öklid, pearson, karesel pearson, minkowski, manhattan (city-block), mahalnobis, hotelling  $t^2$  ve Canberra uzaklığı şeklindedir.



Tablo 3.2. Benzerlik/Farklılık Matrisi

	1. Birim	2. Birim	3. Birim	....	n. Birim
1. Birim	0	$d_{12}$	$d_{13}$	....	$d_{1n}$
2. Birim	SİMETRİ	0	$d_{23}$	....	$d_{2n}$
3. Birim	SİMETRİ	SİMETRİ	0	....	....
...	SİMETRİ	SİMETRİ	SİMETRİ	0	$d_{n-1,n}$
n. Birim	SİMETRİ	SİMETRİ	SİMETRİ	SİMETRİ	0

Kümelerin Oluşturulması (Adım 3): Uygun küme yöntemi yardımı ile benzerlik/farklılık matrisine göre birimlerin/değişkenlerin uygun sayıda kümelere ayrılması sağlanır. Gruplara ayırırken, en sık kullanılan kümeleme yöntemlerinden hiyerarşik ve hiyerarşik olmayan kümeleme yöntemlerinden uygun olan yöntemler tercih edilir. Birimler birbirlerine olan uzaklık veya yakınlıklarına göre uygun sayıda kümeye ayrılır.

Analiz ve Sonuç (Adım 4): Elde edilen kümelerin yorumlanması ve bu kümeleme yapısına dayalı olarak kurulan hipotezlerin doğrulanması için gerekli analitik yöntemlerin uygulanarak, sonuçların duyarlılığının ve anlamlılığının tartışması yapılır. Sonuçların uygun olmaması durumunda (değişkenlerin uygun olmaması ve/veya küme sayısının uygun belirlenmemiş olması nedeniyle) değişkenler, küme sayısı, uzaklık ölçüleri ve kümeleme yöntemleri gibi parametreler gözden geçirilerek süreç başından tekrarlanır (Tatlıdil 1996).

Kümeleme analizi çok sayıda değişik işlevi yerine getiren yöntemler topluluğudur. Bu nedenle farklı amaçlar için farklı yöntemler uygulanır. Ayrıca değişkenlerin ölçü birimlerinin ve ölçüleme tekniklerinin farklı olmasından dolayı birimlerinin benzerliklerinin ortaya konmasında da değişik ölçüler kullanılır.

### 3.2. Kümeleme Analizinde Değişken Türlerine Göre Benzerlik ve Yakınlık Ölçüleri

Bir veri setinde yer alan birimlerin kümelmesi işlemi bu birimlerin birbirleriyle olan benzerlikleri ya da birbirlerine olan uzaklıkları kullanılarak gerçekleştirilmektedir.

Değişkenlerin kesikli ya da sürekli olmalarına ya da değişkenlerin nominal, ordinal, aralık ya da oransal ölçekte olmalarına göre hangi uzaklık ölçüsünün ya da hangi benzerlik ölçüsünün kullanılacağına karar verilir. Uzaklık fonksiyonunun genel özelliklerini şu şekilde sıralayabiliriz.

Uzaklık Fonksiyonun Özellikleri;

- $d(i, j) \geq 0$  ; Uzaklık negatif değil
- $d(i, i) = 0$  ; Her birim kendisine olan uzaklığı sıfırlar.
- $d(i, j) = d(j, i)$  ; Uzaklık fonksiyonu simetriktir.
- $d(i, j) \leq d(i, h) + d(h, j)$  ; iki birimin arasındaki uzaklık bu iki birimin üçüncü bir birime olan uzaklıkları toplamından küçük olamaz (üçgen eşitsizliği)

### 3.2.1. Öklid (Euclidean) Uzaklığı

Öklid uzaklığı formülleri standartlaştırılmış verilerle değil, işlenmemiş verilerle hesaplama yapılır. Öklid uzaklıkları kümeleme analizine sıra dışı olabilecek yeni nesnelerin eklenmesinden etkilenmezler. Ancak boyutlar arasındaki ölçek farklılıkları Öklid uzaklıklarını önemli ölçüde etkilemektedir. Öklid uzaklık formülü en yaygın olarak kullanılan uzaklık hesaplama formülüdür (Demiralay 2005).

Öklid uzaklık ölçüsü kullanılarak iki birim arasındaki uzaklık  $n$  birim sayısı  $v$  e  $p$  değişken sayısı olmak üzere;  $i, j = 1, 2, 3, \dots, n$  ,  $i$ . ve  $j$ . birimin birbirine olan uzaklığı

$$d(i, j) = \sqrt{(x_{i1} - x_{j1})^2 + (x_{i2} - x_{j2})^2 + \dots + (x_{ip} - x_{jp})^2} \quad (3.1)$$

formülü ile hesaplanır.

### 3.2.2. Karesel Öklid (Squared Euclidean) Uzaklığı

Birimlere ait aynı değişkenler arası farkların karelerinin toplanması ile hesaplanır. Uzaklık,

$$d(i, j) = (x_{i1} - x_{j1})^2 + (x_{i2} - x_{j2})^2 + \dots + (x_{ip} - x_{jp})^2 \quad (3.2)$$

şeklinde hesaplanır.

### 3.2.3. Manhattan ( City-Block) Uzaklığı

Bu ölçü de birimler arasındaki mutlak uzaklık kullanılır. Manhattan uzaklığı birimlerin aynı değişkenleri arasındaki mutlak farkların toplanması,

$$d(i, j) = (|x_{i1} - x_{j1}| + |x_{i2} - x_{j2}| + \dots + |x_{ip} - x_{jp}|) \quad (3.3)$$

formülü ile hesaplanır. City-block uzaklık ölçüsü uygulamada bazı sorunlara yol açmaktadır. Bu sorunlardan en belirginini city-block uzaklık ölçüsünün değişkenler arasında ilişki olmadığını varsaymaktadır. Eğer araştırma konusunda değişkenler arasında korelasyon varsa city-block uzaklık ölçüsüyle hesaplanan uzaklık ölçüleri baz alınarak yapılan kümeleme anlamlı olmayacaktır. Sorunlardan bir diğeri de ölçüm yapılan değişkenlerin birimleri farklı olması durumunda standartlaştırılmış karasel öklid uzaklığıyla karşılaştırıldığında City-block uzaklık ölçüsünün anlamlı sonuçlar vermediği görülebilmektedir (Atbaş 2008).

### 3.2.4. Pearson Uzaklığı

Pearson uzaklık ölçüsü kullanılarak iki birim arasındaki uzaklık;

$$d(i, j) = \sqrt{\frac{(x_{i1} - x_{j1})^2}{S_1^2} + \frac{(x_{i2} - x_{j2})^2}{S_2^2} + \dots + \frac{(x_{ip} - x_{jp})^2}{S_p^2}} \quad (3.4)$$

formülü ile hesaplanır. Bu formülde kullanılan  $S_p$ , uzaklığın hesaplandığı değişkene ait varyanstır. Bununla birlikte farklı gruplar hakkında önceden bilgi sahibi olunmadığı için, uzaklık hesaplanmasında  $S$  değerinin kullanılması doğru olmaz. Bu nedenle Pearson uzaklık ölçüsü yerine genellikle Öklid uzaklık ölçüsü tercih edilir (Anonim 2014).

### 3.2.5. Mahalanobis Uzaklığı

Doğrudan birleştirme yapan, standart bir yöntem olan Mahalanobis Uzaklık ölçüsüdür. İki değişken arasında bir ilişki mevcut ise, bu iki değişken arasındaki kovaryans veya korelasyonu göz önüne alan Mahalanobis uzaklığının kullanılması gerekmektedir. p değişkenli bir analizde i ve j gözlemleri arasındaki Mahalanobis uzaklık ölçüsü;

$$d(i, j) = D^2 = (x_i - x_j)' S^{-1} (x_i - x_j) \quad (3.5)$$

Formülü ile hesaplanmaktadır. S, p×p kovaryans matrisini göstermektedir. Mahalanobis uzaklığının avantajı, aykırı noktaları da hesaplamasıdır. Bu yönleriyle Mahalanobis uzaklığı, uzaklık ölçüleri arasında en avantajlı olanıdır denilebilir (Sharma 1996).

### 3.2.6. Minkowski Uzaklığı

Minkowski uzaklık ölçüsü genel bir formüldür. Formülde yer alan m değerinin alacağı farklı değerlere göre yeni formüller türetir. Minkowski uzaklık ölçüsü kullanılarak iki birim arasındaki uzaklık

$$d(i, j) = \left[ |x_{i1} - x_{j1}|^m + |x_{i2} - x_{j2}|^m + \dots + |x_{ip} - x_{jp}|^m \right]^{1/m} \quad (3.6)$$

formülü ile hesaplanır. Minkowski uzaklık ölçüsündeki m değeri büyük ve küçük farklara verilen ağırlığı değiştirir. m=1 değerini alırsa, formül, Manhattan uzaklık ölçüsünün formülüne, m = 2 değerini alırsak, formül Öklid uzaklık ölçüsü formülüne dönüşür (Anderberg 1973).

### 3.2.7. Karesel Pearson Uzaklığı

$$d(i, j) = \frac{(x_{i1} - x_{j1})^2}{S_1^2} + \frac{(x_{i2} - x_{j2})^2}{S_2^2} + \dots + \frac{(x_{ip} - x_{jp})^2}{S_p^2} \quad (3.7)$$

formülü ile hesaplanır.

### 3.2.8. Hotelling T<sup>2</sup> uzaklığı

İki birimin ortalama vektörlerinin karşılaştırılmasında kullanılır.

$$d(i, j) = T^2 = \frac{n_1 \cdot n_2}{n} (\bar{x}_i - \bar{x}_j)' S^{-1} (\bar{x}_i - \bar{x}_j) \quad (3.8)$$

Formülü ile hesaplanmaktadır. S , p×p kovaryans matrisini göstermektedir .

### 3.2.9. Canberra Ölçütü

$$d(i, j) = \frac{|x_{i1} - x_{j1}| + |x_{i2} - x_{j2}| + \dots + |x_{ip} - x_{jp}|}{(x_{i1} + x_{j1}) + (x_{i2} + x_{j2}) + \dots + (x_{ip} + x_{jp})} \quad (3.9)$$

formülü ile hesaplanır.

## 3.3. Kümeleme Yöntemleri

Kümeleme analizinde uzaklık/yakınlık ölçüleri kullanılarak oluşturulan uzaklık/yakınlık matrisindeki değerlerinden faydalanılarak birimlerin kümelere (gruplara) atanması yapılır.

Araştırmacı hangi benzerlik/uzaklık ölçüsünü kullanacağına karar verdikten sonra, kümeleme işleminin nasıl olacağına karar vermek zorundadır. Birimlerin benzerliklerine göre kümelere dâhil edilmesinde kullanılabilir çeşitli yaklaşımlar vardır. Bu yaklaşımlardan biri, en çok benzer iki birimi aynı gruba atamakla başlayıp tüm birimlerin aynı gruba atanması ile biten hiyerarşik bir yaklaşımdır. Bir başka yaklaşım ise tüm verilerin ortalama değerlerine en yakın değerlere sahip birimlerin aynı kümeye atanmasını esas alan yaklaşımdır. En çok kullanılan bu iki yaklaşım dışında diğer yaklaşımlar da mevcuttur. Tüm yaklaşımlarda en önemli ölçüt, kümeler arası farklar ile kümeler içi benzerliklerin maksimum olmasını sağlamaktır. En çok kullanılan kümeleme algoritmaları hiyerarşik ve hiyerarşik olmayan kümeleme adı altında iki kategoride toplanmaktadır (Blashfield and Aldenderfer 1978).

### 3.3.1. Hiyerarşik Kümeleme Yöntemleri

Hiyerarşik kümeleme yöntemleri, veri setinin birimlerinin birbirlerine olan uzaklık değerlerini kullanarak, veri setindeki birimlerin hiyerarşik ayrıştırmasını yapar. Hiyerarşik ayrıştırma sırasında, dendogram olarak bilinen ağaç diyagramı kullanılır. Ağaç diyagramı, hiyerarşik kümeleme yöntemiyle elde edilen kümelerin görselleştirilmesini sağlar. Küme sayısına görsel olarak karar verilir. Gruplayıcı ve bölücü olmak üzere iki yöntem mevcuttur.

Gruplayıcı hiyerarşik yöntemde her birim veya her gözlem başlangıçta bir küme olarak kabul edilir. Daha sonra en yakın iki küme (veya gözlem) yeni bir kümede toplanarak birleştirilir. Böylece her adımda küme sayısı bir azaltılır. Bölücü hiyerarşik yöntemde ise süreç gruplayıcı hiyerarşik yöntemin tam tersidir. Bu yöntemde tüm gözlemlerden oluşan büyük bir küme ile ise başlanır. Benzer olmayan gözlemler ayıklanarak daha küçük kümeler oluşturulur. Her gözlem tek başına küme oluşturana kadar işleme devam edilir (Everitt et al. 2001).

Hiyerarşik kümeleme yöntemi dört adımdan oluşan bir algoritma ile ifade edilebilir. Bunlar,

1.  $n$  tane birey,  $n$  tane küme olmak üzere işleme başlanır.
2. En yakın iki küme ( $d_{ij}$  değeri en küçük olan alınır) birleştirilir.
3. Küme sayısı bir indirgenerek yinelenmiş uzaklıklar matrisi bulunur.
4. 2 ve 3 nolu adımlar  $n-1$  kez tekrarlanır (Tatlıdil 1996).

Analizlerde birçok kümeleme yöntemi denemek sonuçları karşılaştırmak için fayda sağlayabilir. Verilerin özelliklerine bağlı olarak, bazı kümeleme yöntemi diğerlerine göre daha uygun kümeler oluşturabilir. En çok kullanılan 7 hiyerarşik kümeleme yöntemleri şunlardır;

1. Tek Bağlantı Kümeleme Yöntemi (Single Linkage)
2. Tam Bağlantı Kümeleme Yöntemi (Complete Linkage)
3. Ortalama Bağlantı Kümeleme Yöntemi (Average Linkage)
4. McQuitty Bağlantı Kümeleme Yöntemi (McQuitty Linkage)

5. Küresel (Merkezi) Bağlantı Kümeleme Yöntemi (Centroid Linkage)
6. Medyan Bağlantı Kümeleme Yöntemi (Median Linkage)
7. Ward Bağlantı Kümeleme Yöntemi (Ward Linkage)

### 3.3.1.1. Tek Bağlantı Kümeleme Yöntemi (Single Linkage)

En yalın hiyerarşik kümeleme yöntemidir. Aynı zamanda en yakın komşuluk yöntemi olarak ta bilinir. Bu yöntemde uzaklıklar matrisinden faydalanılarak, birbirine en yakın birim veya kümeler birleştirilir, birleştirmelere bütün birimler herhangi bir kümeye atanıncaya kadar devam edilir. Birleştirme yapılırken kümelerin eleman sayısının birden fazla olması koşulu yoktur. Bir birim yalnız başına bir küme oluşturabilir. İki küme arasındaki uzaklık, bir kümedeki bir gözlem ve diğer kümedeki bir gözlem arasındaki minimum mesafedir.

Bu yöntemde eğer i. ve j. birimler birleştirilmiş ise birleştirilen kümenin k. küme ile ilişkisi uzaklık ölçütü olarak,

$$d_{k(i,j)} = \min(d_{ki}, d_{kj}) \quad (3.10)$$

biçiminde ifade edilmektedir.

Eşitlikte;

$d_{k(i,j)}$ , k. kümenin i. ve j. kümelerle olan uzaklığını,

$d_{kj}$ , k. kümenin j. kümeye olan uzaklığını,

$d_{ki}$ , k. kümenin i. küme ile olan uzaklığını göstermektedir.

### 3.3.1.2. Tam Bağlantı Kümeleme Yöntemi (Complete Linkage)

Tam bağlantı kümeleme yöntemine en uzak komşuluk yöntemi de denir. Tam bağlantı yöntemi, iki küme arasındaki uzaklık bir kümedeki bir gözlem ve diğer kümedeki bir gözlem arasındaki maksimum mesafedir. Bu yöntem, bir kümedeki tüm gözlemlerin

maksimum mesafe içinde olmasını sağlar ve benzeri kümeleri üretir. Tek bağlantı yöntemi ile benzerlik göstermekte olup, minimum yerine maksimum uzaklığı baz almaktadır. Küçük ve yoğun küme yapısına sahip gözlemlerde etkili bir yöntem olarak benimsenmiştir. Bu yöntem;

$$d_{k(i,j)} = \max(d_{ki}, d_{kj}) \quad (3.11)$$

Formülüyle hesap edilmektedir. Formülde;

$d_{k(i,j)}$ , k. kümenin i. ve j. kümelerle olan uzaklığını,

$d_{kj}$ , k. kümenin j. kümeyle olan uzaklığını,

$d_{ki}$ , k. kümenin i. küme ile olan uzaklığını göstermektedir.

### 3.3.1.3. Ortalama Bağlantı Kümeleme Yöntemi (Average Linkage)

Bu yöntemde işleme tek bağlantı ve tam bağlantı yöntemlerinde olduğu gibi başlanır. Ancak kümeleme kriteri olarak bir küme içindeki birim ile diğer küme içindeki birimler arasındaki ortalama uzaklıklar kullanılır. Ortalama bağlantı tekniğinde kümeler küçük varyanslar ile birbirlerine bağlıdır. Bu teknik tek bağlantı ve tam bağlantı teknikleri arasında sonuçlar vermesi nedeniyle bir alternatif yöntem olarak önerilmektedir (Hubert 1974).

Tek bağlantılı kümeleme yönteminde işlemlerin uzun sürmesi, tam bağlantılı kümeleme yönteminde ise, aynı küme içersindeki bireylerin uzaklıklarının belli bir değerden küçük olması durumunda tüm kümelerin sağlıklı oluşturulmasının garanti edilememesi, son yıllarda sıkça kullanılan ortalama bağlantı kümeleme yönteminin alternatif olarak önerilmesine sebep olmuştur. Burada iki küme arası mesafe, her biri bir gruptan olacak olan tüm nesne çiftleri arasındaki ortalama mesafedir. Ortalama bağlantılı kümeleme metodunda, aşağıdaki gibi hesaplanır.

$$d_{(k,l)j} = (N_k d_{(k,j)} + N_l d_{(l,j)}) / N_k + N_l \quad (3.12)$$



Burada;

$d_{(k,l)j}$  : k ve l'inci kümenin j'inci kümenin küme ile olan uzaklığını,

$d_{(k,j)}$  : k'inci kümenin j'inci küme olan uzaklığı,

$d_{(l,j)}$  : l'inci kümenin j'inci küme olan uzaklığını göstermektedir.

$N_k$ : k'inci kümedeki toplam birey sayısını,

$N_l$ : l'inci kümedeki toplam birey sayısıdır (Özdamar 2004).

#### 3.3.1.4. McQuitty Bağlantı Kümeleme Yöntemi (McQuitty Linkage)

McQuitty bağlantı kümeleme yöntemi ile iki kümenin birleştirilmesi esnasında bir kümenin diğer kümeye mesafesi yakında birleştirilecek kümelerin mesafelerinin ortalaması olarak hesaplanır. Ağırlıksız ortalama bağlantı kümeleme yöntemi olarak da bilinmektedir. j'nin k ve l kümeleri arasındaki uzaklığı aşağıda verilen eşitlik yardımı ile bulunabilir (Yazgan ve Kayaalp 2002).

$$d_{(k,l)j} = (d_{(k,j)} + d_{(l,j)}) / 2 \quad (3.13)$$

formülüyle hesap edilmektedir. Formülde;

$d_{(k,l)j}$ , k ve l. kümenin j. küme ile olan uzaklığını,

$d_{(k,j)}$ , k. kümenin j. kümeye olan uzaklığını,

$d_{(l,j)}$ , l. kümenin j. küme ile olan uzaklığını göstermektedir.

#### 3.3.1.5. Küresel (Merkezi) Bağlantı Kümeleme Yöntemi (Centroid Linkage)

Merkezi Bağlantı Kümeleme Yöntemi iki küme merkezi veya ortalaması arasındaki mesafeyi baz alarak kümeleme yapar. Ortalama bağlantı yöntemi ile benzerlik gösterir, fakat ona göre daha ortalama bir yöntemdir.

Bu yöntem, ortalama bağlantı kümeleme yönteminin özel bir biçimi olup kümeler arası uzaklıklar ve küme merkezleri arası uzaklıklar olarak tanımlanmaktadır. Kümelerin birleştirilmesi küme merkezleri arasındaki uzaklığa göre yapıldığından kümeler

merkezleri ile ifade edilmekte  $j$ 'nin  $k$  ve  $l$  kümeleri arasındaki uzaklıktır (Renchber, 2002).

Merkezi Ortalama bağlantılı kümeleme yöntemi;

$$d_{(k,l)j} = \frac{(N_k d_{(k,j)} + N_l d_{(l,j)})}{N_k + N_l} - \frac{N_k N_l d_{(k,l)}}{N_k^2 + N_l^2} \quad (3.14)$$

$d_{(k,l)j}$  :  $k$  ve  $l$ 'inci kümenin  $j$ 'inci kümenin küme ile olan uzaklığını,

$d_{(k,j)}$  :  $k$ 'inci kümenin  $j$ 'inci küme olan uzaklığını,

$d_{(l,j)}$  :  $l$ 'inci kümenin  $j$ 'inci küme olan uzaklığını,

$d_{(k,l)}$  :  $k$ 'inci kümenin  $l$ 'inci küme olan uzaklığını,

$N_k$ :  $k$ 'inci kümedeki toplam birey sayısını,

$N_l$ :  $l$ 'inci kümedeki toplam birey sayısını göstermektedir.

### 3.3.1.6. Medyan Bağlantı Kümeleme Yöntemi (Median Linkage)

Medyan bağlantı kümeleme yöntemi, iki kümenin merkezleri arasındaki uzaklığın eşit ağırlıklı olarak hesaplanmasıyla elde edilir (Gower 1967). İki küme arasındaki uzaklık hesaplanırken,

$$d_{(k,l)j} = \frac{(d_{(k,j)} + d_{(l,j)})}{2} - \frac{d_{(k,l)}}{4} \quad (3.15)$$

formülü kullanılır. Burada,

$d_{(k,l)j}$ ,  $k$  ve  $l$ . kümenin  $j$ . küme ile olan uzaklığını,

$d_{(k,j)}$ ,  $k$ . kümenin  $j$ . küme ile olan uzaklığını,

$d_{(l,j)}$ ,  $l$ . kümenin  $j$ . küme ile olan uzaklığını,

$d_{(k,l)}$ ,  $k$ . kümenin  $l$ . küme ile olan uzaklığını ifade etmektedir.

### 3.3.1.7. Ward Bağlantı Kümeleme Yöntemi (Ward Linkage)

Ward Bağlantı kümeleme yöntemi, iki küme arasındaki uzaklığın hesaplanmasında merkezden sapmaları yani varyansı esas alır. 1963 yılında Joe Henry Ward tarafından önerilmiştir. Bu yöntem en küçük varyans yöntemi de denilmektedir. Ward bağlantı kümeleme yönteminde amaç kümeler içi kareler toplamını minimize etmektir. Benzer eleman sayısına sahip küme elde etme eğilimindedir, uç değerlere karşı duyarlıdır. Bu yöntemde j'nin k ve l kümeleri arasındaki uzaklığı aşağıda verilen eşitlik yardımı ile bulunabilir (Özdamar 2004).

$$d_{(k,l)j} = \frac{(N_j + N_k)d_{(k,j)} + (N_j + N_l)d_{(l,j)} + N_j d_{(k,l)}}{N_j + N_k + N_l} \quad (3.16)$$

$d_{(k,l)j}$  : k ve l'inci kümenin j'inci kümenin küme ile olan uzaklığını,

$d_{(k,j)}$  : k'inci kümenin j'inci küme olan uzaklığını,

$d_{(l,j)}$  : l'inci kümenin j'inci küme olan uzaklığını,

$d_{(k,l)}$  : k'inci kümenin l'inci küme olan uzaklığını,

$N_k$ : k'inci kümedeki toplam birey sayısını,

$N_l$ : l'inci kümedeki toplam birey sayısını,

$N_j$ : j'inci kümedeki toplam birey sayısını göstermektedir.

### 3.3.2. Hiyerarşik Olmayan Kümeleme Yöntemleri

Bazı durumlarda küme sayısı önceden bellidir ve araştırmacı bu küme sayısına göre çözümler üretmek durumundadır. Küme sayısı konusunda ön bilgi varsa veya araştırmacı anlamlı olacak küme sayısına karar vermiş ise bu durumda, çok uzun zaman alan hiyerarşik yöntemler yerine hiyerarşik olmayan yöntemler kullanılmaktadır (Anderberg 1973).

Hiyerarşik olmayan kümeleme yöntemleri n adet birimden oluşan veri setini başlangıçta belirlenen  $k \leq n$  olmak üzere k adet kümeye ayırmak için kullanılır. Hiyerarşik olmayan kümeleme yöntemlerinin, hiyerarşik kümeleme yöntemlerinden en önemli farkından

birisi de budur. Hiyerarşik olmayan kümeleme yöntemleri, hiyerarşik kümeleme yöntemlerine oranla daha büyük veri setlerine uygulanabilir. Hiyerarşik olmayan kümeleme yöntemleriyle oluşturulacak  $k$  adet kümede her bir küme en azından bir birim içerir ve her birim yalnızca bir grupta bulunur.

Hiyerarşik kümeleme yönteminde, ilk olarak başlangıç küme merkezleri gelişi güzel olarak seçilir. Birimlerin, belirlenen kümelerin merkezlerine olan uzaklıklarına göre yeni küme merkezleri oluşturulur. Bu işlemler birbirilerinden farklı, kendi içlerinde homojen, birbirileri arasında benzerlik bulunmayan  $k$  adet küme oluşturuluncaya kadar sürdürülür. Hiyerarşik kümeleme yöntemleri arasında en bilinenleri K-ortalamlar kümeleme yöntemidir.

### **3.3.2.1. k- Ortalamalar Kümeleme Yöntemi**

K-ortalamlar yönteminde küme merkezleri oluşturulurken her bir tekrarlama oluşan kümeler için değişkenlerin ortalamaları alınır. İkinci önemli koşul ise, oluşturulacak olan küme sayısının başlangıçta biliniyor olmasıdır.

Amaç diğer kümeleme yöntemlerinde olduğu gibi, gerçekleştirilen kümeleme işlemi sonucunda elde edilen kümelerin, küme içi benzerliklerinin maksimum, kümeler arası benzerliklerinin ise minimum olmasını sağlamaktır. Küme benzerliği, kümenin ağırlık merkezi kabul edilen bir birim ile kümedeki diğer birimler arasındaki uzaklıkların ortalama değeri ile ölçülmektedir (Han and Kamber 2006). K- ortalamlar yönteminde yapılacak işlemler aşağıdaki gibidir.

1.  $k$  adet birim başlangıç küme merkezleri olarak rastgele seçilir.
2. Küme merkezi olmayan birimler belirlenen uzaklık ölçütlerine başlangıç küme merkezlerinin ait oldukları kümelere atanır.
3. Yeni küme merkezleri oluşturulan  $k$  adet başlangıç kümesindeki değişkenlerin ortalamaları alınarak oluşturulur.
4. Birimler en yakın oldukları oluşturulan yeni küme merkezlerine birimlerin uzaklıkları hesaplanarak kümeye atanır.

5. Bir önceki küme merkezlerine olan uzaklıklar ile yeni oluşturulan küme merkezlerine olan uzaklıklar karşılaştırılır.
6. Uzaklıklar makul görülebilir oranda azalmış ise 4. adıma dönülür.
7. Eğer çok büyük bir değişiklik söz konusu olmamış ise, tekrarlama sona erdirilir.

Bu yöntemde bireyler, gruplar içi kareler toplamını en küçük yapacak şekilde k kümeye bölünür  $x_1, x_2, \dots, x_n$  değişkenlerinin her biri p değişkenli gözlem vektörleri, çok boyutlu X uzayında birer nokta ifade ederken, aynı uzayda  $a_{1n}, a_{2n}, \dots, a_{kn}$  her grup birey için küme merkezleri olarak belirlendiğinde, aşağıdaki formüle göre bireyler en küçük uzaklığı veren (en yakın) kümeye sınıflanmaktadır (Tatlidil 1996).

$$W_N = \frac{1}{n} \sum_{i=1}^n \min \|x_i - a_{in}\|^2 \quad (3.17)$$

Formülü ile hesaplanır. Bu yöntem uç verilerden etkilenir.

### 3.4. Küme Sayısının Belirlenmesi

Kümeleme analizinin en kritik konusu küme sayısına karar vermektir. Araştırmacının küme sayısına karar vermede öznelliği minimize etmesi gerekmektedir. Ancak günümüzde yayınlanan birçok makalede bu konuda kesin bulunmuş sonuçlar yoktur. İlk önerilen yaklaşımlardan en çok bilinen eşitlik,

$$k = \sqrt{\frac{n}{2}} \quad (3.18)$$

Biçiminde hesaplanmaktadır. Burada k küme sayısı, n birim sayısını göstermektedir. Küçük örneklemlerli araştırmalarda kullanılması tavsiye edilir. Büyük örneklemlerli araştırmalarda kullanılması durumunda sağlıklı sonuçlara ulaşılması zorlaşır (Everitt 1974).

Diğer bir yöntem ise Mariott tarafından önerilmiştir. Bu yüzden M harfi ile gösterilir.

$$M = k^2|W| \quad (3.19)$$

biçimindedir (Marriot 1971). Burada W, grup içi kareler toplamı matrisidir ve

$$W = \sum_{j=1}^k \sum_{i=1}^{n_j} (x_{ij} - \bar{x}_j)(x_{ij} - \bar{x}_j)' \quad (3.20)$$

$n_j$ ; j. kümedeki birim sayısı

k ; küme sayısı

$x_{ij}$ ; j. kümedeki i. birim değerleri

$\bar{x}_j$  ; j. kümenin örneklem ortalama vektörü

biçiminde hesaplanır.

Küme sayısına karar vermek için Calinski ve Harabasz , Lewis ve Thomas, Wilk's Lamda ölçütü gibi küme sayısını belirlemeye yönelik bir çok yöntem mevcuttur. Araştırmacı oluşan küme yapılarını incelemeli, uygun sayıda küme sayısına elde ettiği sonuçlara göre karar vermelidir.

## 4. KÜMELEME ANALİZİ YÖNTEMLERİNİN HAYVANCILIK VERİLERİNDE KARŞILAŞTIRILMALI OLARAK İNCELENMESİ

Bu bölümde Türkiye'deki 81 il, hayvancılık yönünden hiyerarşik kümeleme analizi yöntemleriyle incelenecektir. Tüm kümeleme yöntemleri ve uzaklık ölçüleri karşılaştırılarak incelenecektir.

### 4.1. Verilerin Özellikleri

Araştırmada, Türkiye'de 81 ilimize ait hayvancılık verileri kullanılmıştır. Veriler Türkiye İstatistik Kurumu (TUIK) internet sitesinde hayvancılık istatistikleri bölümünden dinamik sorgulama ile elde edilmiştir. Veriler 2012 yılına ait Büyükbaş Hayvancılık, Küçükbaş Hayvancılık, Kümes Hayvancılığı ve Arıcılık istatistiklerinden oluşmaktadır.

Analizde kullanılan değişkenler,

- $X_1$ : Yetişkin yerli sığır sayısı,
- $X_2$ : Genç-Yavru yerli sığır sayısı,
- $X_3$ : Sağılan yerli sığır sayısı,
- $X_4$ : Yerli sığırdan elde edilen yıllık süt miktarı,
- $X_5$ : Yetişkin kültür ırkı sığır sayısı,
- $X_6$ : Genç-Yavru kültür ırkı sığır sayısı,
- $X_7$ : Sağılan kültür ırkı sığır sayısı,
- $X_8$ : Kültür ırkı sığırdan elde edilen yıllık süt miktarı,
- $X_9$ : Yetişkin melez sığır sayısı,
- $X_{10}$ : Genç-Yavru melez sığır sayısı,
- $X_{11}$ : Sağılan melez sığır sayısı,
- $X_{12}$ : Melez sığırdan elde edilen yıllık süt miktarı,
- $X_{13}$ : Yetişkin manda sayısı,
- $X_{14}$ : Genç-Yavru manda sayısı,

- X<sub>15</sub>: Sağılan manda sayısı,  
 X<sub>16</sub>: Mandadan elde edilen yıllık süt miktarı,  
 X<sub>17</sub>: Yerli koyun sayısı,  
 X<sub>18</sub>: Genç-Yavru yerli koyun sayısı,  
 X<sub>19</sub>: Sağılan yerli koyun sayısı,  
 X<sub>20</sub>: Yerli koyundan elde edilen yıllık süt miktarı,  
 X<sub>21</sub>: Yerli koyundan elde edilen yıllık yapağı miktarı,  
 X<sub>22</sub>: Merinos koyun sayısı,  
 X<sub>23</sub>: Genç-Yavru merinos koyun sayısı,  
 X<sub>24</sub>: Sağılan merinos koyun sayısı,  
 X<sub>25</sub>: Merinos koyundan elde edilen yıllık süt miktarı,  
 X<sub>26</sub>: Merinos koyundan elde edilen yıllık yapağı miktarı,  
 X<sub>27</sub>: Kıl keçisi sayısı,  
 X<sub>28</sub>: Genç-Yavru kıl keçisi sayısı,  
 X<sub>29</sub>: Sağılan kıl keçisi sayısı,  
 X<sub>30</sub>: Kıl keçisinden elde edilen yıllık süt miktarı,  
 X<sub>31</sub>: Kıl keçisinden elde edilen yıllık tiftik miktarı,  
 X<sub>32</sub>: Tiftik keçisi sayısı,  
 X<sub>33</sub>: Genç-Yavru tiftik keçisi sayısı,  
 X<sub>34</sub>: Sağılan tiftik keçisi sayısı,  
 X<sub>35</sub>: Tiftik keçisinden elde edilen yıllık süt miktarı,  
 X<sub>36</sub>: Tiftik keçisinden elde edilen yıllık tiftik miktarı,  
 X<sub>37</sub>: Yumurta tavuğu sayısı,  
 X<sub>38</sub>: Et tavuğu sayısı,  
 X<sub>39</sub>: Hindi sayısı,  
 X<sub>40</sub>: Ördek sayısı,  
 X<sub>41</sub>: Kaz Sayısı,  
 X<sub>42</sub>: Toplam kovan sayısı,  
 X<sub>43</sub>: Üretilen yıllık bal miktarı,  
 X<sub>44</sub>: Üretilen yıllık balmumu miktarı,  
 Olmak üzere 44 adet veri setinden oluşmaktadır.



## 4.2. Analizde Kullanılan Kümeleme Yöntemleri ve Programlar

Araştırmada veriler kümeleme analizi yöntemlerinden hiyerarşik kümeleme yöntemlerinden tek bağlantı kümeleme yöntemi, ortalama bağlantı kümeleme yöntemi, tam bağlantı kümeleme yöntemi, merkezi bağlantı kümeleme yöntemi, mcquitty bağlantı kümeleme yöntemi, medyan bağlantı kümeleme yöntemi, ward bağlantı kümeleme yöntemleri kullanılmıştır. Uzaklık ölçülerinden öklid, karesel öklid, pearson, karesel pearson ve manhattan ölçüleri kullanılarak kümeler arası uzaklıklar hesaplanmıştır.

Veriler Minitab 16.2.1 (Lead Technologies 2010) ile analiz edilmiştir. Analizde bütün kümeleme yöntemleri, bütün uzaklık ölçülerine göre kümelere ayrılmıştır. Kümeler 5, 6, 7, 8, 9 ve 10'lu kümeler şeklinde oluşturulmuştur.

Ancak kümeleme yöntemlerinin ve uzaklıklarının bir arada karşılaştırılması için küme sayısı 7 olarak alınacaktır. Küme sayısı belirlenirken genel formüle başvurulmuş  $44^{1/2}=6,63$  yani yaklaşık olarak 7 belirlenmiştir. 5,6,8,9 ve 10' lu kümelerin oluşturduğu tablolara ise ekte yer verilecektir.

## 4.3. Uzaklık Ölçülerinin Karşılaştırılması

81 ilimize ait 44 adet değişkenden oluşan veri seti öklid uzaklığı, karesel öklid uzaklığı, Manhattan uzaklığı, pearson uzaklığı ve karesel pearson uzaklığından oluşan uzaklık ölçüleri yedi hiyerarşik kümeleme yöntemi ile karşılaştırılmıştır. Aşağıdaki tablolar incelendiğinde uzaklık ölçülerinden öklid ve karesel öklid uzaklık ölçüsü ward hiyerarşik kümeleme yöntemi dışındaki bütün hiyerarşik kümeleme yöntemlerinde aynı kümeleri oluşturmuştur. Ward kümeleme yönteminde ise dört tane aynı kümeden oluşturmuştur. Pearson ve karesel pearson uzaklıkları diğer uzaklık ölçülerinden farklı kümeler oluşturmuştur. Pearson ve karesel pearson uzaklıkları ile oluşan kümeler tek Bağlantı, tam bağlantı, ortalama bağlantı, medyan bağlantı kümeleme yöntemlerinde bütün kümeler aynı kümelerden oluşurken merkezi bağlantı ve mcquitty bağlantı kümeleme yöntemlerinde beş küme aynı illerden oluşmuştur. Ward bağlantı kümeleme yönteminde ise dört küme aynı illerden oluşmuşlardır. Manhattan uzaklık ölçüsü öklid ve karesel öklid uzaklığına büyük ölçüde benzer küme sonuçları vermiştir.

Öklid uzaklığı, Karesel Öklid uzaklığı, Manhattan uzaklığı benzer kümeler oluştururken, Pearson ve karesel pearson uzaklığı benzer sonuçlar vermiştir. Aşağıda uzaklıkları hiyerarşik kümeleme yöntemlerinde karşılaştıracamız.

#### 4.3.1. Tek Bağlantı Kümeleme Yöntemi ile Uzaklık Ölçülerinin Karşılaştırılması

Tablo 4.1 incelendiğinde öklid uzaklığı ve karesel öklid uzaklığına göre Afyon, Konya bir küme oluştururken, Balıkesir, Sakarya, Bolu, İzmir, Manisa tek başına ayrı ayrı küme oluşturmuştur. Diğer 74 il bir kümede toplanmıştır.

Manhattan uzaklığında ise öklid uzaklığı ve karesel öklid uzaklığından farklı olarak Konya, Afyon, Sakarya, Bolu, İzmir tek başlarına bir küme, Manisa ile Balıkesir bir küme, diğer 74 il bir küme oluşturmuşlardır. Pearson ve karesel pearson kümeleme yöntemlerinde Erzurum, Kars, Konya, Eskişehir, Ankara, Samsun birer küme diğer 75 il bir kümede toplanmıştır.

Tablo 4.1. Tek bağlantı kümeleme yöntemi ile oluşan kümeler

Küme	Öklid Uzaklığı	Karesel Öklid Uzaklığı	Manhattan Uzaklığı	Pearson Uzaklığı	Karesel Pearson Uzaklığı
1. Küme	Afyon Konya	Afyon Konya	Afyon	Erzurum	Erzurum
2.Küme	Balıkesir	Balıkesir	Balıkesir Manisa	Kars	Kars
3.Küme	Sakarya	Sakarya	Sakarya	Konya	Konya
4.Küme	Bolu	Bolu	Bolu	Eskişehir	Eskişehir
5. Küme	İzmir	İzmir	İzmir	Ankara	Ankara
6.Küme	Manisa	Manisa	Konya	Samsun	Samsun
7.Küme	Diğer İller	Diğer İller	Diğer İller	Diğer İller	Diğer İller

#### 4.3.2. Tam Bağlantı Kümeleme Yöntemi ile Uzaklık Ölçülerinin Karşılaştırılması

Tablo 4.2 incelendiğinde tam bağlantı kümeleme yönteminde, öklid uzaklığı ve karesel öklid uzaklığında bütün kümeler aynı illerden oluşmuştur. Manhattan uzaklığı öklid ve karesel öklid uzaklıklarına yakın kümeler oluşturmuştur, üç kümesi bu iki uzaklık ölçüsü ile aynı illerden oluşurken Sakarya ve Bolu bir kümede birleşmiştir. Uşak, Kocaeli,

Düzce kalabalık il grubundan ayrılarak yeni bir küme oluşturmuştur. Pearson ve karesel pearson uzaklığı aynı kümeleri oluşturmuşlardır.

Tablo 4.2. Tam bağlantı kümeleme yöntemi ile oluşan kümeler

Kümeler	Öklid Uzaklığı	Karesel Öklid Uzaklığı	Manhattan Uzaklığı	Pearson Uzaklığı	Karesel Pearson Uzaklığı
1. Küme	Afyon Konya	Afyon Konya	Afyon Konya	Afyon Konya Balıkesir İzmir	Afyon Konya Balıkesir İzmir
2.Küme	Balıkesir Manisa	Balıkesir Manisa	Balıkesir Manisa	Erzurum Kars	Erzurum Kars
3.Küme	Sakarya	Sakarya	Sakarya Bolu	Ağrı Van Şanlıurfa	Ağrı Van Şanlıurfa
4.Küme	Bolu	Bolu	Uşak Kocaeli Düzce	Muş Diyarbakır İstanbul Samsun	Muş Diyarbakır İstanbul Samsun
5. Küme	İzmir	İzmir	İzmir	Eskişehir	Eskişehir
6.Küme	Çanakkale Zonguldak Eskişehir Adana Mersin Bursa Ankara Uşak Kocaeli Düzce	Çanakkale Zonguldak Eskişehir Adana Mersin Bursa Ankara Uşak Kocaeli Düzce	Çanakkale Zonguldak Eskişehir Adana Mersin Bursa Ankara	Ankara	Ankara
7.Küme	Diğer İller	Diğer İller	Diğer İller	Diğer İller	Diğer İller

#### 4.3.3. Ortalama Bağlantı Kümeleme Yöntemi ile Uzaklık Ölçülerinin Karşılaştırılması

Tablo 4.3 incelendiğinde; ortalama bağlantı kümeleme yönteminde öklid uzaklığı, karesel öklid uzaklığı tamamen aynı kümeleri oluştururken, manhattan uzaklık ölçüsü iki küme dışında bu uzaklık ölçüleri ile aynı kümeleri oluşturmuşlardır, farklı olan bu iki kümenin de çok farklı olmadığı tablo incelendiğinde fark edilecektir. Uşak, Kocaeli ve Düzce illerinin bir kümeden diğer kümeye geçmesi ile kümede farklılıklar oluşmuştur. Pearson ve karesel pearson tüm kümeleri aynı illerden oluşturmuştur.

Tablo 4.3. Ortalama bağlantı kümeleme yöntemi ile oluşan kümeler

Kümeler	Öklid Uzaklığı	Karesel Öklid Uzaklığı	Manhattan Uzaklığı	Pearson Uzaklığı	Karesel Pearson Uzaklığı
1. Küme	Afyon Konya	Afyon Konya	Afyon Konya	Afyon Konya Balıkesir İzmir	Afyon Konya Balıkesir İzmir
2.Küme	Balıkesir Manisa	Balıkesir Manisa	Balıkesir Manisa	Erzurum Kars	Erzurum Kars
3.Küme	Sakarya	Sakarya	Sakarya	Ağrı Van Şanlıurfa	Ağrı Van Şanlıurfa
4.Küme	Bolu	Bolu	Bolu	Muş Diyarbakır İstanbul Samsun	Muş Diyarbakır İstanbul Samsun
5. Küme	İzmir	İzmir	İzmir	Eskişehir	Eskişehir
6.Küme	Çanakkale Zonguldak Eskişehir Adana Mersin Bursa Ankara Uşak Kocaeli Düzce	Çanakkale Zonguldak Eskişehir Adana Mersin Bursa Ankara Uşak Kocaeli Düzce	Çanakkale Zonguldak Eskişehir Adana Mersin Bursa Ankara	Ankara	Ankara
7.Küme	Diğer İller	Diğer İller	Diğer İller	Diğer İller	Diğer İller

#### 4.3.4. Merkezi Bağlantı Kümeleme Yöntemi ile Uzaklık Ölçülerinin Karşılaştırılması

Tablo 4.4 incelendiğinde; merkezi bağlantı kümeleme yönteminde öklid uzaklığı ve karesel öklid uzaklığı ile yapılan hiyerarşik kümeleme analizleri sonucunda bütün kümeler aynı kümelerden oluşturmuştur.

Manhattan uzaklık ölçüsü öklid ve karesel öklid uzaklık ölçüsüne benzer şekilde kümeler oluştururken sadece iki kümede kısmi farklılıklar oluşmuştur. Pearson ve karesel pearson uzaklığı beş kümede aynı illerden oluşurken iki kümede farklı illerden oluşmuştur.

Tablo 4.4. Merkezi bağlantı kümeleme yöntemi ile oluşan kümeler

Kümeler	Öklid Uzaklığı	Karesel Öklid Uzaklığı	Manhattan Uzaklığı	Pearson Uzaklığı	Karesel Pearson Uzaklığı
1. Küme	Afyon Konya	Afyon Konya	Afyon	Erzurum	Erzurum
2.Küme	Balıkesir Manisa	Balıkesir Manisa	Balıkesir Manisa	Kars	Kars
3.Küme	Sakarya	Sakarya	Sakarya	Konya	Konya
4.Küme	Bolu	Bolu	Bolu	Eskişehir	Eskişehir
5. Küme	İzmir	İzmir	İzmir	Ankara	Ankara
6.Küme	Çanakkale Zonguldak Eskişehir Adana Mersin Bursa Ankara Uşak Kocaeli Düzce	Çanakkale Zonguldak Eskişehir Adana Mersin Bursa Ankara Uşak Kocaeli Düzce	Konya	Samsun	Antalya Mersin
7.Küme	Diğer İller	Diğer İller	Diğer İller	Diğer İller	Diğer İller

#### 4.3.5. Medyan Bağlantı Kümeleme Yöntemi ile Uzaklık Ölçülerinin Karşılaştırılması

Tablo 4.5 incelendiğinde medyan bağlantı kümeleme yönteminde öklid uzaklığı ve karesel öklid uzaklığı bütün kümeleri aynı illerden oluşturmuştur. Manhattan uzaklığı her iki öklid uzaklığına benzer beş küme oluştururken iki kümede kısmi farklılıklar oluşturmuştur. Pearson ve karesel pearson uzaklıkları beş kümede benzer illerden oluşurken iki kümede kısmi farklılıklar oluşturmuştur.

Tablo 4.5. Medyan bağlantı kümeleme yöntemi ile oluşan kümeler

Kümelere	Öklid Uzaklığı	Karesel Öklid Uzaklığı	Manhattan Uzaklığı	Pearson Uzaklığı	Karesel Pearson Uzaklığı
1. Küme	Afyon Konya	Afyon Konya	Afyon Konya	Erzurum	Erzurum
2. Küme	Balıkesir Manisa	Balıkesir Manisa	Balıkesir Manisa	Kars	Kars
3. Küme	Sakarya	Sakarya	Sakarya	Konya	Balıkesir İzmir Afyon Konya
4. Küme	Bolu	Bolu	Bolu	Eskişehir	Eskişehir
5. Küme	İzmir	İzmir	İzmir	Ankara	Ankara
6. Küme	Çanakkale Zonguldak Eskişehir Adana Mersin Bursa Ankara Uşak Kocaeli Düzce	Çanakkale Zonguldak Eskişehir Adana Mersin Bursa Ankara Uşak Kocaeli Düzce	Uşak Kocaeli Düzce	Antalya Mersin	Antalya Mersin
7. Küme	Diğer İller	Diğer İller	Diğer İller	Diğer İller	Diğer İller

#### 4.3.6. McQuitty Bağlantı Kümeleme Yöntemi ile Uzaklık Ölçülerinin Karşılaştırılması

Tablo 4.6 incelendiğinde; McQuitty bağlantı kümeleme yönteminde öklid, karesel öklid ve manhattan uzaklık ölçüleri aynı illerden oluşan kümeler oluşturmuştur. Pearson ve karesel pearson uzaklığı beş kümede aynı illerden oluşurken sadece iki kümede küçük farklılıklar oluşturmuştur.

Tablo 4.6. McQuitty bağlantı kümeleme yöntemi ile oluşan kümeler

Kümeler	Öklid Uzaklığı	Karesel Öklid Uzaklığı	Manhattan Uzaklığı	Pearson Uzaklığı	Karesel Pearson Uzaklığı
1. Küme	Afyon Konya	Afyon Konya	Afyon Konya	Erzurum Ardahan Sivas Kars	Erzurum Ardahan Kayseri Tokat Sivas Kars
2.Küme	Balıkesir Manisa	Balıkesir Manisa	Balıkesir Manisa	Ağrı Van Şanlıurfa	Ağrı Van Şanlıurfa
3.Küme	Sakarya	Sakarya	Sakarya	Balıkesir İzmir Afyon Konya	Balıkesir İzmir Afyon Konya
4.Küme	Bolu	Bolu	Bolu	Eskişehir	Eskişehir
5. Küme	İzmir	İzmir	İzmir	Ankara	Ankara
6.Küme	Çanakkale Zonguldak Eskişehir Adana Mersin Bursa Ankara Uşak Kocaeli Düzce	Çanakkale Zonguldak Eskişehir Adana Mersin Bursa Ankara Uşak Kocaeli Düzce	Çanakkale Zonguldak Eskişehir Adana Mersin Bursa Ankara Uşak Kocaeli Düzce	Muş Diyarbakır İstanbul Samsun	Muş Diyarbakır İstanbul Samsun
7.Küme	Diğer İller	Diğer İller	Diğer İller	Diğer İller	Diğer İller

#### 4.3.7. Ward Bağlantı Kümeleme Yöntemi ile Uzaklık Ölçülerinin Karşılaştırılması

Tablo 4.7 incelendiğinde ward bağlantı kümeleme yönteminde öklid ve karesel öklid uzaklıklarına göre oluşan kümeler dört kümede aynı illerden oluşurken üç kümede kısmi farklı illerden oluşmuştur.

Manhattan uzaklığı iki kümede öklid uzaklıkları ile aynı kümeler oluştururken, beş kümede kısmi farklılıklar oluşturmuştur. Pearson ve karesel pearson dört kümeyi aynı illerden oluştururken üç kümede kısmi farklı illerden oluşmuştur.

Tablo 4.7. Ward bağlantı kümeleme yöntemi ile oluşan kümeler

Kümeler	Öklid Uzaklığı	Karesel Öklid Uzaklığı	Manhattan Uzaklığı	Pearson Uzaklığı		Karesel Pearson Uzaklığı
1. Küme	Afyon Konya	Afyon Konya	Afyon Konya	Erzurum Ardahan Sivas Kars Muş	Diyarbakır İstanbul Tokat Samsun	Erzurum Kars
2. Küme	Balıkesir Manisa	Balıkesir Manisa	Balıkesir Manisa	Balıkesir İzmir Afyon Konya Eskişehir		Balıkesir İzmir Afyon Konya Eskişehir
3. Küme	İzmir Sakarya	Sakarya	Sakarya Bolu	Muğla Adana Ordu Antalya Mersin		Muğla Adana Ordu Antalya Mersin
4. Küme	Bolu	Bolu	İzmir Bursa Ankara Kayseri Çorum	Ankara		Ankara
5. Küme	Malatya Çankırı İstanbul Bilecik Bartın Hatay Kırşehir Elazığ Aydın Denizli Samsun Kayseri Çorum	İzmir	Malatya Çankırı Hatay Kırşehir İstanbul Bilecik Karabük Bartın Elazığ Aydın Samsun Denizli	Ağrı Van Şanlıurfa		Ağrı Van Şanlıurfa
6. Küme	Çanakkale Zonguldak Eskişehir Adana Mersin Bursa Ankara Uşak Kocaeli Düzce	Çanakkale Zonguldak Eskişehir Adana Mersin Bursa Ankara Uşak Kocaeli Düzce	Çanakkale Zonguldak Eskişehir Adana Mersin Uşak Kocaeli Düzce	Erzincan Malatya Osmaniye Bayburt Çankırı Trabzon Kocaeli Düzce Bartın Giresun	Bilecik Nevşehir Kırşehir Kırıkkale Yalova Rize Zonguldak Karabük Gümüşhane Artvin Sinop	Muş Diyarbakır İstanbul Tokat Samsun
7. Küme	Diğer İller	Diğer İller	Diğer İller	Diğer İller		Diğer İller



#### 4.4. Kümeleme Yöntemlerinin Karşılaştırılması

Bu kısımda yedi hiyerarşik kümeleme yönteminin birbirlerine benzerlik veya farklılıkları beş uzaklık yöntemine göre analiz edilerek tablolar haline getirilmiştir. Tablolarda “ simgesi ortalama bağlantı kümeleme yöntemi ile oluşan kümelerle aynı illerden oluşan kümeleri ifade etmek için kullanılmıştır, illerin isimleri tekrardan tabloya yazılmamıştır. Aşağıdaki tablolarda oluşturulan bütün hiyerarşik kümeleme yöntemleri uzaklık ölçülerinde ayrıntılı olarak analiz edilecektir.

##### 4.4.1. Öklid Uzaklığına Göre Kümeleme Yöntemlerinin Karşılaştırılması

Tablo 4.8 incelendiğinde; öklid uzaklık ölçüsünde ortalama, tam, merkezi, medyan, mcquitty bağlantı kümeleme yöntemlerinde bütün kümeler aynı illerden oluşmuştur. Ward ve tek bağlantı kümeleme yöntemi dört kümede diğer bağlantı kümeleri ile aynı illerden oluşurken üç kümede farklılıklar göstermiştir.

Tablo 4.8. Öklid uzaklığına göre kümeleme yöntemlerinin karşılaştırılması

K.Yöntemi	1.Küme	2.Küme	3.Küme	4.Küme	5.Küme	6.Küme	7.Küme
Ortalama	Afyon Konya	Balıkesir Manisa	Sakarya	Bolu	İzmir	Çanakkale Zonguldak Eskişehir Adana Mersin Bursa Ankara Uşak Kocaeli Düzce	Diğerleri
Tam	“	“	“	“	“	“	“
Merkezi	“	“	“	“	“	“	“
Medyan	“	“	“	“	“	“	“
McQuitty	“	“	“	“	“	“	“
Tek	“	Balıkesir	“	“	“	Manisa	Diğerleri
Ward	“	“	Malatya Çankırı İstanbul Bilecik Bartın Hatay Kırşehir Elazığ Aydın Denizli Samsun Kayseri Çorum	“	İzmir Sakarya	“	Diğerleri

#### 4.4.2. Karesel Öklid Uzaklığına Göre Kümeleme Yöntemlerinin Karşılaştırılması

Tablo 4.9 incelendiğinde karesel öklid uzaklığına göre ortalama, tam, merkezi, medyan, mcquitty ve ward bağlantı kümeleme yöntemlerinde bütün kümeler aynı illerden oluşmuştur. Sadece tek bağlantı kümeleme yöntemi bu uzaklıkta diğer kümeleme yöntemlerinden farklılık göstermiştir. Tek bağlantı kümeleme yönteminin dört kümesi diğer kümeleme yöntemleri ile aynı illerden oluşmuştur. Üç kümesinde ise kısmi farklılık göstermiştir.

Tablo 4.9. Karesel öklid uzaklığına göre kümeleme yöntemlerinin karşılaştırılması

K.Yöntemi	1.Küme	2.Küme	3.Küme	4.Küme	5.Küme	6.Küme	7.Küme
Ortalama	Afyon Konya	Balıkesir Manisa	Sakarya	Bolu	İzmir	Çanakkale Zonguldak Eskişehir Adana Mersin Bursa Ankara Uşak Kocaeli Düzce	Diğerleri
Tam	“	“	“	“	“	“	“
Merkezi	“	“	“	“	“	“	“
Medyan	“	“	“	“	“	“	“
McQuitty	“	“	“	“	“	“	“
Tek	“	Balıkesir	“	“	“	Manisa	Diğerleri
Ward	“	“	“	“	“	“	“

#### 4.4.3. Manhattan Uzaklığına Göre Kümeleme Yöntemlerinin Karşılaştırılması

Tablo 4.10 incelendiğinde manhattan uzaklığına göre ortalama ve mcquitty bağlantı kümeleme yöntemi aynı illerden oluşan kümeler oluştururken, medyan bağlantı kümeleme yöntemi beş, merkezi ve tek bağlantı kümeleme yöntemi dört kümede, ward bağlantı kümeleme yöntemi iki, ortalama bağlantı kümeleme yöntemi ile aynı illerden oluşan kümeler oluşturmuşlardır.

Tablo 4.10. Manhattan uzaklığına göre kümeleme yöntemlerinin karşılaştırılması

K.Yöntemi	1.Küme	2.Küme	3.Küme	4.Küme	5.Küme	6.Küme	7.Küme
Tam	“	“	Sakarya Bolu	Uşak Kocaeli Düzce	“	Çanakkale Zonguldak Eskişehir Adana Mersin Bursa Ankara	Diğerleri
Merkezi	Afyon	“	“	“	“	Konya	Diğerleri
Medyan	“	“	“	“	“	Uşak Kocaeli Düzce	Diğerleri
McQuitty	“	“	“	“	“	“	“
Tek	Afyon	“	“	“	“	Uşak Kocaeli Düzce	Diğerleri
Ward	“	“	Sakarya Bolu	Malatya Çankırı Hatay Kırşehir İstanbul Bilecik Karabük Bartın Elazığ Aydın Samsun Denizli	İzmir Bursa Ankara Kayseri Çorum	Çanakkale Mersin Eskişehir Zonguldak Adana Uşak Kocaeli Düzce	Diğerleri
Ortalama	Afyon Konya	Balıkesir Manisa	Sakarya	Bolu	İzmir	Çanakkale Zonguldak Eskişehir Adana Mersin Bursa Ankara Uşak Kocaeli Düzce	Diğerleri

#### 4.4.4. Pearson Uzaklığına Göre Kümeleme Yöntemlerinin Karşılaştırılması

Tablo 4.11 incelendiğinde pearson uzaklık ölçüsüne göre merkezi, medyan ve tek bağlantı kümeleme yöntemleri aynı illerden oluşan kümeleri oluşturmuşlardır. Ortalama ve Tam bağlantı kümeleme yöntemleri 5 kümede aynı illerden kümeler oluşturmuşlardır. Ward bağlantı kümeleme yöntemi ortalama ve tam bağlantı kümeleme yöntemi ile aynı illerden oluşan iki küme oluşturmuştur.

Tablo 4.11. Pearson uzaklığına göre kümeleme yöntemlerinin karşılaştırılması

<b>K. Yöntemi</b>	<b>1.Küme</b>	<b>2.Küme</b>	<b>3.Küme</b>	<b>4.Küme</b>	<b>5.Küme</b>	<b>6.Küme</b>	<b>7.Küme</b>
<b>Ortalama</b>	Erzurum Kars	Ağrı Van Şanlıurfa	Balıkesir İzmir Afyon Konya	Eskişehir	Ankara	Antalya Mersin	Diğerleri
<b>Tam</b>	“	“	“	“	“	Muş Diyarbakır İstanbul Samsun	Diğerleri
<b>Merkezi</b>	Erzurum	Kars	Konya	“	“	Samsun	Diğerleri
<b>Medyan</b>	Erzurum	Kars	Konya	“	“	Samsun	Diğerleri
<b>McQuitty</b>	Erzurum Ardahan Sivas Kars	“	“	“	“	Muş Diyarbakır İstanbul Samsun	Diğerleri
<b>Tek</b>	Erzurum	Kars	Konya	“	“	Samsun	Diğerleri
<b>Ward</b>	Erzurum Ardahan Sivas Kars Muş Diyarbakır İstanbul Tokat Samsun	“	Balıkesir İzmir Afyon Konya Eskişehir	Erzincan Malatya Osmaniye Bayburt Çankırı Trabzon Kocaeli Düzce Bartın Giresun Bilecik Nevşehir Kırşehir Kırıkkale Yalova Rize Zonguldak Karabük Gümüşhane Artvin Sinop	“	Muğla Adana Ordu Antalya Mersin	Diğerleri

#### 4.4.5. Karesel Pearson Uzaklığına Göre Kümeleme Yöntemlerinin Karşılaştırılması

Tablo 4.12 incelendiğinde karesel pearson uzaklığına göre merkezi, medyan ve tek bağlantı kümeleme yöntemleri beş kümede aynı illerden oluşan kümeler oluşturmuştur. Tam ve ortalama bağlantı kümeleme yöntemleri beş kümede aynı illerden oluşan kümeler oluşturmuşlardır. McQuitty ve tam bağlantı kümeleme yöntemi beş kümede aynı illerden oluşan kümeler oluşturmuşlardır. Ward bağlantı kümeleme yöntemi tam ve ortalama bağlantı kümeleme yöntemi ile aynı illerden oluşan üç küme oluşturmuştur.

Tablo 4.12. Karesel Pearson uzaklığına göre kümeleme yöntemlerinin karşılaştırılması

K. Yöntemi	1.Küme	2.Küme	3.Küme	4.Küme	5.Küme	6.Küme	7.Küme
Ortalama	Erzurum Kars	Ağrı Van Şanlıurfa	Balıkesir İzmir Afyon Konya	Eskişehir	Ankara	Antalya Mersin	Diğerleri
Tam	“	“	“	“	“	Muş Diyarbakır İstanbul Samsun	Diğerleri
Merkezi	Erzurum	Kars	Konya	“	“	“	Diğerleri
Medyan	Erzurum	Kars	“	“	“	“	Diğerleri
McQuitty	Erzurum Ardahan Kayseri Tokat Sivas Kars	“	“	“	“	Muş Diyarbakır İstanbul Samsun	Diğerleri
Tek	Erzurum	Kars	Konya	“	“	Samsun	Diğerleri
Ward	“	“	Balıkesir İzmir Afyon Konya Eskişehir	Muğla Adana Ordu Antalya Mersin	“	Muş Diyarbakır İstanbul Samsun	

#### 4.5. Kümeleme Yöntemleri ve Uzaklık Ölçülerine Göre Ortak Kümeler

Tüm tablolar incelendiğinde Öklid, karesel öklid ve manhattan uzaklık ölçülerine göre oluşan kümeler ile pearson ve karesel pearson uzaklıklarına göre oluşturulan kümelere farklılıklar olduğu ortaya çıkmıştır. Hangi uzaklık yöntemine göre kümelerin oluşumunun referans alınacağına etki eden değişkenlerin incelenmesiyle karar verilebilir.

Tablo 4.13 öklid, karesel öklid ve manhattan uzaklık ölçülerine göre incelendiğinde, tüm hiyerarşik kümeleme yöntemlerinde Afyon ve Konya illerinin birinci kümeyi, Balıkesir ve Manisa illerinin ikinci kümeyi, Sakarya ilinin tek başına 4. Kümeyi, Bolu ilinin tek başına 5. Kümeyi, Çanakkale, Zonguldak, Eskişehir, Adana, Mersin, Bursa, Ankara, Uşak, Kocaeli, Düzce illerinin 6. kümeyi, geriye kalan 64 ilinde 7. kümeyi oluşturduğu görülmektedir.

Tablo 4.13. Öklid, karesel öklid, manhattan ortak kümeleri

Küme	İller
1. Küme	Afyon,Konya
2.Küme	Balıkesir, Manisa
3.Küme	Sakarya
4.Küme	Bolu
5.Küme	İzmir
6.Küme	Çanakkale, Zonguldak, Eskişehir, Adana, Mersin, Bursa, Ankara, Uşak, Kocaeli, Düzce
7. Küme	Diğer İller

Tablo 4.14 pearson ve karesel pearson uzaklıklarına göre incelendiğinde, tüm hiyerarşik kümeleme yöntemlerine göre Erzurum ve Kars 1. kümeyi, Ağrı, Van, Şanlıurfa illerinin 2. kümeyi, Balıkesir, İzmir, Afyon ve Konya illerinin 3. kümeyi, Eskişehir ilinin tek başına 4. kümeyi, Ankara ilinin tek başına 5. kümeyi, Muş, Diyarbakır, İstanbul, Samsun illerinin 6. kümeyi, geriye kalan 66 ilin 7. kümeyi oluşturduğu görülmektedir.

Tablo 4.14. Pearson ve karesel pearson uzaklık ölçüleri ortak kümeleri

Küme	İller
1. Küme	Erzurum, Kars
2.Küme	Ağrı ,Van, Şanlıurfa
3.Küme	Balıkesir, İzmir, Afyon, Konya,
4.Küme	Eskişehir
5.Küme	Ankara
6.Küme	Muş, Diyarbakır, İstanbul, Samsun
7. Küme	Diğer İller

#### 4.6. Değişkenlerin Kümeleme Yöntemlerine Etkisi

İllerin kümelenmesine değişkenlerin etkisi bu kısımda incelenecektir. Tablo 4.15 ve Tablo 4.16 incelendiğinde illerin değişkenlere göre büyükten küçüğe doğru sıralanışına göre ilk on il tablo haline getirilmiştir. Daha fazlasına ihtiyaç duyulmadığından sadece ilk sıradaki on il ile yetinilmiştir. Kümeleme analizinde uzaklık ölçüleriyle yapılan hiyerarşik kümeleme yöntemleri ile oluşturulan kümelerin de illerin sıralaması doğrultusunda olduğu görülmektedir.

Tablo 4.15. İllerin değişken büyüklüklerine göre sıralanışı 1. kısım

	Koyun (Yerli)	Koyun (Merinos)	Keçi (Kıl)	Keçi (Tiftik)	Sığır (Kültür)	Sığır (Melez)	Sığır (Yerli)
1	Van	Ankara	Mersin	Ankara	Balıkesir	Erzurum	Kars
2	Konya	Eskişehir	Antalya	Karaman	Konya	Ardahan	Ağrı
3	Ağrı	Afyon	Adana	Siirt	İzmir	Sivas	Diyarbakır
4	Şanlıurfa	Karaman	Mardin	Eskişehir	Aydın	Kars	Erzurum
5	Muş	Balıkesir	Bitlis	Bolu	Denizli	Konya	Muş
6	Diyarbakır	Konya	Siirt	Kastamonu	Afyon	Adana	Van
7	Balıkesir	Bursa	Çanakkale	Kırıkkale	Burdur	Samsun	Samsun
8	Bitlis	Edirne	Denizli	Mardin	Çanakkale	Kayseri	Yozgat
9	Manisa	Burdur	Diyarbakır	Konya	Kırklareli	Ağrı	Ankara
10	Elazığ	Antalya	Balıkesir	İzmir	Aksaray	Ankara	Kastamonu

Tablo 4.16. İllerin değişken büyüklüklerine göre sıralanışı 2. kısım

	Manda	Tavuk (Et)	Tavuk (Yumurta)	Hindi	Kaz	Ördek	Aricılık
1	Samsun	Bolu	Afyon	Bolu	Kars	Muş	Ordu
2	İstanbul	Balıkesir	Konya	Manisa	Muş	Samsun	Muğla
3	Diyarbakır	Manisa	Manisa	İzmir	Ardahan	Balıkesir	Adana
4	Tokat	Sakarya	Balıkesir	Sakarya	Batman	Diyarbakır	Mersin
5	Muş	İzmir	İzmir	Kocaeli	Afyon	Ağrı	Aydın
6	Bitlis	Düzce	Çorum	Şanlıurfa	Ağrı	Edirne	İzmir
7	Afyon	Uşak	Bursa	Eskişehir	Samsun	Kars	Balıkesir
8	Sivas	Kocaeli	Kayseri	İstanbul	Yozgat	Tekirdağ	Sivas
9	Kayseri	Mersin	Ankara	Diyarbakır	Kütahya	Afyon	Antalya
10	Düzce	Ankara	Denizli	Muş	Konya	Muğla	Denizli

Bir önceki başlıkta oluşan ortak kümelerin tablosu, illerin değişkenlerinin sıralanmasıyla oluşan tablo karşılaştırılarak incelendiğinde oluşan kümeleri iki başlıkta değerlendireceğiz.

#### 4.6.1. Öklid, Karesel Öklid ve Manhattan Uzaklıklarına Göre Oluşan Kümelere Değişkenlerin Etkisi

1. Kümeyi oluşturan Afyon ve Konya illerinin değişkenleri Tablo 4.16'da incelendiğinde merinos koyununa ait veriler, kültür sığır ırkına ait veriler ve yumurta tavukçuluğu verileri bakımından benzerlik gösterdiği tablolardan anlaşılmaktadır. Bu iki ilin özellikle yumurta tavukçuluğunda ön planda olduğu gözükmektedir. Yumurta tavukçuluğu verilerinin kümenin oluşmasında etkin olduğu ortaya çıkmaktadır.

2. Kümeyi oluşturan Balıkesir ve Manisa illerinin değişkenleri incelendiğinde yerli koyun ırkına ait veriler, et tavukçuluğuna ait veriler, yumurta tavukçuluğuna ait veriler bakımından benzerlikler gösterdiği gözlenmektedir. Özelde et ve yumurta tavukçuluğunun bu kümenin oluşumunda etkin olduğu gözlenmektedir.

3. Kümeyi oluşturan Sakarya ilinin değişkenleri incelendiğinde et tavukçuluğu ve hindiciliğe ait veriler bakımından tek başına bir küme oluşturduğu gözlenmektedir.

4. Kümeyi oluşturan Bolu ilinin et tavukçuluğu ve hindiciliğe ait veriler bakımından ön sırada olduğu tablolardan anlaşılmaktadır. Ayrıca tiftik keçisine ait veriler bakımından da diğer illerden ayrıştığı gözlenmektedir.

5. Kümeyi oluşturan İzmir ilinin kültür sığırcılığına ait veriler, et ve yumurta tavukçuluğuna ait veriler, hindi yetiştiriciliğine ait veriler, arıcılığa ait veriler bakımından ön plana çıkması diğer illerden İzmir ilini ayıran önemli özellik olarak ortaya çıkmaktadır.

6. Kümeyi oluşturan Çanakkale, Zonguldak, Eskişehir, Adana, Mersin, Bursa, Ankara, Uşak, Kocaeli, Düzce illerine ait veriler incelendiğinde merinos koyuncululuğu, kıl keçisi, sığırcılık, et ve yumurta tavukçuluğu, hindicilik ve arıcılığa ait veriler bakımından benzerlik gösterdikleri görülmektedir.

7. Kümeyi oluşturan 64 ilin birçok değişkene göre sınıflandıkları gözlenmektedir. Küme sayısının artırılmasıyla illerin farklı kümeleneyeceği dendogramlar incelendiğinde görülecektir.

#### **4.6.2. Pearson ve Karesel Pearson Uzaklıklarına Göre Oluşan Kümelerde Değişkenlerin Etkisi**

1. Kümeyi oluşturan Erzurum ve Kars illerinin değişkenleri incelendiğinde sığır (melez) ve sığır (yerli) verileri bakımından benzerlik gösterdikleri görülmektedir. Kümelerin oluşumunda bu verilerin ön plana çıktığı görülmektedir.



2. Kümeyi oluşturan Ağrı, Van ve Şanlıurfa illerine ait hayvancılık verilerinden yerli koyun, yerli sığır verileri bakımından kümenin oluştuğu gözlenmektedir. En önemli etkenin yerli koyunculığa ait veriler üzerinden gerçekleştiği görülmektedir.

3. Kümeyi oluşturan Balıkesir, İzmir, Afyon, Konya illerine ait değişkenlerden yerli koyunculığa ait veriler , tiftik keçiciliğine ait veriler, et ve yumurta tavukçuluğuna ait veriler, kaz ve ördek verileri , arıcılık verileri , merinos koyuncululuğuna ait veriler, kültür sığırcılığına ait verilere göre kümelendiği gözlemlenmektedir. Kültür sığırcılığının ve yumurta tavukçuluğunun kümenin oluşumunda en büyük etken olduğu anlaşılmaktadır.

4. Kümeyi oluşturan Eskişehir iline ait merinos koyuculuğu, tiftik keçiciliği ve hindi yetiştiriciliğine ait verilere ait değişkenlere göre tek başına bir küme oluşturduğu gözlenmektedir. Daha çok merinos koyunculığa ait verinin etkin olduğu gözlemlenmektedir.

5. Kümeyi oluşturan Ankara iline ait verilere göre merinos koyuculuğu, tiftik keçiciliği, melez sığırcılık, yerli sığırcılık, et ve yumurta tavukçuluğuna ait değişkenlere göre ilin tek başına küme oluşturduğu gözlenmiştir. Ankara ili merinos koyunu ve tiftik keçisine ait verilerle ön plana çıkmaktadır.

6. Kümeyi oluşturan Muş, Diyarbakır, İstanbul, Samsun illerine ait verilerden yerli koyun, yerli sığır, manda, hindi, kaz ve ördek yetiştiriciliğine ait değişkenlere göre kümenin şekillendiği görülmektedir. Küme özellikle manda, hindi ve ördek yetiştiriciliğine göre şekillenmiştir. Kümeye giren İstanbul ili istatistiklerin incelenmesi neticesinde bu kümeden çıkartılarak 7. Kümeye kaydırılması uygun olacaktır.

7. Kümeyi oluşturan 66 ilin dağılımları küme sayısının arttırılması ile değişebilir.

#### **4.7. İllerin Birbirine Uzaklık ve Yakınlıkları**

İllerin birbirine olan uzaklık ve yakınlıkları Ortalama bağlantı ve öklid uzaklığı baz alınarak yapılan analizde hayvancılık verilerinde birbirine en yakın iki il %99,8683 oran ile Artvin ve Giresun olurken, en uzak iki il %30,4912 Adana ve Balıkesir olarak tespit edilmiştir. Diğer illerin durumu da tablo 4.17.' de ayrıntılı olarak verilmiştir.

Tablo 4.17. İllerin Birbirine Uzaklık ve Yakınlıkları

99,8683	Artvin	Giresun	98,5243	Ağrı	Şanlıurfa
99,7668	Gümüşhane	Yalova	98,5056	Bitlis	Diyarbakır
99,7239	Sinop	Bayburt	98,4743	Amasya	Gaziantep
99,698	Artvin	Rize	98,455	Bitlis	Niğde
99,6907	Tunceli	Şırnak	98,4206	Adıyaman	Kars
99,6584	Gümüşhane	Sinop	98,1835	Adıyaman	Ordu
99,6137	Artvin	Trabzon	98,1379	Aydın	Elazığ
99,5923	Bingöl	Hakkari	98,11	Adıyaman	Bitlis
99,5409	Artvin	Gümüşhane	98,0017	Antalya	Nevşehir
99,5401	Nevşehir	Kırkkale	97,91	Kocaeli	Uşak
99,4814	Bingöl	Tunceli	97,8625	Adıyaman	Antalya
99,4724	Bitlis	Siirt	97,8198	Bilecik	İstanbul
99,4431	Edirne	Kırklareli	97,7464	Adıyaman	Erzurum
99,3984	Batman	Iğdır	97,6443	Bilecik	Çankırı
99,3932	Gaziantep	Kütahya	97,4507	Çorum	Kayseri
99,3902	Adıyaman	Isparta	97,3953	Çanakkale	Zonguldak
99,3549	Kastamonu	Tokat	97,3228	Ağrı	Van
99,285	Bilecik	Bartın	96,9999	Adıyaman	Muğla
99,2676	Kilis	Osmaniye	96,9312	Denizli	Samsun
99,255	Adıyaman	Edirne	96,9288	Ankara	Bursa
99,2211	Tekirdağ	Yozgat	96,522	Adıyaman	Bilecik
99,2044	Adıyaman	Kastamonu	96,4663	Adana	Eskişehir
99,1626	Bitlis	Batman	96,3857	Çanakkale	Mersin
99,1501	Çankırı	Malatya	96,2579	Adıyaman	Amasya
99,1437	Nevşehir	Tekirdağ	95,914	Aydın	Denizli
99,1224	Kahramanmaraş	Aksaray	95,3617	Adana	Çanakkale
99,0678	Artvin	Kilis	95,1557	Adıyaman	Ağrı
99,0584	Adıyaman	Burdur	93,2776	Afyon	Konya
99,0433	Hatay	Kırşehir	93,2195	Kocaeli	Düzce
99,0099	Artvin	Bingöl	93,0034	Adıyaman	Aydın
98,9982	Erzincan	Kahramanmaraş	90,0368	Balıkesir	Manisa
98,9898	Bitlis	Mardin	88,8864	Adana	Kocaeli
98,9842	Bilecik	Hatay	87,1557	Adıyaman	Çorum
98,9506	Adıyaman	Erzincan	86,303	Adana	Ankara
98,9481	Amasya	Karaman	80,5363	Adana	Adıyaman
98,9374	Diyarbakır	Muş	74,4325	Balıkesir	Sakarya
98,7264	Adıyaman	Artvin	70,7023	Balıkesir	Bolu
98,712	Kars	Ardahan	63,0092	Balıkesir	İzmir
98,7077	Bilecik	Karabük	62,6912	Adana	Afyon
98,6705	Erzurum	Sivas	30,4912	Adana	Balıkesir

## 5. SONUÇLAR VE ÖNERİLER

Araştırmada hiyerarşik kümeleme analizi yöntemlerinden tek bağlantı kümeleme yöntemi, ortalama bağlantı kümeleme yöntemi, tam bağlantı kümeleme yöntemi, merkezi bağlantı kümeleme yöntemi, mcquitty bağlantı kümeleme yöntemi, medyan bağlantı kümeleme yöntemine göre yapılan analizlerde oluşan kümelerin benzer kümeler oluşturduğu, ward bağlantı kümeleme yönteminin bu bağlantı kümeleme yöntemlerinden kısmi farklılıkta kümeler oluşturduğu tespit edilmiştir.

Uzaklık ölçülerine bağlı olarak yapılan analizlerde öklid, karesel öklid ve manhattan ölçüleri kullanılarak oluşturulan kümelerde kümeler arası mesafenin benzer uzaklıkta veya yakınlıkta olduğu tespit edilmiştir. Pearson ve karesel pearson uzaklık ölçüleri kullanılarak oluşturulan kümeler arası uzaklığın ve yakınlığın benzer olduğu görülmüştür. Uzaklık ölçülerinin analizlerde çıkan sonuçlara göre iki uzaklık ölçüsü grubu olduğu tespit edilmiştir. Birinci grubun öklid, karesel öklid ve manhattan olduğu, diğer grubun pearson ve karesel pearson olduğu gözlemlenmiştir.

Türkiye'deki 81 ilimize ait kümeleme analizinde bazı illerimizin diğer illere göre önemli derecede farklılıklar gösterdiği gözlemlenmiştir. Afyon, Konya, Balıkesir, Bolu, Manisa, İzmir illeri yumurta tavukçuluğunda ön plana çıktığı oluşan kümelerden gözlenmektedir. Yumurta birliklerinin de bu illerde kurulmuş olması kümeleme analizi sonuçlarını doğrulamaktadır. Manda yetiştiriciliğinde Muş, Diyarbakır ve Samsun illerinin ön plana çıktığı oluşan kümelerde gözlemlenmiştir. Manda yetiştiriciliğinde bu illerin ön plana çıkmış olması kümelerin isabetli bir şekilde oluştuğunu göstermektedir. Erzurum ve Kars illeri yerli ve melez sığır yetiştiriciliğinde ön plana çıkarken, Ağrı, Şanlıurfa ve Van koyun yetiştiriciliğinde ön plana çıkmıştır. Merinos koyuncululuğunda, Ankara ve Eskişehir illerinin öne çıktığı gözlemlenmiştir. Arıcılıkta Ordu, Muğla ve Adana illerinin uygun ortamlar sağladığı görülmektedir.

Hayvancılık ile ilgili illerimizin mevcut potansiyelinin değerlendirilerek, ilgili alanlarda teşvikler verilerek ilgili hayvancılık kolunda var olan potansiyelin güçlendirilmesi gerektiği düşünülmektedir.

Verilerin sınıflandırılmasında, kümeleme analizinin başarılı sonuçlar verdiği gözlemlenmiştir. Kümeleme analizi araştırmacıya verileri düzenli, yorumlanabilir ve grafiksel zenginlikte sunduğundan dolayı önerilebilir bir yöntem olarak karşımıza çıkmaktadır. Araştırmacı hiyerarşik yedi kümeleme yönteminden hangisini tercih etmesi ile ilgili “şu yöntem” tercih edilmelidir şeklindeki kesin hüküm içeren ifadelerden kaçınılması gerektiği sonucuna varılmıştır. Verilerin analizinde bütün hiyerarşik kümeleme yöntemleri denenmeli ve karşılaştırılarak incelenmeli oluşan kümelerin kesişimi daha doğru analizlerin yapılmasına imkân sağlamaktadır. Uzaklık ölçüleri olarak ta bütün uzaklık ölçülerine göre analiz yapılması araştırmacının verilerini sınıflandırmasında büyük kolaylıklar sağlayacaktır.

Sonuç olarak günümüzde bütün kümeleme yöntemleri ve bütün uzaklık ölçülerine göre istatistiksel analizleri kolaylıklı yapmamızı sağlayan paket programların ve çeşitli yazılımların varlığı araştırmacının işini kolaylaştırmaktadır. Kümeleme analizinde Minitab (Lead Technologies 2010) paket programı ortaya koyduğu dendogramların yalınlığı ve görsel zenginliği bakımından araştırmacılara tavsiye edilebilir bir paket program olarak karşımıza çıkmaktadır.

## KAYNAKLAR

Ada, A. “Kümeleme Analizi ile AB Ülkeleri ve Türkiye’nin Sürdürülebilir Kalkınma Açısından Değerlendirilmesi”, Dumlupınar Üniversitesi Sosyal Bilimler Dergisi, 29: 319-332, 2011.

Anderberg, M.R. “Cluster Analysis for applications”, Academic Press, 553-555, New York, 1973.

Atalay, A., Torum, A. “Türkiye’deki illerin 1997-2006 Yılları Arası Trafik Kazalarına Göre Kümeleme Analizi”, Pamukkale Üniversitesi Mühendislik Bilimleri Dergisi, 16(3): 335-343, 2010.

Atbaş, A.C.G. “Kümeleme Analizinde Küme Sayısının Belirlenmesi Üzerine Bir Çalışma”, Yüksek Lisans Tezi, Ankara Üniversitesi, Fen Bilimleri Enstitüsü, 14, 2008.

Banoğlu, K., Çakır M. “Liselerin Öğrenme Ortamı Profillerinin Belirlenmesine Yönelik Bir Kümeleme Analizi Çalışması”, Eğitim Bilimleri ve Uygulama Dergisi, 10(20): 161-179, 2011.

Blashfield, R.K., Aldenderfer, M.S. “ The Literature on Cluster Analysis”, Multivariate Behavioral Research, 13, 271-295, 1978.

Çelik, B., Akçapınar, H. “Ankara Keçisinin Tiftik Özellikleri Yönünden Kümeleme Analizi İle İncelenmesi”, Lalahan Hayvancılık Araştırma Enstitüsü Dergisi, 46(1): 19-27, 2006.

Çelik, Ş. “Kümeleme Analizi İle Sağlık Göstergelerine Göre Türkiye’deki İllerin Sınıflandırılması”, Doğu Üniversitesi Dergisi, 14(2): 175-194, 2013.

Demiralay, M., Çamurcu, Y. “Cure, Agnes ve K-means Algoritmalarındaki Kümeleme Yeteneklerinin Karşılaştırılması”, İstanbul Ticaret Üniversitesi Fen Bilimleri Dergisi, 8(2): 1-18, 2005.

Dinler, M., Doğan, Z. “Doğu Anadolu İllerinin Koyunculuk Yönünden Hiyerarşik Kümeleme Yöntemleri ile İncelenmesi”, 9. Ulusal Zootekni Öğrenci Kongresi, Erzurum, Türkiye, 217-224, 2013.

Doğan, İ. “Kümeleme Analiz ile seleksiyon”, Turkish Journal of Veterinary and Animal Sciences, 26: 47-53, 2002.

Everitt, S.B. “Cluster Analysis”, Heinmann Educational Books, 115, London, 1974.

Everitt, S.B., Landau, S., Leese, M. "Cluster Analysis", Oxford University Press Inc, 122, New York, 2001.

Fırat, M., Dikbaşı, F., Koç, A., Güngör, M. "K-Ortalamlar Yöntemi ile Yıllık Yağışların Sınıflandırılması ve Homojen Bölgelerin Belirlenmesi", İMO Teknik Dergi, 383: 6037-6050, 2012.

Galimberti, G., Soffritti, G. "Model-Based Methods to Identify Multiple Cluster Structures In a Data Set", Computational Statistics & Data Analysis, 52: 520-536, 2007.

Gevrekçi, Y., Ataç, F.E., Takma, Ç., Akbaş, Y., Taşkın, T. "Koyunculuk Açısından Batı Anadolu İllerinin Sınıflandırılması", Kafkas Üniversitesi Veterinerlik Fakültesi Dergisi, 17 (5): 755-760, 2011.

Gower, J.C. "A general coefficient of similarity and some of its properties", Biometrics, 27(4): 857-872, 1971.

Gürcan, İ.S., Akçapınar, H. "Alman Et ve Karacabey Merinosu Koyunlarının Canlı Ağırlık, Vücut Ölçüleri ve Yapağı İnceliği Yönünden Kümeleme Analizi ile İncelenmesi", Turkish Journal of Veterinary and Animal Sciences, 26, 1255-1261, 2002.

Hair, J.F., Anderson, R.E., Tatham R.L., Black, W.C. "Multivariate Data Analysis With Readings", NY Macmillan, Newyork, 1995.

Han, J., Kamber, M. "Data Mining Concepts and Techniques", Morgan Kaufmann Publishers Inc, 402, San Francisco, 2006.

<http://tuikapp.tuik.gov.tr/hayvancilikapp/hayvancilik.zul> (erişim tarihi: 01.11.2013)

<http://www.ist.yildiz.edu.tr/dersler/dersnotu/kum-analiz.doc> (erişim tarihi: 01.02.2014)

Hubert, L. "Approximate Evaluation Techniques for the Single-Link and Complete-Link Hierarchical Clustering Procedures", Journal of the American Statistical Association, 69: 698-704, 1974.

Karabulut, M., Gürbüz, M., Sandal, E.K., "Hiyerarşik Kluster (küme) Tekniği Kullanılarak Türkiye'de İllerin Sosyo-ekonomik Benzerliklerinin Analizi", Ankara Üniversitesi Coğrafi Bilimler Dergisi, 2(2): 65-78, 2004.

Kaufman, L., Rousseeuw, P.J. "Finding Groups in Data: An Introduction to Cluster Analysis", Wiley-Interscience, New York, 1990.

Kılıç, İ., Özbeyaz C. "Bulanık Kümeleme Analizinin Koyun Yetiştiriciliğinde Kullanımı ve Bir Uygulama", Kocatepe Veteriner Dergisi, 3(2): 31-37, 2004.

Marriot, F.H.C. "Practical Problems in a Method of Cluster Analysis", Biometrics, 27: 501-514, 1971.

Öz, B., Taban, S., Kar, M. “Kümeleme Analizi ile Türkiye ve AB Ülkelerinin Beşeri Sermaye Göstergeleri Açısından Karşılaştırılması”, Eskişehir Osmangazi Üniversitesi Sosyal Bilimler Dergisi, 10(1): 1-30, 2009.

Özbeyaz, C., Yıldız, M., Çamdeviren, H. “Türkiye’ de Yetiştirilen Çeşitli Sığır Irkları Arasındaki Genetik İlişkiler”, Lalahan Hayvancılık Araştırma Enstitüsü Dergisi, 39(1): 17-32, 1999.

Özdamar, K. “Paket Programlar ile İstatistiksel Veri Analizi (Çok Değişkenli Analizler)”, Kaan Kitabevi, 279, Eskişehir, 2004.

Özgen, N., Çelik, C. “Üniversite Öğrencilerinin Coğrafya Dersine Yönelik Tutumlarının Kümeleme Analizi İle Belirlenmesi (Siirt Eğitim Fakültesi Örneği)”, Marmara Coğrafya Dergisi, 17: 67-78, 2008.

Öztürk, I., Yıldız, N. “n<p Boyutlu Biyolojik verilerde farklı kümeleme yöntemlerinin karşılaştırmalı olarak incelenmesi”, KSÜ Doğa Bilim Dergisi, 15(4): 26-36, 2012.

Rencher, A.C. “Methods of Multivariate Analysis, Second Edition”, A John Wiley & Sons, Inc. Publications, 451-481, USA, 2002.

Sangün, L., Akar M. “Aşamalı Kümeleme Yöntemlerinin Su Ürünlerinde Kullanımı”, Ç.Ü Fen ve Mühendislik Bilimleri Dergisi, 26(3): 1-8, 2011.

Sharma, S., “Applied Multivariate Techniques”, Wiley-Interscience, New York, 1996.

Tatlıdil, H. “Uygulamalı Çok Değişkenli İstatistiksel Analiz”, Akademi Matbaası, 329-343, Ankara, 1996.

Turhan, K., Engin, Y. , Kamburoğlu, S., Kurt, B., Örem, A. “Potasyum (K) ve Hemoliz İndeksi (HI) Arasındaki İlişkinin Kümeleme Analizi Yöntemi ile İrdelenmesi”, 9. Ulusal Tıp Bilişimi Kongresi, Antalya, Türkiye, s. 106-112, 2012.

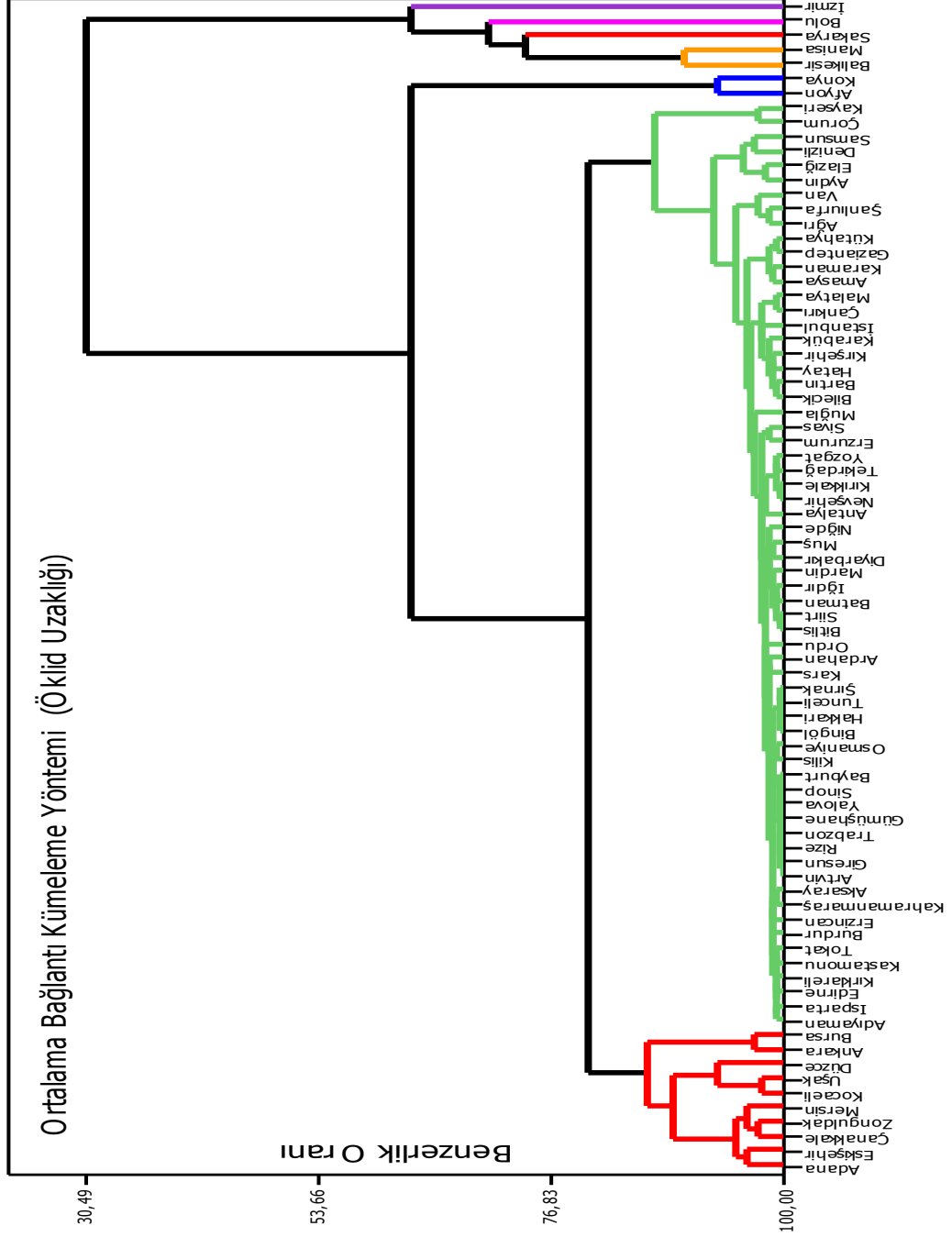
Urfalıoğlu, F., Seyfullahoğulları, A. “İllerin Bazı Sosyo-Ekonomik Kriterler Altında Gelişmişlik Sınıflaması”, Marmara Üniversitesi İktisadi İdari Bilimler Fakültesi Dergisi, 19(1): 209-232, 2004.

Yazgan, E., Kayaalp, G.T. “Kümeleme (Cluster ) Analizi Yöntemlerinin Karşılaştırmalı Olarak İncelenmesi ve Tarımsal Araştırmalarda Kullanılması”, Çukurova Üniversitesi Ziraat Fakültesi Dergisi, 17 ( 3 ): 7- 14, 2002.

Yılandı, V. “Bulanık Kümeleme Analizi ile Türkiye’deki İllerin Sosyoekonomik Açından Sınıflandırılması”, Süleyman Demirel Üniversitesi İktisadi ve İdari Bilimler Fakültesi Dergisi, 15(3): 453-470, 2010.

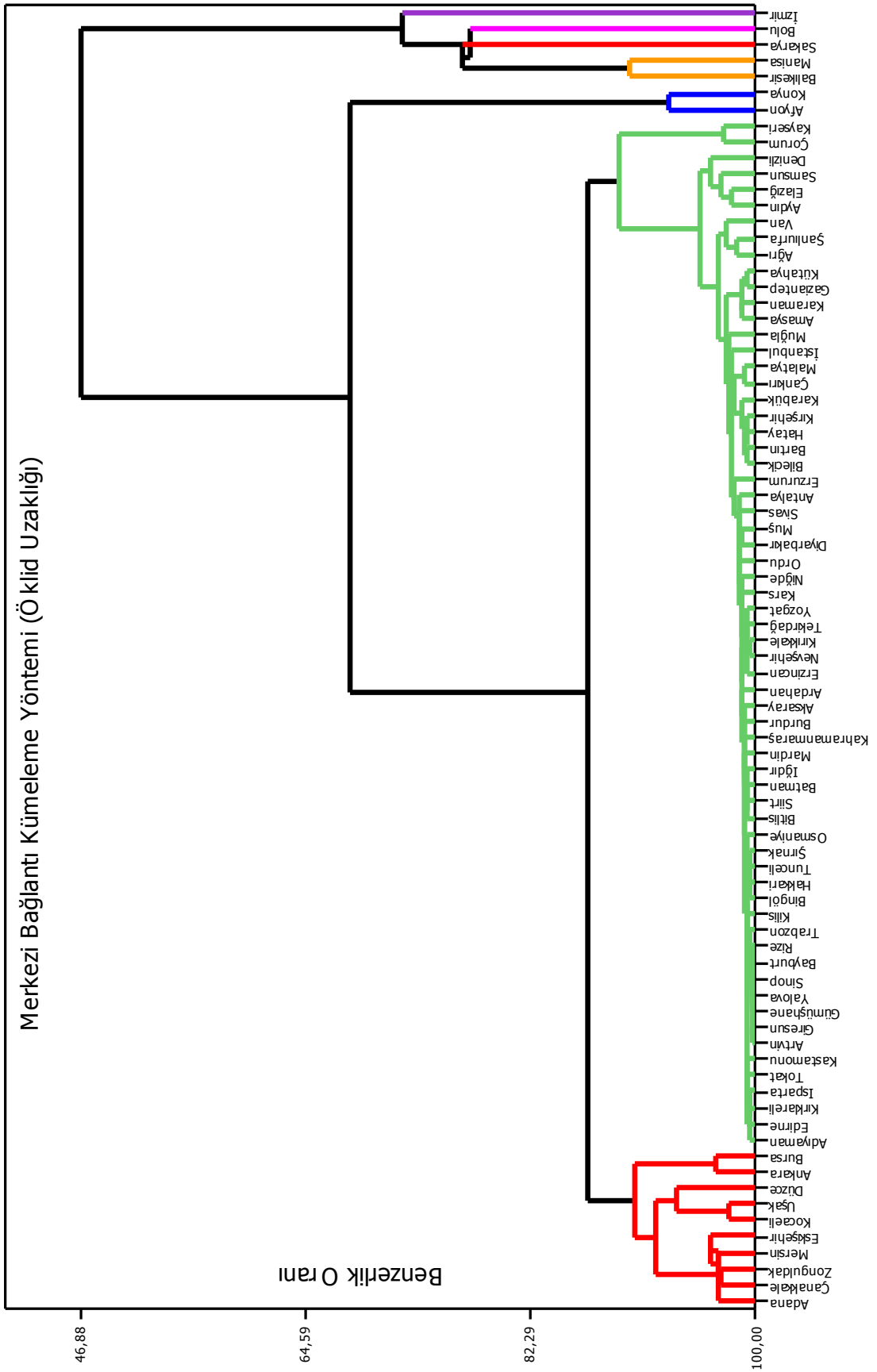
Yılmaz, Ş., Patır, S. “Kümeleme Analizi ve Pazarlamada Kullanımı”, Akademik Yaklaşımlar Dergisi, 2(1): 91-113, 2011.

## EKLER

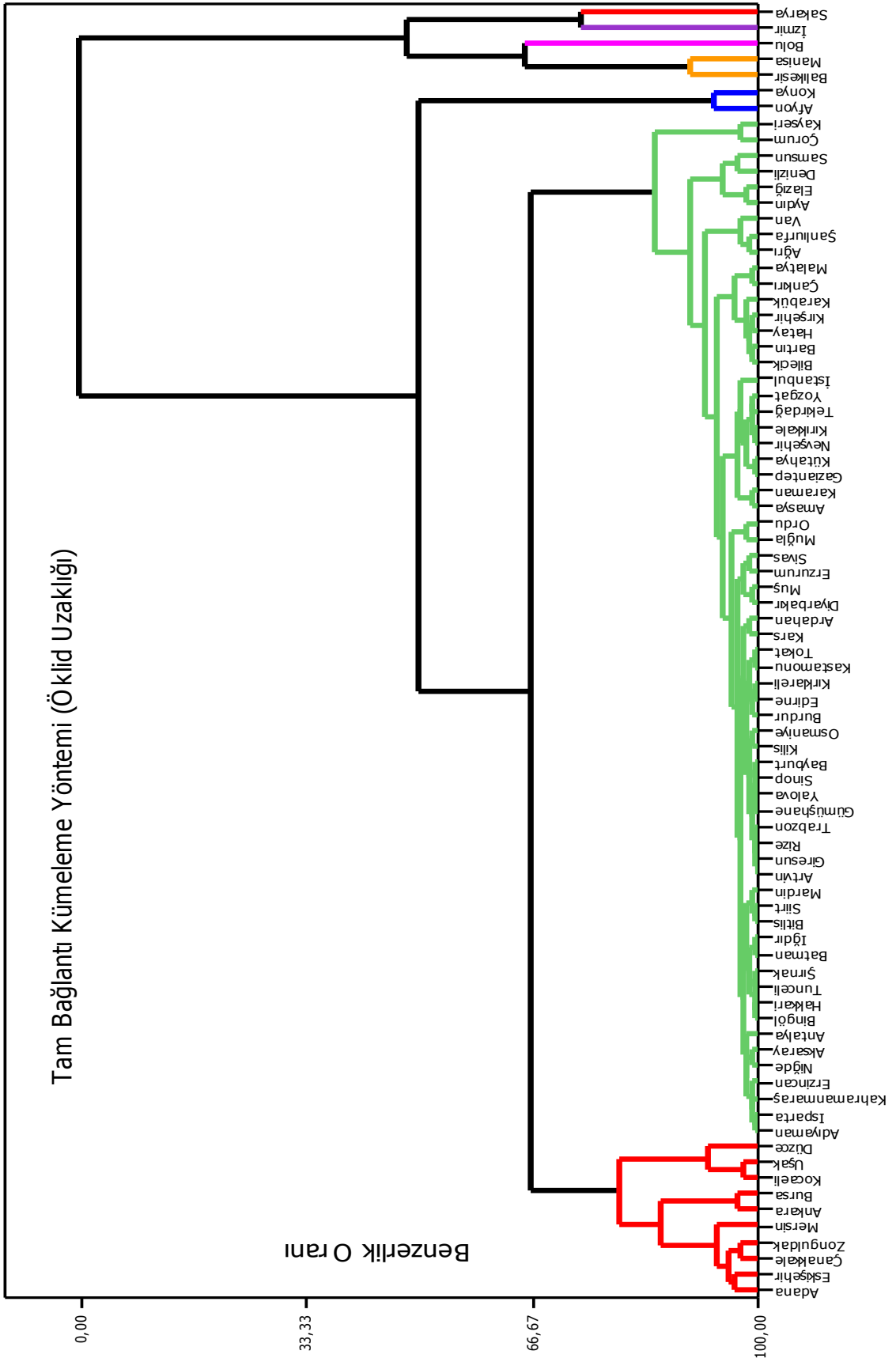


Ek Şekil 1. Ortalama Bağlantı Kümeleme Yöntemi (Öklid Uzaklığı)

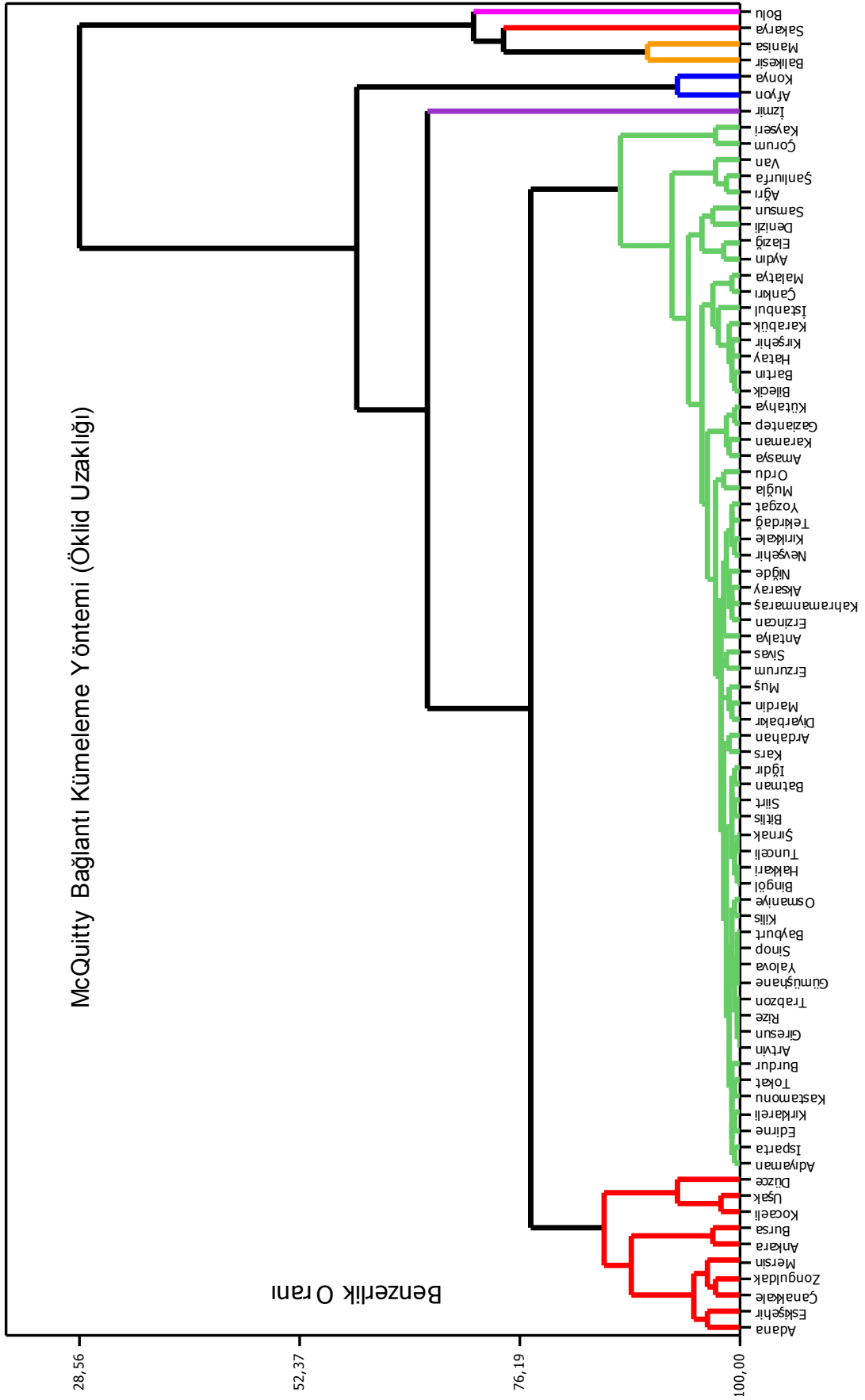




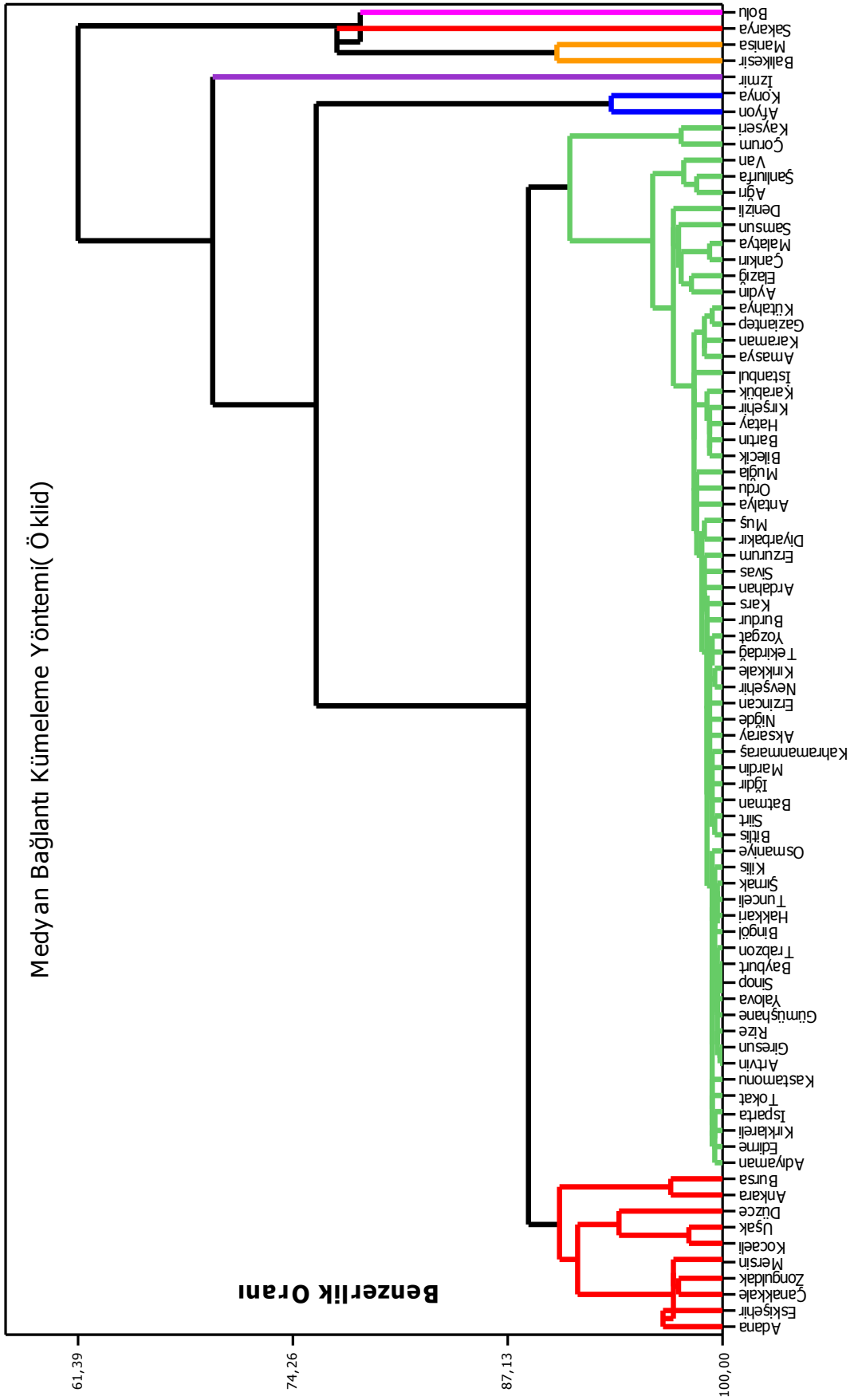
Ek Şekil 2. Merkezi Bağlantı Kümeleme Yöntemi (Öklid Uzaklığı)



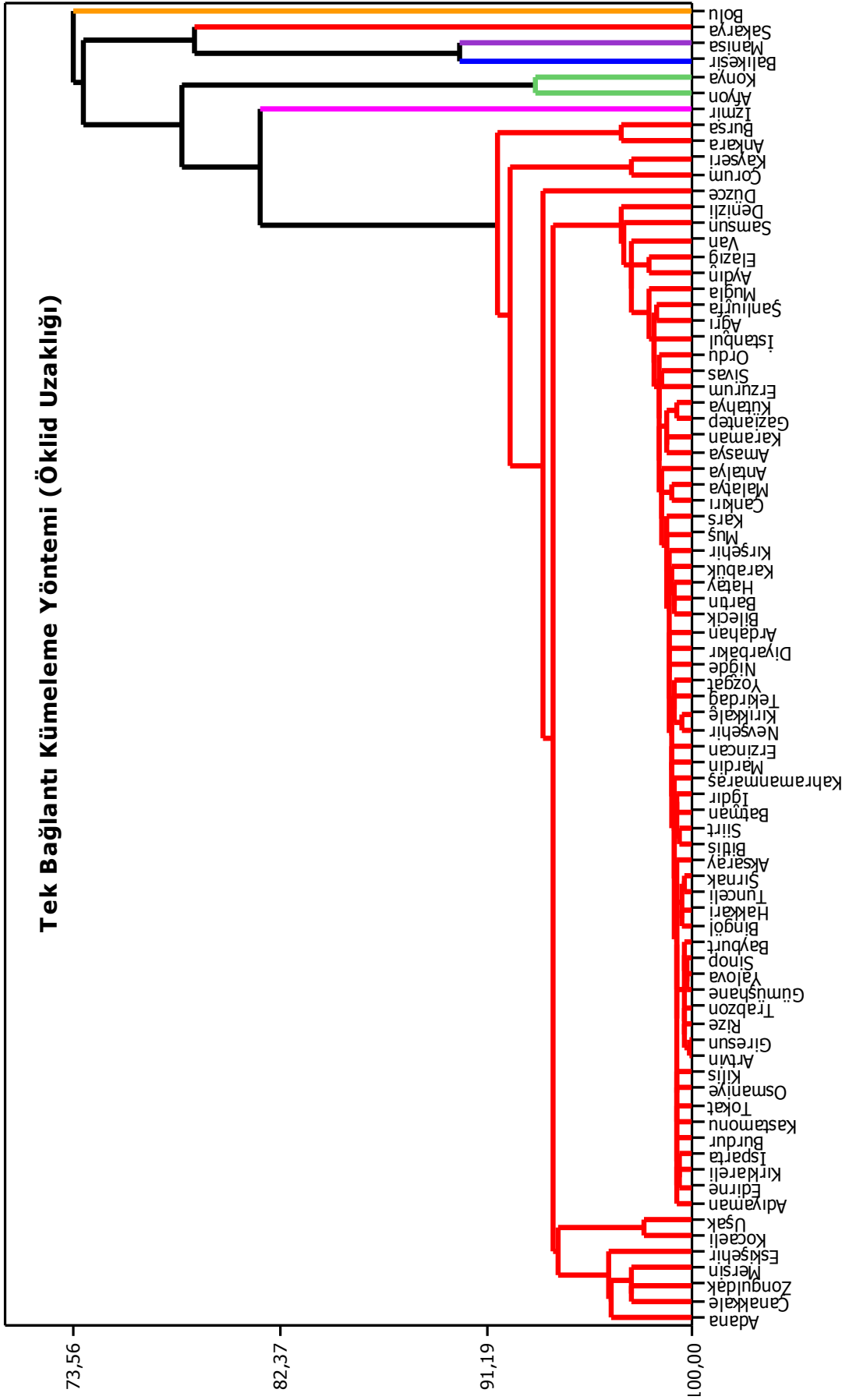
Ek Şekil 3. Tam Bağlantı Kümeleme Yöntemi (Öklid Uzaklığı)



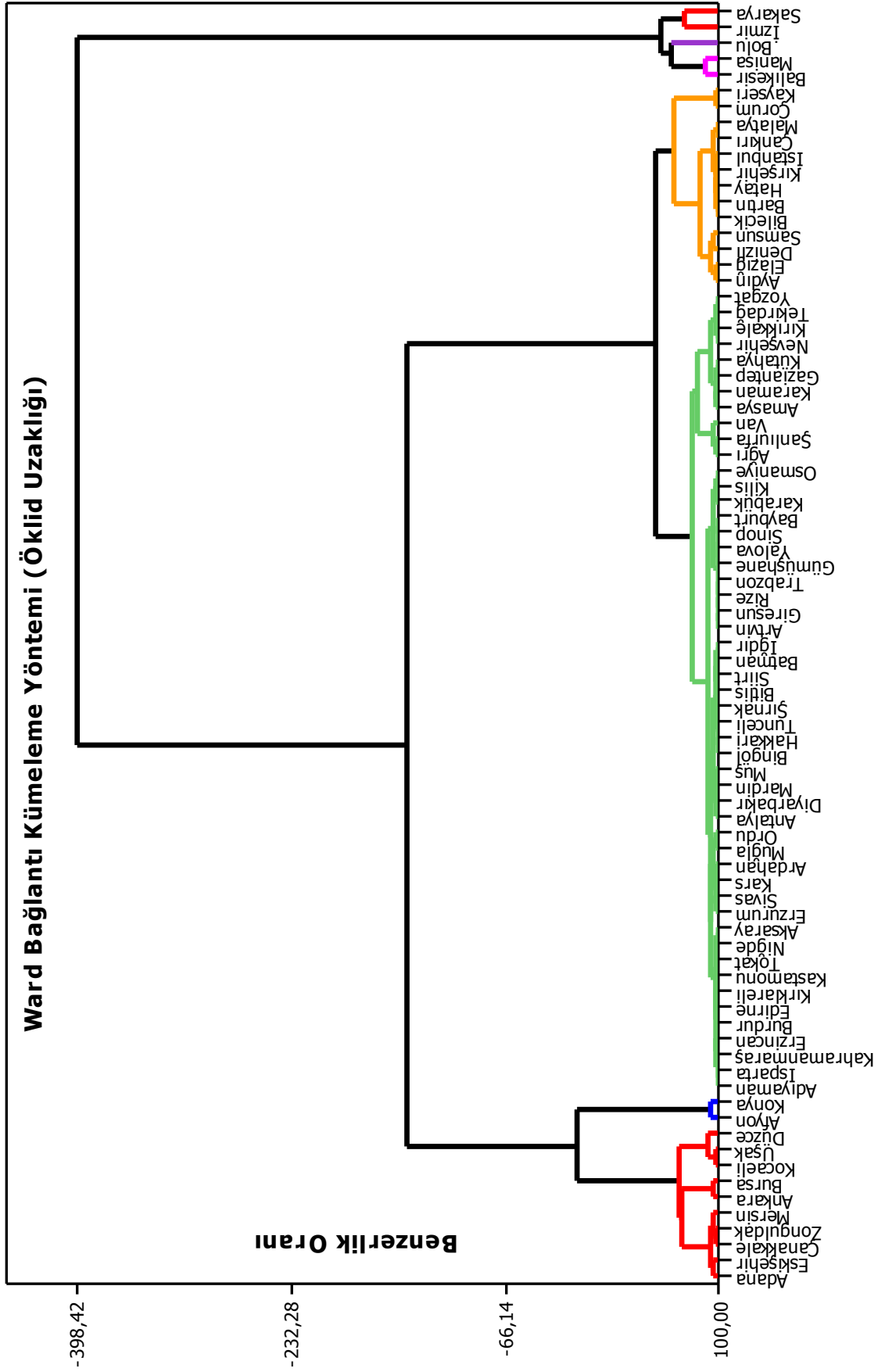
Ek Şekil 4. McQuitty Bağlantı Kümeleme Yöntemi (Öklid Uzaklığı)



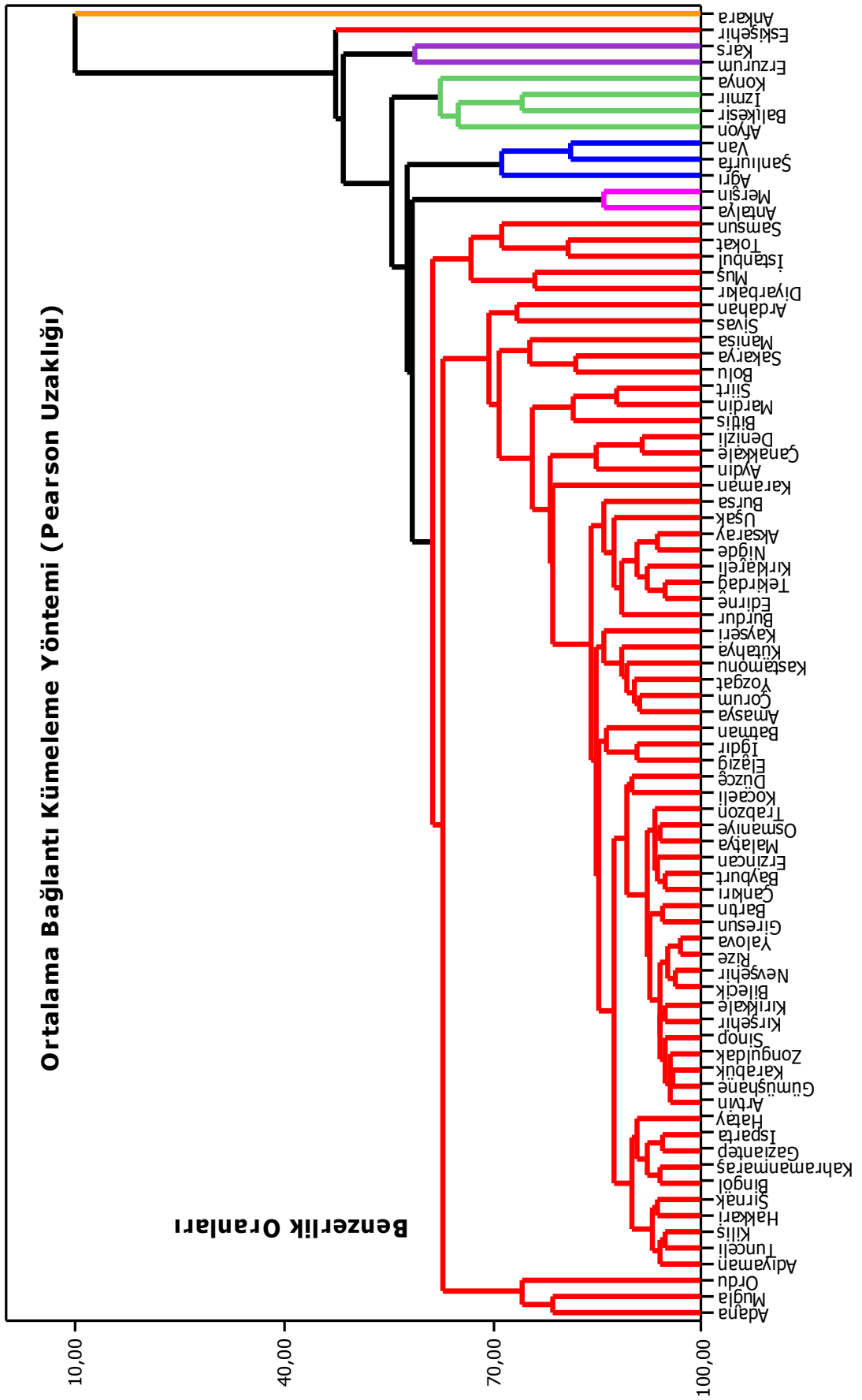
Ek Şekil 5. Medyan Bağlantı Kümeleme Yöntemi (Öklid Uzaklığı)



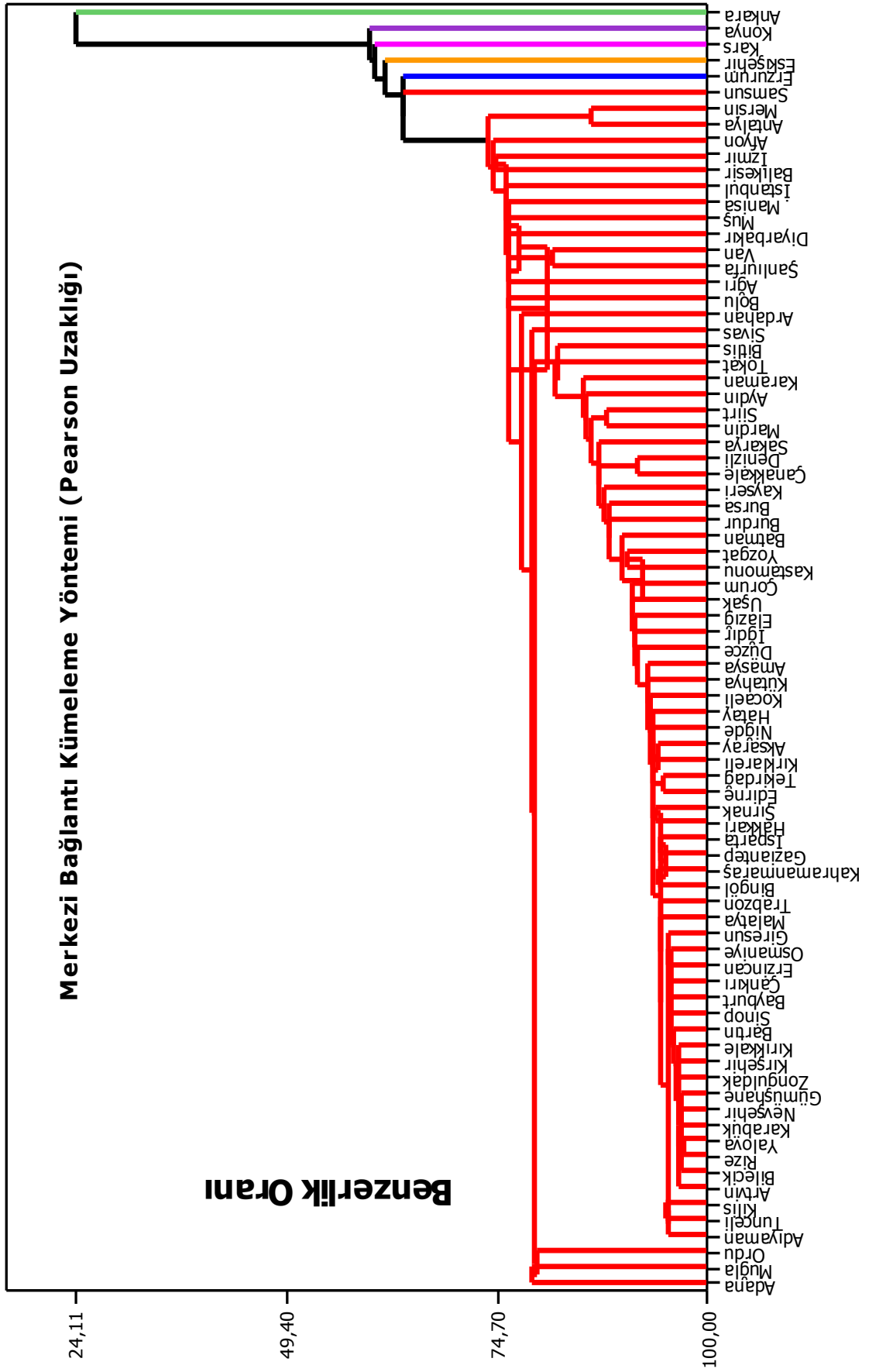
Ek Şekil 6. Tek Bağlantı Kümeleme Yöntemi (Öklid Uzaklığı)



Ek Şekil 7. Ward Bağlantı Kümeleme Yöntemi (Öklid Uzaklığı)

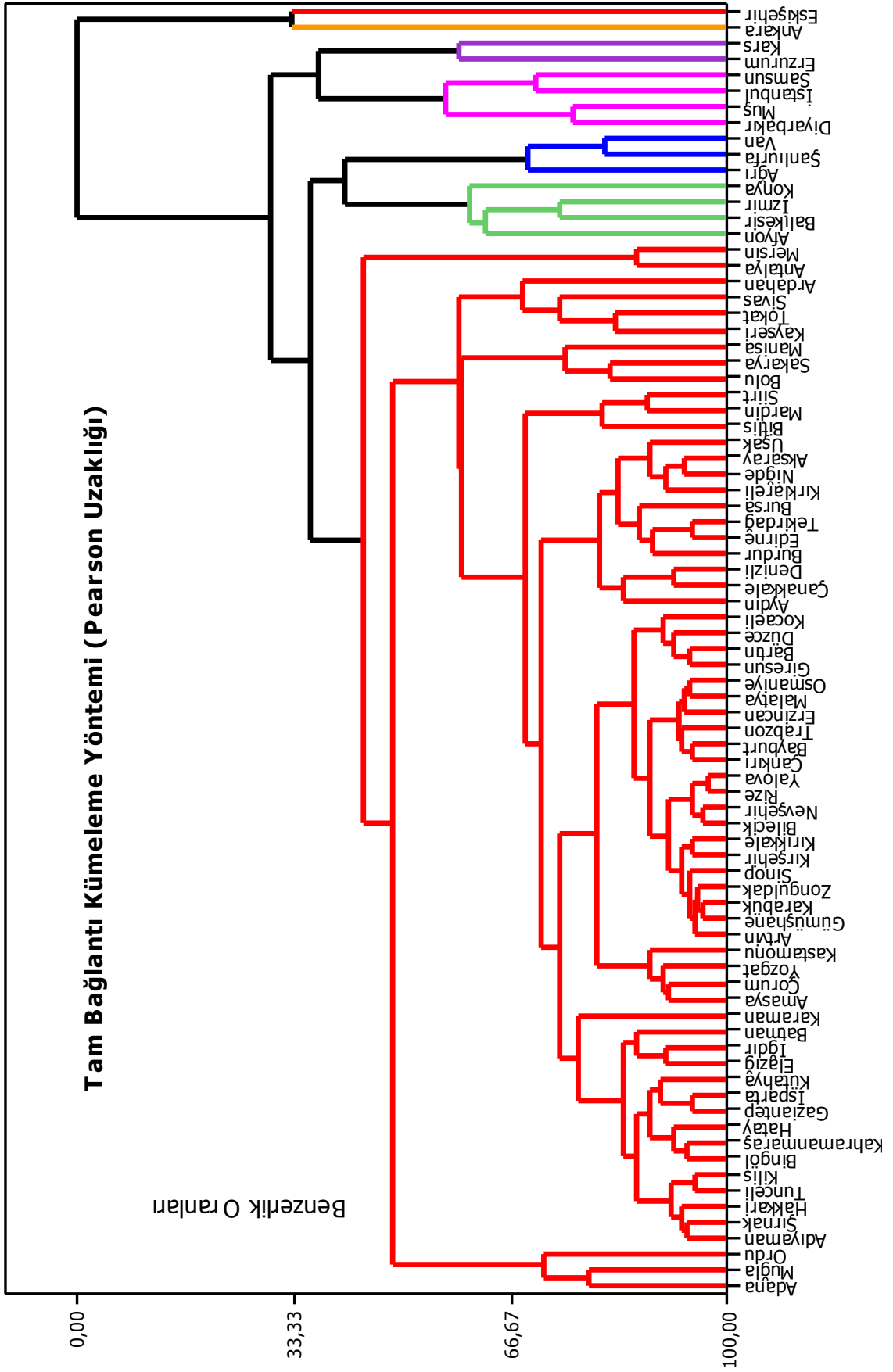


Ek Şekil 8. Ortalama Bağlantı Kümeleme Yöntemi (Pearson Uzaklığı)

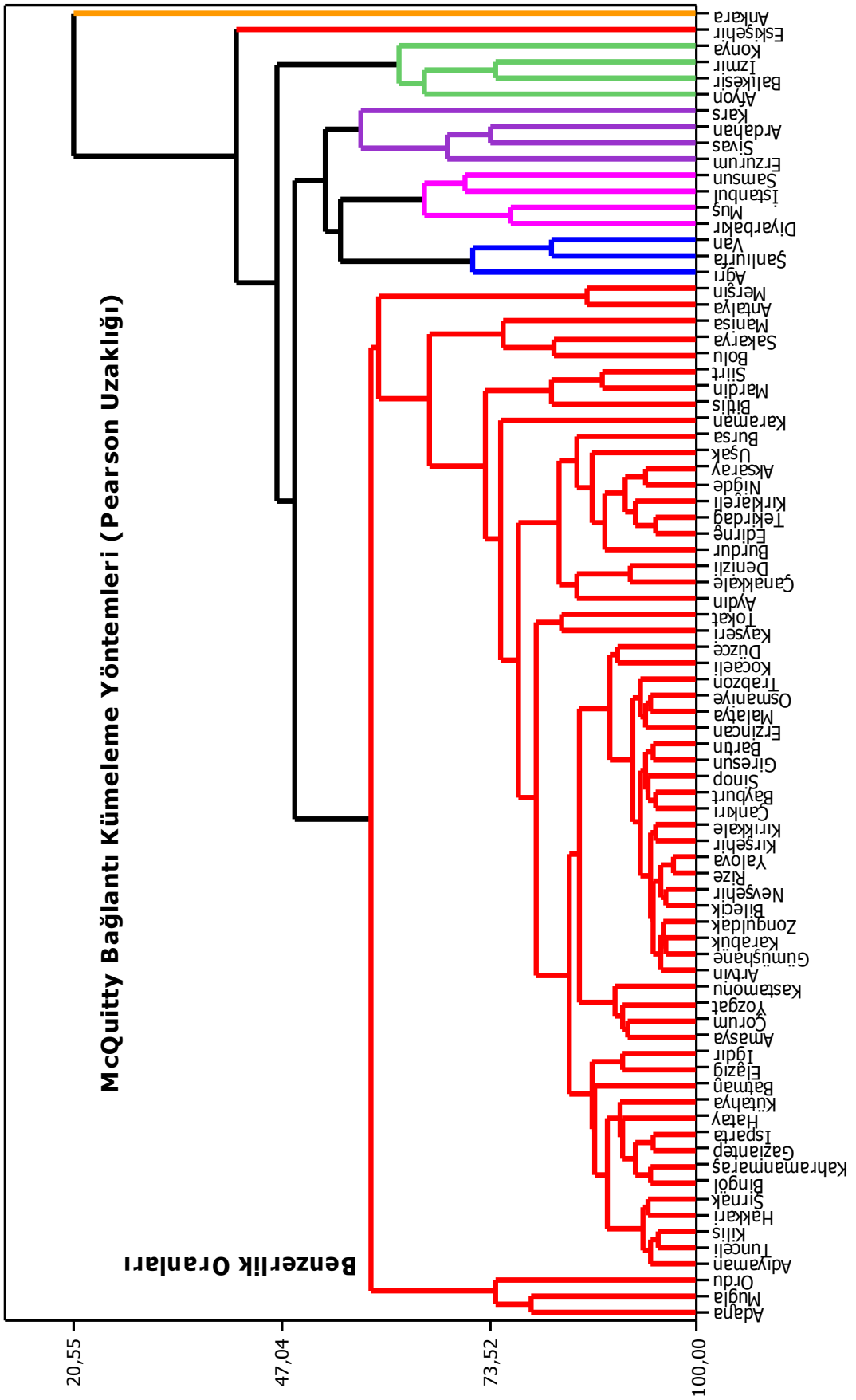


Ek Şekil 9. Merkezi Bağlantı Kümeleme Yöntemi (Pearson Uzaklığı)

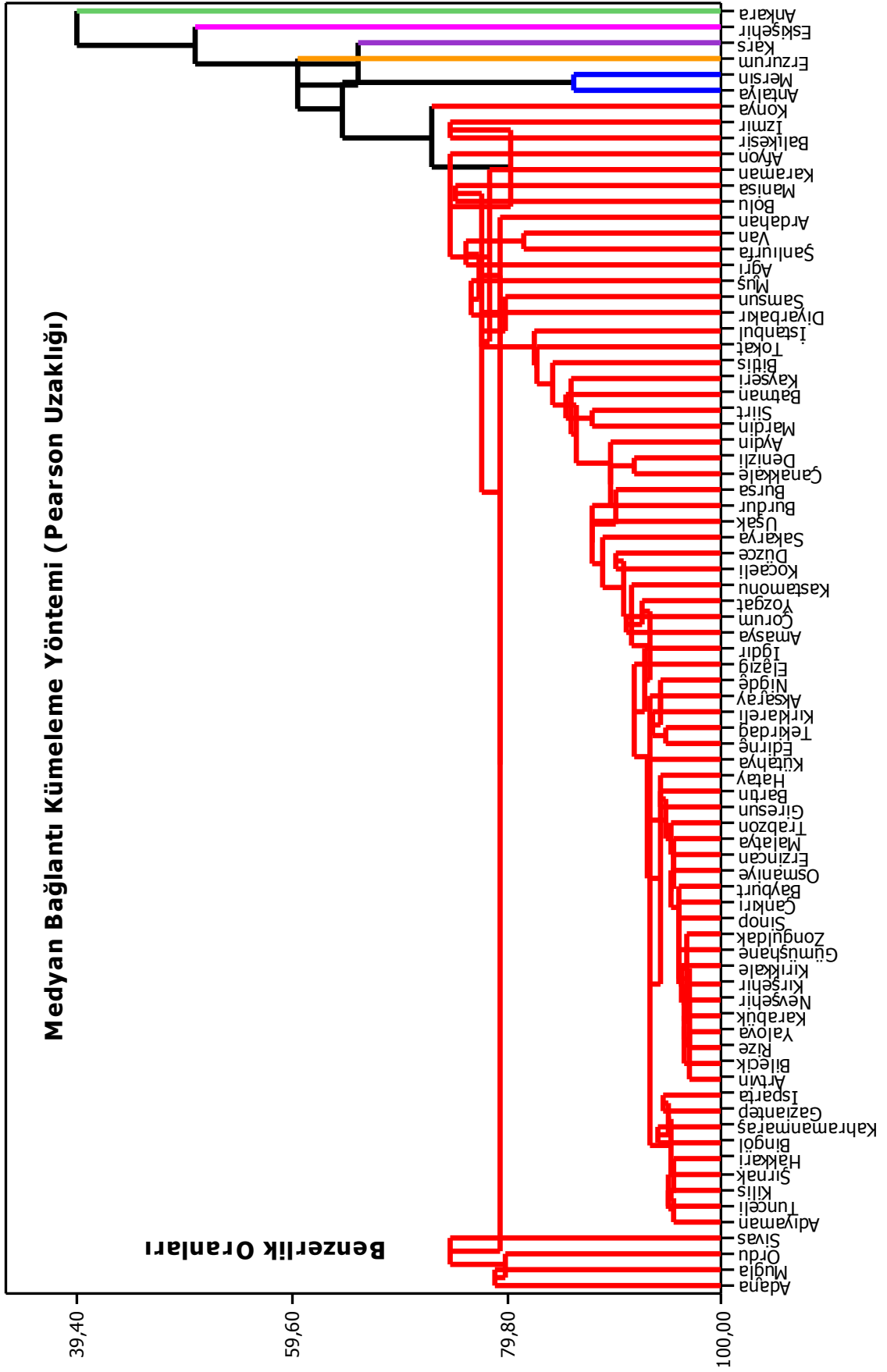




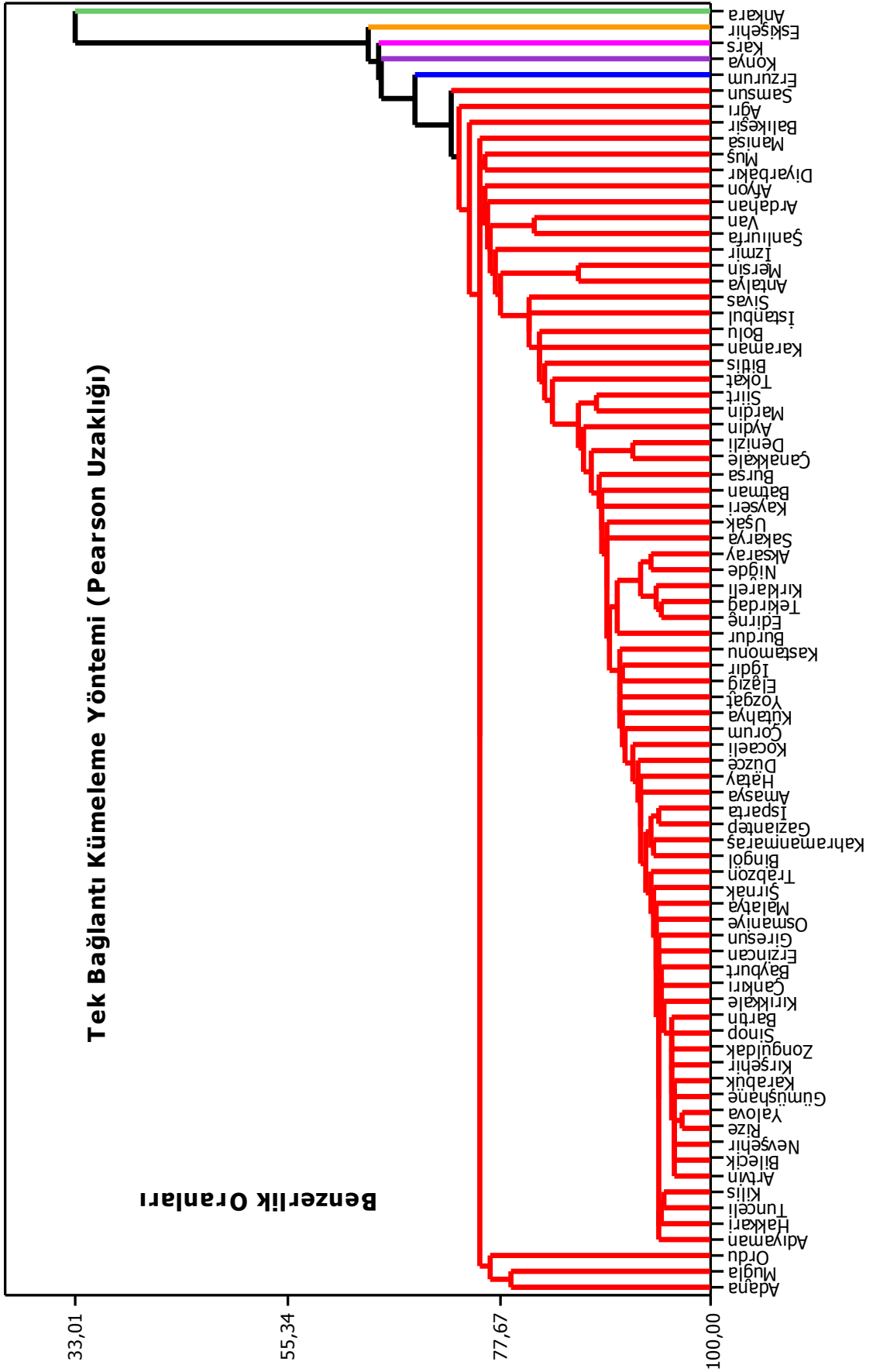
Ek Şekil 10. Tam Bağlantı Kümeleme Yöntemi (Pearson Uzaklığı)



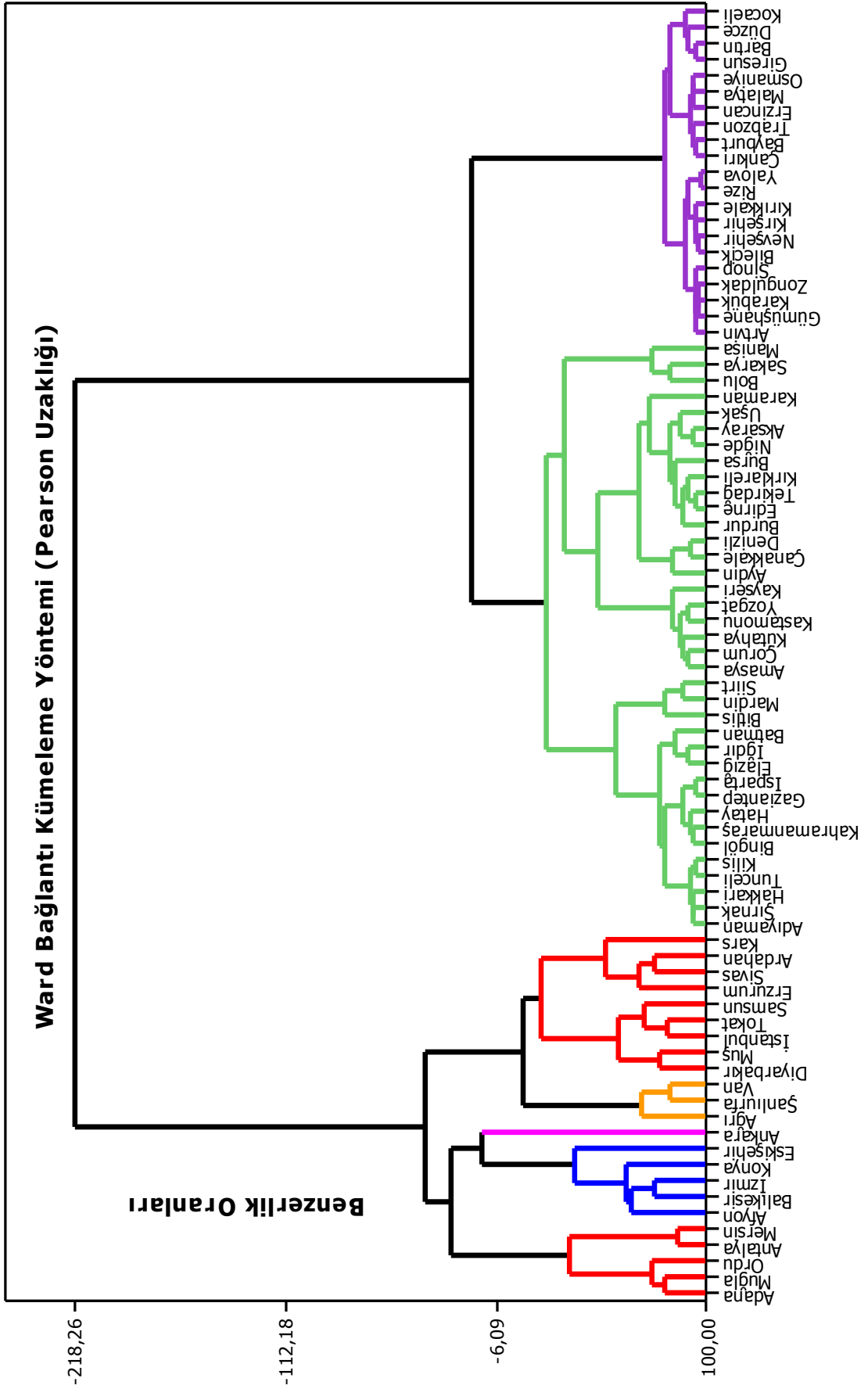
Ek Şekil 11. McQuitty Bağlantı Kümeleme Yöntemi (Pearson Uzaklığı)



Ek Şekil 12. Medyan Bağlantı Kümeleme Yöntemi (Pearson Uzaklığı)



Ek Şekil 13. Tek Bağlantı Kümeleme Yöntemi (Pearson Uzaklığı)



Ek Şekil 14. Ortalama Bağlantı Kümeleme Yöntemi (Pearson Uzaklığı)

Ek Tablo 1. Yerli Koyunculuk İstatistikleri (TUİK 2012 – İlk 25 İl)

İl Adı	Koyun (Yerli Yetişkin)	Genç-Yavru	Sağılan hayvan sayısı (baş)	Süt (Ton)	Yün kıl tiftik (ton)
Van	1.531.180	615.959	965.606	82.077	3.334,51
Konya	1.332.472	301.944	957.340	75.630	2.647,75
Ağrı	1.003.366	182.343	712.597	53.445	2.211,35
Şanlıurfa	926.934	434.437	640.262	55.063	2.534,87
Muş	641.248	203.749	455.280	35.512	1.444,95
Diyarbakır	512.809	207.239	310.192	25.436	1.247,84
Balıkesir	473.463	118.977	285.621	22.278	1.175,40
Bitlis	460.863	146.841	266.244	18.903	1.112,10
Manisa	435.500	163.268	237.905	18.081	1.014,31
Elazığ	433.812	144.264	288.264	22.196	1.121,47
İğdır	433.429	127.111	276.879	23.258	969,734
Afyon	418.305	151.547	236.905	18.479	926,01
Erzurum	416.751	97.259	265.062	19.880	1.028,02
Niğde	400.320	42.258	288.508	21.638	704,584
Mardin	397.662	152.183	238.225	17.867	972,126
Batman	364.039	141.902	236.825	18.946	996,704
Siirt	354.532	116.435	187.674	12.574	1.017,29
Sivas	351.459	85.791	252.309	17.662	868,816
İzmir	349.296	130.148	216.161	17.293	858,205
Ankara	349.129	76.367	216.109	16.424	811,846
Kayseri	333.212	118.346	208.218	17.074	903,116
Çanakkale	314.965	78.928	208.594	15.853	664,104
Aksaray	308.082	44.323	179.684	15.093	528,608
Mersin	301.162	88.222	208.044	16.019	671,687
Denizli	299.245	103.274	208.259	17.910	801,013

Ek Tablo 2. Merinos Koyunculuk İstatistikleri (TUİK 2012 – İlk 25 İl)

İl Adı	Koyun (Merinos) Yetişkin	Genç- Yavru	Sağılan hayvan sayısı (baş)	Süt (Ton)	Yün kıl tiftik (ton)
Ankara	279.979	65.473	174.634	9.256	1.101,30
Eskişehir	238.207	102.589	146.795	6.165	1.158,71
Afyon	106.149	41.297	54.884	2.470	442,338
Karaman	102.429	6.265	61.217	2.877	326,082
Balıkesir	79.735	23.516	47.178	2.595	249,248
Konya	74.924	24.315	49.528	1.981	297,717
Bursa	56.179	18.284	37.540	1.802	223,389
Edirne	29.906	12.602	17.482	944	138,151
Burdur	24.789	15.901	14.174	638	122,07
Antalya	24.662	11.666	14.453	867	105,932
Kütahya	17.457	9.720	8.931	482	99,495
Tekirdağ	15.449	4.126	9.182	459	57,688
Muğla	13.001	8.013	6.720	336	63,042
Kastamonu	11.438	4.418	5.970	299	54,37
Bolu	9.679	4.399	4.695	188	51,54
Kocaeli	7.096	2.980	4.239	170	29,694
Aksaray	5.853	882	3.607	159	20,205
Bilecik	5.360	2.908	2.601	133	26,681
Karabük	5.303	1.914	2.833	122	24,747
Kırşehir	5.109	1.200	3.147	142	18,927
Çanakkale	4.934	1.681	3.090	170	18,509
Mersin	3.855	70	2.911	143	11,445
Şırnak	3.850	3.050	1.535	78	23,66
Aydın	3.015	962	1.718	84	11,128
Osmaniye	2.376	1.175	1.454	87	10,355

Ek Tablo 3. Kıl Keçisi İstatistikleri (TÜİK 2012 – İlk 25 İl)

İl Adı	Keçi (kıl) Yetişkin	Genç- Yavru	Sağılan hayvan sayısı (baş)	Süt (Ton)	Yün kıl tiftik (ton)
Mersin	480.569	179.756	300.674	34.277	240,285
Antalya	399.204	113.417	226.399	26.942	263,475
Adana	255.278	93.450	126.801	14.582	145,508
Mardin	245.610	94.779	146.381	15.956	122,805
Bitlis	234.197	62.956	131.631	12.768	179,161
Siirt	213.617	84.285	112.183	12.228	106,809
Çanakkale	177.884	41.014	121.320	12.860	88,942
Denizli	171.265	25.518	106.698	10.883	121,084
Diyarbakır	154.480	65.960	80.113	8.412	92,688
Balıkesir	146.549	41.119	86.771	10.760	73,275
Muğla	136.489	45.826	72.939	8.242	68,245
Van	135.014	90.346	71.832	7.758	102,611
Konya	134.266	47.171	79.880	7.828	100,7
Isparta	132.778	36.601	79.619	8.201	73,028
Gaziantep	130.266	38.118	75.097	8.110	78,16
Şanlıurfa	127.847	48.278	76.515	8.111	78,882
İzmir	127.383	56.281	60.583	6.422	63,692
Kahramanmaraş	123.890	72.390	67.435	7.283	74,334
Manisa	121.763	55.644	51.438	5.298	75,493
Muş	121.548	34.683	79.786	9.096	72,929
Şırnak	120.980	58.465	65.165	6.842	64,119
Bingöl	107.106	33.694	67.081	7.178	60,515
Batman	103.531	35.388	77.752	7.309	51,766
Adıyaman	96.287	28.825	60.759	7.048	51,995
Ağrı	94.323	27.390	48.177	5.203	46,784



Ek Tablo 4. Tiftik Keçisi İstatistikleri (TUİK 2012 – İlk 25 İl)

İl Adı	Keçi(Tiftik) Yetişkin	Genç- Yavru	Sağılan hayvan sayısı (baş)	Süt (Ton)	Yün kıl tiftik (ton)
Ankara	64.039	21.710	31.867	1.243	115,847
Karaman	10.893	835	5.216	219	13,616
Siirt	7.780	4.332	5.453	104	12,993
Eskişehir	7.539	2.747	3.951	119	15,078
Bolu	7.328	1.659	5.964	250	11,923
Kastamonu	4.749	1.163	2.001	50	6,378
Kırıkkale	3.363	1.547	1.762	51	6,084
Mardin	3.310	1.800	1.748	42	4,217
Konya	2.417	1.153	1.054	45	2,417
İzmir	1.479	584	760	14	2,676
Çankırı	1.431	358	821	32	2,454
Şırnak	1.040	500	333	10	1,352
Çorum	902	369	401	9	1,139
Afyon	664	347	294	5	1,035
Aksaray	628	71	356	11	0,669
Bilecik	357	143	190	4	0,545
Kütahya	275	75	185	8	0,57
Ardahan	165	20	90	2	0,199
Karabük	131	43	59	1	0,132
Giresun	60	25	30	1	0,071
Niğde	35	0	29	1	0,088
Samsun	8	2	0	0	0,009
İstanbul	8	0	0	0	0,01
Bursa	6	12	1	0	0,007
Mersin	0	0	0	0	0

Ek Tablo 5. Kültür Sığır İstatistikleri (TÜİK 2012 – İlk 25 İl)

İl Adı	Yetişkin (Sığır Kültür)	Genç- Yavru	Sağılan hayvan sayısı (baş)	Süt (Ton)
Balıkesir	285.760	128.732	154.518	611.426
Konya	284.777	114.122	142.363	568.454
İzmir	264.274	80.229	144.825	571.768
Aydın	169.656	71.865	86.597	343.356
Denizli	159.512	61.235	74.318	295.193
Afyon	151.652	63.702	75.323	299.031
Burdur	130.268	44.725	78.197	308.485
Çanakkale	121.043	47.492	69.365	274.684
Kırklareli	99.273	33.745	59.429	232.187
Aksaray	97.095	24.268	58.980	233.855
Tekirdağ	94.805	34.168	49.172	193.394
Bursa	93.419	37.373	48.022	186.133
Edirne	88.860	32.661	52.775	209.832
Kayseri	87.055	28.719	41.167	160.387
Sivas	78.669	19.681	53.858	210.371
Niğde	77.770	18.732	46.436	183.144
Kastamonu	70.238	21.527	38.720	152.091
Ankara	67.577	19.670	27.819	108.356
Manisa	62.561	21.477	33.630	127.728
Isparta	62.373	17.000	36.854	146.864
Kütahya	62.134	15.960	33.732	132.264
Yozgat	61.624	23.179	30.090	110.489
Diyarbakır	58.937	13.896	25.528	99.226
Uşak	58.732	50.031	35.413	140.908
Gaziantep	58.715	10.966	27.105	101.453

Ek Tablo 6. Melez Sığır İstatistikleri (TÜİK 2012 – İlk 25 İl)

İl Adı	Yetişkin Sığır (Melez)	Genç-Yavru	Sağılan hayvan sayısı (baş)	Süt (Ton)
Erzurum	312.754	135.037	164.226	477.899
Ardahan	212.910	42.886	106.842	297.554
Sivas	179.395	45.136	119.353	331.442
Kars	160.417	48.176	94.569	280.114
Konya	142.874	50.566	65.844	182.783
Adana	134.783	48.186	72.479	188.952
Samsun	131.952	48.813	68.765	188.417
Kayseri	102.364	34.748	45.595	126.162
Ağrı	98.386	19.688	44.029	118.614
Ankara	95.585	25.588	33.023	82.723
Muş	93.355	30.110	45.233	122.130
Çorum	89.237	23.200	43.930	118.742
Manisa	85.346	32.078	41.466	109.844
Ordu	83.521	21.332	53.113	147.919
Diyarbakır	83.219	26.140	34.054	94.873
İzmir	82.754	25.539	41.984	118.143
Tokat	81.884	32.124	39.203	108.867
Kastamonu	80.912	24.249	46.186	129.137
Yozgat	77.912	20.808	39.889	106.305
Balıkesir	77.421	32.705	31.671	83.453
Hatay	71.277	23.475	35.955	102.976
Sakarya	70.825	23.122	37.020	100.548
Trabzon	64.556	16.309	38.423	103.626
Malatya	63.450	13.802	39.183	107.988
Muğla	62.917	19.393	27.654	78.400

Ek Tablo 7. Yerli Sığır İstatistikleri (TUİK 2012 – İlk 25 İl)

İl Adı	Yetişkin(Sığır yerli)	Genç-Yavru	Sağılan hayvan sayısı (baş)	Süt (Ton)
Kars	187.789	42.584	106.042	139.551
Ağrı	140.205	46.857	63.561	87.078
Diyarbakır	113.240	35.550	43.578	54.734
Erzurum	86.338	21.415	46.741	61.137
Muş	83.692	20.088	52.330	71.849
Van	69.399	25.781	33.833	45.506
Samsun	67.412	23.508	32.031	39.014
Yozgat	51.932	15.646	26.571	35.340
Ankara	51.576	14.954	17.768	23.187
Kastamonu	49.746	14.046	25.797	34.182
Mardin	49.255	5.633	27.659	38.059
Tokat	49.196	17.830	22.948	31.783
Şanlıurfa	46.584	16.096	22.987	28.619
Amasya	42.118	8.552	18.220	23.285
Konya	38.615	15.629	16.299	21.548
Ardahan	32.015	16.698	16.275	21.239
Iğdır	31.186	6.370	19.225	26.339
Çorum	30.717	8.510	16.556	22.450
Sinop	30.180	7.184	14.428	19.766
Aydın	27.279	10.975	12.416	17.345
Gümüşhane	23.167	6.067	11.032	14.133
Giresun	22.527	3.151	10.116	13.191
Afyon	21.070	8.859	10.462	13.151
Bitlis	20.769	4.113	11.980	14.926
Düzce	19.534	4.985	10.314	12.840

Ek Tablo 8. Manda İstatistikleri (TÜİK 2012 – İlk 25 İl)

İl Adı	Yetişkin(Manda)	Genç-Yavru	Sağılan hayvan sayısı (baş)	Süt (Ton)
Samsun	11.112	2.929	6.373	6.309
İstanbul	8.384	2.129	4.993	5.362
Diyarbakır	7.258	1.647	4.802	4.562
Tokat	5.870	1.939	2.843	2.698
Muş	5.168	711	3.331	3.221
Bitlis	4.281	1.318	1.540	1.416
Afyon	3.885	1.200	2.132	2.346
Sivas	3.193	398	2.168	1.951
Kayseri	2.686	972	1.381	1.367
Düzce	2.607	748	1.141	1.064
Amasya	2.193	612	1.045	1.088
Balıkesir	1.785	410	1.253	1.269
Giresun	1.727	455	888	871
Bartın	1.505	287	921	838
Çorum	1.461	362	841	818
Kütahya	1.442	333	821	801
Yozgat	1.431	385	872	872
Kırklareli	1.336	211	841	795
Kocaeli	1.163	248	630	567
Çankırı	1.009	264	485	501
Bolu	935	346	475	478
Erzurum	898	215	404	392
İğdır	857	416	407	407
Karabük	843	102	243	233
Erzincan	832	310	424	411

Ek Tablo 9. Et ve Yumurta Tavuğu İstatistikleri (TUIK 2012 – İlk 25 İl)

İl Adı	Et Tavuğu (Adet)	İl Adı	Yumurta Tavuğu
Bolu	25.910.750	Afyon	11.434.700
Balıkesir	21.177.966	Konya	10.053.182
Manisa	21.099.077	Manisa	8.763.265
Sakarya	18.360.900	Balıkesir	6.066.074
İzmir	11.959.755	İzmir	4.275.022
Düzce	8.695.292	Çorum	4.207.032
Uşak	6.969.864	Bursa	4.203.460
Kocaeli	6.692.450	Kayseri	3.578.723
Mersin	5.266.356	Ankara	3.516.538
Ankara	4.785.360	Denizli	1.646.361
Çanakkale	4.738.222	Amasya	1.437.078
Bursa	4.423.957	Samsun	1.332.145
Eskişehir	4.250.985	Karaman	1.274.591
Zonguldak	4.221.400	Eskişehir	1.266.037
Adana	3.674.000	Kütahya	1.049.632
Aydın	2.082.150	Gaziantep	1.017.973
Elazığ	2.025.881	Sakarya	899.861
Samsun	1.835.950	İstanbul	834.853
Malatya	1.401.026	Adana	742.890
Çankırı	1.242.000	Nevşehir	727.900
Denizli	1.161.700	Tekirdağ	705.200
Hatay	911.060	Yozgat	645.189
Bilecik	893.950	Aydın	633.097
Bartın	712.500	Kırıkkale	623.540
Kırşehir	697.000	Muğla	573.787

Ek Tablo 10. Kaz, Ördek ve Hindi İstatistikleri (TUİK 2012 – İlk 25 İl)

İl Adı	Kaz	İl Adı	Ördek	İl Adı	Hindi
Kars	111.150	Muş	32.147	Bolu	436.988
Muş	89.018	Samsun	27.534	Manisa	426.731
Ardahan	74.004	Balıkesir	21.301	İzmir	257.697
Batman	36.302	Diyarbakır	15.434	Sakarya	188.196
Afyon	32.300	Ağrı	13.780	Kocaeli	133.459
Ağrı	28.670	Edirne	13.510	Şanlıurfa	107.044
Samsun	24.293	Kars	12.030	Eskişehir	91.117
Yozgat	18.205	Tekirdağ	10.691	İstanbul	68.767
Kütahya	17.957	Afyon	9.655	Diyarbakır	64.766
Konya	16.395	Muğla	8.861	Muş	60.982
Diyarbakır	16.116	Çanakkale	8.393	Konya	59.513
Erzurum	13.332	Konya	8.392	Balıkesir	58.620
Eskişehir	10.749	Batman	8.209	Antalya	56.250
Edirne	9.460	Bursa	7.664	Ağrı	53.750
Şanlıurfa	9.405	Kırklareli	7.056	Bursa	53.703
Çorum	8.561	Aksaray	6.210	Mardin	43.741
Bolu	7.463	Hatay	5.700	Batman	41.868
Aksaray	7.188	Kahramanmaraş	5.335	Tekirdağ	39.070
Tokat	7.130	Amasya	5.132	Burdur	38.850
Erzincan	6.878	Ankara	4.908	Bingöl	27.263
Bingöl	6.385	Ardahan	4.709	Afyon	26.220
Çanakkale	6.159	Adana	4.454	Kars	24.400
Kahramanmaraş	6.155	Şanlıurfa	4.408	Samsun	21.373
Niğde	5.504	Erzurum	4.379	Gaziantep	19.333
Ankara	5.101	Bolu	4.341	Yozgat	18.616

Ek Tablo 11. Arıcılık İstatistikleri (TÜİK 2012 – İlk 25 İl)

İl Adı	Toplam kovan	Bal üretimi (ton)	Balmumu üretimi (ton)
Ordu	487.214	11.457,65	220,318
Muğla	710.949	10.765,26	570,537
Adana	428.059	8.320,71	450,376
Mersin	214.521	3.159,48	205,588
Aydın	208.550	2.907,30	114,495
İzmir	182.017	2.873,02	150,377
Balıkesir	140.625	2.485,45	73,363
Sivas	202.529	2.363,67	130,055
Antalya	195.188	2.354,77	152,533
Denizli	90.770	1.633,86	68,076
Bingöl	94.952	1.523,10	22
Siirt	89.573	1.350,42	24,467
Hakkari	123.303	1.296,04	31
Şanlıurfa	61.196	1.281,26	12,961
Trabzon	128.731	1.246,66	99,131
Erzurum	119.771	1.186,25	62,768
Samsun	76.286	1.185,39	69,276
Diyarbakır	74.993	1.178,24	39,371
Giresun	85.902	1.086,20	73,682
Hatay	80.093	1.074,18	61,14
Erzincan	74.058	1.073,55	45,157
Elazığ	96.696	1.070,12	104,077
Çanakkale	57.341	1.048,46	55,284
Van	61.933	991,636	83,214
Osmaniye	81.379	916,54	48,755



## ÖZGEÇMİŞ

1978 yılında Bingöl ili Genç ilçesinde doğdu. İlk ve ortaokulu Mersin’de, liseyi İstanbul Üsküdar Cumhuriyet Lisesi’nde tamamladı. 1995 yılında kazandığı Karadeniz Teknik Üniversitesi Fen Edebiyat Fakültesi Matematik Bölümünden 1999 yılında mezun oldu. 1999 yılında Bingöl ilinde matematik öğretmeni olarak göreve başladı. 2008 yılından itibaren Bingöl Milli Eğitim Müdürlüğü bünyesinde Fatih Projesi il koordinatörü olarak görevine devam etmektedir. 2011 yılında kayıt yaptırdığı Bingöl Üniversitesi Fen Bilimleri Enstitüsü Zootekni Ana Bilim Dalı Yüksek Lisans Programına devam etmektedir. Evli ve iki çocuk babasıdır.