

**YILDIZ TEKNİK ÜNİVERSİTESİ  
FEN BİLİMLERİ ENSTİTÜSÜ**

## **CHAID ANALİZİ VE BİR UYGULAMA**

**İstatistikçi Gamze PEHLİVAN**

**FBE İstatistik Anabilim Dalında Hazırlanan**

**YÜKSEK LİSANS TEZİ**

**Tez Danışmanı: Yrd. Doç. Dr. Doğan YILDIZ**

**İSTANBUL,2006**

# İÇİNDEKİLER

Sayfa

|                                                                    |      |
|--------------------------------------------------------------------|------|
| SİMGE LİSTESİ .....                                                | iv   |
| KISALTMA LİSTESİ .....                                             | v    |
| ŞEKİL LİSTESİ .....                                                | vi   |
| ÇİZELGE LİSTESİ .....                                              | vii  |
| ÖNSÖZ .....                                                        | viii |
| ÖZET .....                                                         | ix   |
| ABSTRACT .....                                                     | x    |
| 1. GİRİŞ .....                                                     | 11   |
| 2. VERİ MADENCİLİĞİ'NDE KARAR AĞAÇLARI .....                       | 13   |
| 2.1 En Sık Kullanılan Karar Ağaçları Algoritmaları .....           | 20   |
| 2.1.1 C&RT (Classification and Regression Tree ) .....             | 20   |
| 2.1.2 QUEST ( Quick, Unbiased, Efficient, Statistical Tree ) ..... | 21   |
| 3. AID ANALİZİ .....                                               | 23   |
| 3.1 AID Analizinin Genel Yapısı .....                              | 23   |
| 3.2 AID Analizinin Algoritması .....                               | 26   |
| 3.3 AID Analizinin Limitleri ve Dezavantajları .....               | 28   |
| 4. CHAID ANALİZİ .....                                             | 31   |
| 4.1 CHAID Analizinin Genel Yapısı .....                            | 31   |
| 4.2 CHAID Analizinin Algoritması .....                             | 32   |
| 4.2.1 Birleştirme .....                                            | 33   |
| 4.2.2 Dağıtma .....                                                | 33   |
| 4.2.3 Durdurma .....                                               | 33   |
| 4.3 Açıklayıcı Değişkenler Önemliliği .....                        | 34   |
| 4.3.1 Monoton Açıklayıcı Değişkenler .....                         | 34   |
| 4.3.2 Serbest Açıklayıcı Değişkenler .....                         | 34   |
| 4.3.3 Hareketli Açıklayıcı Değişkenler .....                       | 35   |
| 5. AID ANALİZİ ile CHAID ANALİZİNİN KARŞILAŞTIRILMASI .....        | 36   |
| 6. UYGULAMA .....                                                  | 38   |

|                                                                                                                                                                                                                      |    |
|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----|
| 7. SONUÇLAR VE DEĞERLENDİRME .....                                                                                                                                                                                   | 40 |
| 7.1 CHAID Analizi 1.....                                                                                                                                                                                             | 40 |
| 7.2 CHAID Analizi 2.....                                                                                                                                                                                             | 43 |
| 7.3 CHAID Analizi 3.....                                                                                                                                                                                             | 45 |
| KAYNAKLAR.....                                                                                                                                                                                                       | 51 |
| Ek 1 Anket Soru Formu.....                                                                                                                                                                                           | 52 |
| EK2 Eğitim, İşe Alma, Performans Değerlendirme, Ücretlendirme, Yan Faydalar, Kariyer Planlanma, Özlük İşleri Ve Bordrolama, Kurum İçi İletişim, Bölge Adı ve Unvanlar ile Faaliyet Alanına Göre Ağaç Diyagramı ..... | 56 |
| EK3 İnsan Kaynakları ile İlgili Görüşler, Bölge Adı ve Unvanlar ile Faaliyet Alanına Göre Ağaç Diyagramı.....                                                                                                        | 60 |
| EK4 Bütün Değişkenler ile Faaliyet Alanına Göre Ağaç Diyagramı.....                                                                                                                                                  | 64 |
| ÖZGEÇMİŞ.....                                                                                                                                                                                                        | 69 |

## SİMGE LİSTESİ

|             |                                                                         |
|-------------|-------------------------------------------------------------------------|
| B           | Bonferroni çarpanı                                                      |
| c           | Açıklayıcı değişkenin kategori sayısı                                   |
| e           | Hata terimi                                                             |
| j           | = 2,3.....c                                                             |
| n           | Gözlem sayısı                                                           |
| p           | Önem istatistiği                                                        |
| $p_j$       | Açıklayıcı değişkenin j.satırının belli bir sınıfa dahil olma olasılığı |
| $T_j^{(i)}$ | j x d tablosunu oluşturmadaki i. metot için $\chi^2$ istatistiği        |
| x           | Açıklayıcı değişken (predictor)                                         |
| y           | Bağımlı değişken ( hedef değişken )                                     |
| $\chi^2$    | Ki-Kare test istatistiği                                                |

## **KISALTMA LİSTESİ**

|       |                                              |
|-------|----------------------------------------------|
| AID   | Otomatik Etkileşim Belirleme Analizi         |
| CHAID | Otomatik Ki-Kare Etkileşim Belirleme Analizi |
| C&RT  | Sınıflama ve Regresyon Ağaçları (CART)       |
| QUEST | Hızlı, Yansız, Etkili İstatistiksel Ağaçlar  |

## ŞEKİL LİSTESİ

|                                                                                 | Sayfa |
|---------------------------------------------------------------------------------|-------|
| Şekil 3.1 Okuma süreleri üzerinde; cinsiyet ve eğitim arasındaki etkileşim..... | 24    |

## ÇİZELGE LİSTESİ

|                                                                                                                               | Sayfa |
|-------------------------------------------------------------------------------------------------------------------------------|-------|
| Çizelge 6.1 Personelin bölgelere göre dağılımı.....                                                                           | 38    |
| Çizelge 7.1 C1a ile C8b arasındaki cevapların ve bölge adı, unvanların açıklayıcı değişken olması durumunda risk matrisi..... | 42    |
| Çizelge 7.2 S9 ile S29 arasındaki cevapların ve bölge adı, unvanların açıklayıcı değişken olması durumunda risk matrisi ..... | 45    |

## **ÖNSÖZ**

İstatistik eğitimim boyunca ve tez çalışmam esnasında her zaman desteği ve yardımı ile yoluma devam etmem gerektiğini hatırlatan ve beni cesaretlendiren hocam Yrd. Doç. Dr. Doğan Yıldız' a öncelikle teşekkür etmek istiyorum.

Tezime devam etme konusunda her türlü yardımını bütün inceliği, anlayışı, sıcaklığı ile gösteren asistan arkadaşım Fatma Noyan'a teşekkür ediyorum.

İstatistiği bana sevdiren, her zaman yol gösterici, yardımcı, destekleyici tavırları ile eğitimimi tamamlamama katkısı olan tüm hocalarıma teşekkürü bir borç bilirim.

Her ne olursa olsun benimle olan ve olacak, her türlü başarıma katkıları büyük olan ablam Emine'ye ve babama, anneme teşekkür etmek istiyorum.



## ÖZET

Bilimsel çalışmalarda önemli istatistiksel problemlerden biri olayı önemli derece etkileyen faktörleri veya bu faktörlerin hangi seviyesinde etkinin yüksek olduğunun belirlenmesidir. Veri madenciliğinde karar ağaçları metotları bu problemi çözmek için geniş veri setlerine uygulanması, görsel açıdan kolay anlaşılır olması, varsayımlarının diğer istatistiksel yöntemlere göre az olması gibi nedenlerden dolayı tercih edilmektedir. Karar ağaçları içinde CHAID kategorik değişkenlerdeki karmaşık etkileşim veya kombinasyonları bulan bir yöntemdir. Yöntem, anakütleyi bağımlı değişkeni en önemli açıklayıcı değişkene göre alt gruplara veya bölümlere tekrarlı olarak ayırmaktadır. İlk temelleri AID yöntemi ile atılan karar ağacı modelleri arasında en çok kullanılan ve bilgisayar ortamında hesaplamaları olan teknikler CHAID, C&RT ve QUEST'dir. AID yönteminde bağımlı değişkeni en iyi açıklayan açıklayıcı değişkene göre ikili alt gruplar oluşturulmaktadır. CHAID yöntemi AID yönteminin geliştirilmiş hali olup bağımlı değişkeni çoklu alt gruplara ayrılabilir. Alt gruplara ayırma işlemini yaparken açıklayıcı değişkenlerin kategorileri ki-kare istatistikleri kullanılarak birleştirilir ve yeni kategoriler oluşturulur. CHAID yöntemi müşteri ilişkileri yönetimi (CRM) çalışmalarında, müşterilerin belli özelliklerine göre sınıflamalarda, tıp alanında hastaların sınıflandırılmasında sıkça kullanılmaktadır.

**Anahtar Kelimeler:** CHAID analizi, AID analizi, kategorik değişken, çoklu alt grup, karar ağaçları.

## **ABSTRACT**

One of the main problems in scientific studies is to find the factors that mostly effect the phenomeon or which level these factors have highest effect. To solve this problem decision tree methods in data mining is prefered that they can be applied in large size data-sets, they are easily understandable in visual way and their assumptions are less than other statistical methods. CHAID analysis was developed as a method which determines the complex interactions and combinations among the categorical variables. Method is repeatedly partitioning the population into different subgroups or segments according to predictor which is most significant variable for dependent variable. Between decision tree models, which AID analysis is published firstly, methods that are most used and have software computation is CHAID, C&RT and QUEST. AID method is forming binary splits according to most explanatory precidtor of dependent variable. CHAID is published as development form of AID and it can form multi-way splits. When CHAID is partitioning subgroups, it uses new categories which is applied splittig the predictors categories by chi-square statistic. CHAID method is mostly used in costumer relationship management (CRM) in to classify the costumer by some qualifciations and in medicine to classify the patients.

**Keywords:** CHAID analysis, AID analysis, categorical variable, multi-way split, decision tree.

## 1. GİRİŞ

Veri madenciliği çeşitli şekillerde elde edilmiş veriyi analiz ederek anlaşılır bir yapıya dönüştürmeyi hedeflemektedir. Kullanılan birçok metot vardır. Karar ağaçları bu metotlardan sadece bir tanesidir ancak varsayımlarının kısıtlı olması, görsel sunumun ön planda olması kolay anlaşılır olması nedenleriyle tercih edilen bir metottur. Karar ağaçları içinde en çok kullanılan ve bilgisayar ortamında geliştirilmiş olan analizler CHAID yani Otomatik Ki-Kare Etkileşim Belirleme Analizi, C&RT yani Sınıflama ve Regresyon Ağaçları ve QUEST yani Hızlı, Yansız, Etkili İstatistiksel Ağaçlardır.

CHAID analizi AID yani Otomatik Etkileşim Belirleme Analizinin geliştirilmiş halidir. Her iki analizde geniş veri setlerinin anlaşılır bir şekilde yorumlanabilmesi için kategorik yapıdaki bir bağımlı değişkenin açıklayıcı değişkenlerle alt gruplara ayırmaya çalışmaktadır. Ancak AID analizinin çeşitli kısıtlamaları olması nedeniyle alt gruplara ayırmada ki-kare önem istatistiklerini kullanan CHAID analizi kullanılmaktadır.

Bu çalışmada öncelikle veri madenciliği hakkında kısa bilgi verilerek daha sonra karar ağaçlarının genel özelliklerinden, kullanım alanlarından, güçlü yönlerinden bahsedilmiştir. Karar ağaçlarında CHAID dışında en çok kullanılan C&RT ve QUEST analizlerinin tarihsel gelişimi, yapısı ve algoritmalarına genel olarak değinilmiştir. Daha sonra ise asıl konu olan AID ve CHAID analizinin genel yapısı, algoritması üzerinde durulmuştur. Beşinci bölümde de bu iki analizinin diğer bölümlerde verilen özellikleri ve algoritmaları dikkate alınarak güçlü ve zayıf yönlerinin karşılaştırılması yapılmıştır.

Çalışma bir şirketin personelinin sağlanan eğitim, işe alma, performans değerlendirme, ücretlendirme, yan faydalar, kariyer planlanma, özlük işleri ve bordrolama, kurum içi iletişim ve insan kaynakları departmanı hakkındaki görüşlerini içeren anket formu uygulanarak elde edilmiş veri setinde bir uygulama yapılarak sonlandırılmıştır. CHAID analizi SPSS' de Answer Tree 3.0 modülü kullanılarak çözümlenmiştir. Veri seti genişliği 3231 personeldir. Öncelikle faaliyet alanı değişkeni eğitim, işe alma, performans değerlendirme, ücretlendirme, yan faydalar, kariyer planlanma, özlük işleri ve bordrolama, kurum içi iletişim, bölge adı ve unvanlar değişkenleri açıklayıcı değişkenler alınarak alt gruplara ayrılmış daha sonra insan

kaynakları departmanı hakkındaki görüşlerini, bölge adı ve unvanlar açıklayıcı değişkenler alınarak alt gruplara ayrılmıştır. Son olarak faaliyet alanı tüm değişkenlerin açıklayıcı değişken olması durumdaki alt grupları bulunmuştur.

Sonuçlar ve değerlendirmeler başlığı altında da yapılan analizlerin karşılaştırılması ve açıklanmasına yer verilmiştir.

## 2. VERİ MADENCİLİĞİ'NDE KARAR AĞAÇLARI

Günümüzde birçok sektör bilgi depolama, kayıtlar tutma konusuna çok değer vermektedir. Özellikle birebir müşterilerinin ihtiyaçları doğrultusunda ürün ve hizmet üreten şirketler için bilgilerin toplanması, analiz edilmesi pazar içinde rekabet açısından önemli olmaktadır. Şirketler tarafından bilgisayar sistemleri sayesinde tutulan her türlü bilgi işlenmediği sürece bir çöp yığını olmaktan başka işe yaramıyor. İşte “ veri madenciliği ” bu bilginin sistemli tutulması, analiz edilip doğru şekillerde kullanılması noktasında önem kazanıyor.

Veri madenciliği, verilerden daha önceden bilinmeyen, gizli, muhtemelen faydalı bilginin monoton olmayan bir süreçte çıkartılması işlemi olarak tanımlamaktadır. [ [www.isletme.istanbul.edu.tr/surekli\\_yayinlar/dergiler/nisan2000/1.htm](http://www.isletme.istanbul.edu.tr/surekli_yayinlar/dergiler/nisan2000/1.htm) ]

Veri madenciliği ile büyük veri yığınlarından oluşan veri sistemleri içerisinde gizli kalmış bilgilerin çekilmesi sağlanır. Bu işlem, istatistik, matematik disiplinleri, modelleme teknikleri, database teknolojisi ve çeşitli bilgisayar programları kullanılarak yapılır. Gartner Grup tarafından yapılan tanımda ise veri madenciliği, istatistik ve matematik teknikleriyle birlikte ilişki tanıma teknolojilerini kullanarak, depolama ortamlarında saklanmış bulunan veri yığınlarının elenmesi ile anlamlı yeni ilişki ve eğilimlerin keşfedilmesi sürecidir. Veri madenciliği kendi başına bir çözüm değil çözüme ulaşmak için verilecek karar sürecini destekleyen, problemi çözmek için gerekli bilgileri sağlamaya yarayan bir araçtır. Veri madenciliği; analiste, iş yapma aşamasında oluşan veriler arasındaki şablonları ve ilişkileri bulması konusunda yardım etmektedir. [ [www.bilgiyonetimi.org/cm/](http://www.bilgiyonetimi.org/cm/) ]

Sıklıkla kullanıldıkları alanlar şöyle özetlenebilir.

Pazarlamada;

- Müşterilerin satın alma örüntülerinin belirlenmesi,
- Müşterilerin demografik özellikleri arasındaki bağlantıların bulunması,
- Posta kampanyalarında cevap verme oranının artırılması,
- Mevcut müşterilerin elde tutulması, yeni müşterilerin kazanılması,
- Pazar sepeti analizi (Market Basket Analysis)
- Müşteri ilişkileri yönetimi (Customer Relationship Management)
- Müşteri değerlendirme (Customer Value Analysis)

- Satış tahmini (Sales Forecasting).

#### Bankacılıkta;

- Farklı finansal göstergeler arasında gizli korelasyonların bulunması,
- Kredi kartı dolandırıcılıklarının tespiti,
- Kredi kartı harcamalarına göre müşteri gruplarının belirlenmesi,
- Kredi taleplerinin değerlendirilmesi,
- Risk yönetimi,
- Risk analizleri.

#### Sigortacılıkta;

- Yeni poliçe talep edecek müşterilerin tahmin edilmesi,
- Sigorta dolandırıcılıklarının tespiti,
- Riskli müşteri tipinin belirlenmesi.

#### Perakendecilikte;

- Satış noktası veri analizleri,
- Alış-veriş sepeti analizleri,
- Tedarik ve mağaza yerleşim optimizasyonu.

#### Borsada;

- Hisse senedi fiyat tahmini,
- Genel piyasa analizleri,
- Alım-satım stratejilerinin optimizasyonu.

#### Telekomünikasyonda;

- Kalite ve iyileştirme analizleri,
- Hisse tespitleri,
- Hatların yoğunluk tahminleri,

#### Sağlık ve İlaçta;

- Test sonuçlarının tahmini,
- Ürün geliştirme,
- Tıbbi teşhis,
- Tedavi sürecinin belirlenmesi.

#### Endüstride;

- Kalite kontrol analizleri,
- Lojistik,
- Üretim süreçlerinin optimizasyonu.

Bilim ve Mühendislikde;

- Ampirik veriler üzerinde modeller kurarak bilimsel ve teknik problemlerin çözümlenmesi.

Veri madenciliği teknikleri direkt ve direkt olmayan diye ikiye ayrılır. Direkt tekniklerin amacı önceden belirlenmiş hedef değişkenini tahmin etmek, sınıflamak, karakterize etmektir. Direkt olmayan tekniklerin amacı ise bir bütün olarak veri setinin yapısının belirlenmesidir.

Bu teknikler, Sınıflama (Classification) ve Regresyon (Regression), Kümeleme (Clustering), Birliktelik Kuralları (Association Rules) ve Ardışık Zamanlı Örüntüler (Sequential Patterns) başlıkları altında ele alınabilir. Sınıflama ve regresyon modelleri direkt; kümeleme, birliktelik kuralları ve ardışık zamanlı örüntü modelleri direkt olmayan tekniklerdir.

**Sınıflama ve Regresyon Modelleri:** Mevcut verilerden hareket ederek geleceğin tahmin edilmesinde faydalanılan ve veri madenciliği teknikleri içerisinde en yaygın kullanıma sahip olan sınıflama ve regresyon modelleridir. Direkt tekniklerin içinde yer alan sınıflama veriyi kategorik olarak ayırmak anlamına gelir. Veri seti, doğru karar verme yönünde, içerisinde çok fazla bilgi gizlemektedir. Veri analizinin iki formu olan sınıflama ve tahmin verinin anlaşılabilirliğini büyük ölçüde artırır. Sınıflama ve tahmin arasındaki fark ise; sınıflama ile kategorik seviyeler elde edilirken tahmin ile sürekli değerli fonksiyonlar modellenir. Ancak çok terimli lojistik regresyon (multinomial logistic regression) gibi kategorik değerlerin de tahmin edilmesine olanak sağlayan tekniklerle, her iki model giderek birbirine yaklaşmakta ve bunun bir sonucu olarak aynı tekniklerden yararlanılması mümkün olmaktadır. Sınıflama ve regresyon modellerinde kullanılan başlıca teknikler,

- Karar Ağaçları (Decision Trees),
- Yapay Sinir Ağları (Artificial Neural Networks),
- Genetik Algoritmalar (Genetic Algorithms),
- K-En Yakın Komşu (K-Nearest Neighbor),
- Bellek Temelli Nedenleme (Memory Based Reasoning),
- Naïve-Bayes,
- Lojistik Regresyondur (Logistic Regression)
- Faktör ve Ayırma analizleri (factor and discriminant)

**Kümeleme Modelleri:** Kümeleme modellerinde amaç üyelerinin birbirlerine çok benzediği ancak özellikleri birbirlerinden çok farklı olan kümelerin bulunması ve veri tabanındaki kayıtların bu farklı kümelere bölünmesidir. Kümeleme analizinde; veri tabanındaki kayıtların hangi kümelere ayrılacağı veya kümelemenin hangi değişken özelliklerine göre yapılacağı konunun uzmanı olan bir kişi tarafından belirtilebileceği gibi veri tabanındaki kayıtların hangi kümelere ayrılacağını geliştirilen bilgisayar programları da yapabilmektedir.

**Birliktelik Kuralları ve Ardışık Zamanlı Örüntüler :** Bir alışveriş sırasında veya birbirini izleyen alışverişlerde müşterinin hangi mal veya hizmetleri satın almaya eğilimli olduğunun belirlenmesi, müşteriye daha fazla ürünün satılmasını sağlama yollarından biridir. Satın alma eğilimlerinin tanımlanmasını sağlayan birliktelik kuralları ve ardışık zamanlı örüntüler, pazarlama amaçlı olarak pazar sepeti analizi (Market Basket Analysis) adı altında veri madenciliğinde yaygın olarak kullanılmaktadır. Bununla birlikte bu teknikler, tıp, finans ve farklı olayların birbirleri ile ilişkili olduğunun belirlenmesi sonucunda değerli bilgi kazanımının söz konusu olduğu ortamlarda da önem taşımaktadır. Birliktelik kuralları eş zamanlı olarak gerçekleşen ilişkilerin tanımlanmasında kullanılır. Ardışık zamanlı örüntüler birbirleri ile ilişkisi olan ancak birbirini izleyen dönemlerde gerçekleşen ilişkilerin tanımlanmasında kullanılır.

Bilimsel çalışmalardan elde edilen verilerin analizinde sınıflama (diskriminant, lojistik regresyon, kümeleme analizleri gibi) ve regresyon modelleri sıkça kullanılmaktadır. Ancak bu tür modellerin gerektirdiği varsayımlar pek çok alanda istatistiksel analiz imkanlarını kısıtlamaktadır. İncelenen veri seti üzerinde hiçbir varsayım gerektirmemesi nedeniyle, Sınıflama ve Regresyon Modelleri bu tür istatistiksel sınıflama ve regresyon tekniklerine karşı güçlü bir alternatif olarak ortaya çıkmaktadır. \*

Sınıflama ve Regresyon Modelleri' nin bir yöntemi olan Karar Ağaçları sınıflandırma ve tahmin için kullanılabilir. Kök, dallardan ve yapraklardan oluşan ağaca benzeyen bir yapısı vardır, bu yüzden karar ağaçları diye adlandırılmıştır. Veri setinin çok karmaşık olduğu durumlarda bile, bağımlı değişkeni etkileyen değişkenleri ve bu değişkenlerin modeldeki önemini basit bir ağaç yapısı ile görsel olarak sunabilmektedir. Örnekteki tüm gözlemleri kapsayan bir kök ile başlar aşağıya doğru indikçe veriyi alt gruplara ayıran dallara ayrılırlar.

---

\* Orekici Temel,G., Çamdeviren,H., ve Akkuş,Z., ( 2005 ), “ Sınıflama Ağaçları Yardımıyla Legs Syndrome Hastalarına Tanı Koyma ”, İnönü Üni. Tıp Fak. Dergisi, 12 (2):111.



Bu kökten dallara doğru büyüyen ağaç yapısında her boğum “ düğüm”dür, oluşan ağaçlarda homojen olmayan düğümlere “ çocuk düğümü ( child node ) ”, homojen düğümlere ise “ terminal düğüm ( parent node )” adı verilir. Herhangi bir düğümün heterojenlik değeri safsızlık (impurity) ölçüsü olarak adlandırılır ve bu değer safsızlık fonksiyonu kullanılarak hesaplanır. Safsızlık ölçüsü sıfır değerini alıyorsa düğüm tamamen homojendir. Ağaçlarda kullanılabilecek birçok alternatif safsızlık ölçüsü (Gini, Twoing, Chi-square, G-square) vardır. Ayırma fonksiyonundan anlaşılacağı gibi, kullanılan safsızlık ölçüsü herhangi bir t düğümü için en iyi ayırmanın seçimini önemli bir şekilde etkilemektedir. En yaygın olarak kullanılan ayırma kriterleri Gini Değişim İndeksi (Gini Diversity Index) ve Twoing Kuralı'dır.

$$\text{gini}(T) = 1 - \sum_{j=1}^n p_j^2$$

Tümüyle saf bir düğümün (node) Gini indeks değeri 0 olur.

Bir sınıflama modelinde yanlış olarak sınıflanan olay sayısının, tüm olay sayısına bölünmesi ile hata oranı, doğru olarak sınıflanan olay sayısının tüm olay sayısına bölünmesi ile ise doğruluk oranı hesaplanır ( Doğruluk Oranı = 1 - Hata Oranı ). Verilerin sınıflandırılması için oluşturulan modellerin hata oranlarına karar vermek için risk matrisi kullanılmaktadır. \* Bir karar ağacı oluşturulduktan sonra doğruluğunun belirlenmesi önemlidir. Ağaçtaki hatalı sınıflamalar “ risk ” olarak adlandırılır. Ağacın riskini tahmin etmek için üç alternatif metot vardır; yeniden yerine koyma tahmini, test örneği tahmini ve çapraz geçerlilik testidir. Risk tahminleri ve yanlış sınıflama tabloları son ağacın doğruluğunu gösterir.

Sınıflama ağaçları modellerinde tekrarlı ikili bölünmelerle homojen alt gruplar elde edilir ve ağaç bu şekilde büyümeye devam eder. Sınıflama Ağaçları’ nda meydana gelen büyüme;

1. Her çocuk düğümündeki gözlem sayısı sadece bir gözlem ise veya on gözlem ise,
2. Her düğümde grup içi homojenlik söz konusu ise,
3. Ağacın düzey sayısında analizi yürüten kişi tarafından bir sınırlama yapıldıysa,
4. Yeni oluşacak düğümlerde fazla bir değişiklik yaratmıyorsa durur.

Ağaç inşası sonunda elde edilen ağaç maksimum ağaç olarak adlandırılır. Ancak maksimum ağaç pratikte iki dezavantaja sahiptir.

---

\* Orekici Temel,G., Çamdeviren,H., ve Akkuş,Z., ( 2005 ), “ Sınıflama Ağaçları Yardımıyla Legs Syndrome Hastalarına Tanı Koyma ”, İnönü Üni. Tıp Fak. Dergisi, 12 (2):113.

1. Maksimum ağaç başlangıç veri setini kusursuz biçimde tanımlar çünkü eklenen her bağımsız değişken hatalı sınıflama oranını düşürür. Bu durumda, maksimum ağaç veri için olması gerekenden daha iyi bir tahmin modeli (overfitting) sunar. Ancak, başlangıç veri setine aşırı uyumlu maksimum ağaçlar farklı bir veri seti söz konusu olduğunda iyi bir tahmin sağlayamazlar.
2. Bir sınıflama ağacının karmaşıklık ölçüsü o ağacın terminal düğüm sayısına eşittir. Terminal düğüm sayıları ve dolayısıyla karmaşıklığı yüksek olan maksimum ağacın anlaşılması ve yorumlanması güçtür. Bu yüzden büyütme ve budamadan sonra basitlik ve doğruluğu dengede tutan en iyi ağacı seçmek gerekir. Bu ağaca “ optimum ağaç ” denir.

Karar ağaçları yöntemlerin güçlü yönleri şunlardır:

- Sınıflama yaparken çok küçük bir hesaplama gerektirir,
- Hem kategorik hem de sürekli değişkenler ile çalışabilir ayrıca karmaşık veri setlerine kolaylıkla uygulanabilir,
- Sınıflama ve tahmin için hangi değişkenlerin daha önemli olduklarını açık belirtiler ile gösterir,
- Parametrik olmayan bir model olduğu için varsayımları çok kısıtlıdır,
- Bağımlı ve bağımsız değişkenler arasındaki ilişki görsel sunuma sahip olduğundan, ağaç şeklindeki model sonuçları çok fazla istatistik bilgisine gerek duyulmadan kolay bir şekilde yorumlanabilir.
- Tanımlanan bağımlı değişken için olabilecek bütün bağımsız değişkenleri ve onların tüm kombinasyonlarını modele katar ve mümkün olan en doğru sınıflandırmayı yapar. Değişken kombinasyonlarına da bakıldığı için interaksiyonlar da değerlendirilmiş olur.
- Hem bağımlı hem de bağımsız değişkenler için kayıp veya eksik değerler ile aşırı uç değerlerden etkilenmeyen bir metottur.

Tahmin edici ve tanımlayıcı özelliklere sahip olan karar ağaçları, veri madenciliğinde kuruluşlarının ucuz olması, yorumlanmalarının kolay olması, veri tabanı sistemleri ile kolayca entegre edilebilmeleri ve güvenilirliklerinin daha iyi olması nedenleri ile sınıflama modelleri içerisinde en yaygın kullanıma sahiptir.

Karar ağacı temelli analizlerin yaygın olarak kullanıldığı sahalar,

- Belirli bir sınıfın muhtemel üyesi olacak elemanların belirlenmesi (Segmentation),
- Çeşitli vakaların yüksek, orta, düşük risk grupları gibi çeşitli kategorilere ayrılması (Stratification),
- Gelecekteki olayların tahmin edilebilmesi için kurallar oluşturulması,
- Parametrik modellerin kurulmasında kullanılmak üzere çok miktardaki değişken ve veri kümesinden faydalı olacakların seçilmesi,
- Sadece belirli alt gruplara özgü olan ilişkilerin tanımlanması,
- Kategorilerin birleştirilmesi ve sürekli değişkenlerin kesikliye dönüştürülmesidir.

Karar ağacı temelli tipik uygulamalar ise,

- Hangi demografik grupların mektupla yapılan pazarlama uygulamalarında yüksek cevaplama oranına sahip olduğunun belirlenmesi (Direct Mail),
- Bireylerin kredi geçmişlerini kullanarak kredi kararlarının verilmesi (Credit Scoring),
- Geçmişte işletmeye en faydalı olan bireylerin özelliklerini kullanarak işe alma süreçlerinin belirlenmesi,
- Tıbbi gözlem verilerinden yararlanarak en etkin kararların verilmesi
- Hangi değişkenlerin satışları etkilediğinin belirlenmesi,
- Üretim verilerini inceleyerek ürün hatalarına yol açan değişkenlerin belirlenmesidir.

İlk temelleri AID yöntemi ile atılan karar ağacı modelleri çeşitli algoritmalar ile sürdürülmüştür. Geliştirilen bu algoritmalar içerisinde **CHAID** (Chi-Squared Automatic Interaction Detector; G.V. Kass; 1980), **C&RT** (Classification and Regression Trees; Breiman, Friedman, Olshen ve Stone; 1984), **ID3** (Quinlan; 1986), **Exhaustive CHAID** (Biggs, de Ville ve Suen; 1991), **C4.5** (Quinlan; 1993), **MARS** (Multivariate Adaptive Regression Splines; Friedman), **QUEST** (Quick, Unbiased, Efficient Statistical Tree; Loh ve Shih, 1997), **C5.0** (Quinlan), **SLIQ** (Supervised Learning in Quest; Mehta, Agarwal ve Rissanen), **SPRINT** (Scalable Parallelizable Induction of Decision Trees; Shafer, Agrawal ve Mehta) başlıcalarıdır. **CHAID**, **CART (C&ART)**, **QUEST** en sık kullanılan algoritmalarlardır.

[ [www.isletme.istanbul.edu.tr/surekli\\_yayinlar/dergiler/nisan2000/1.htm](http://www.isletme.istanbul.edu.tr/surekli_yayinlar/dergiler/nisan2000/1.htm) ]

## 2.1 En Sık Kullanılan Karar Ağaçları Algoritmaları

### 2.1.1 C&RT (Classification and Regression Tree )

Ele alınan bağımlı değişkenin (hedef değişken, target variable ) kategorik yapıda ise yöntem sınıflama ağaçları (Classification Tree, CT), sürekli ise regresyon ağaçları (Regression Tree, RT) olarak adlandırılmaktadır. Bu araçlar çoğunlukla tıp alanında ve pazarlama amaçlarını belirlemede kullanılmaktadır. 1984 yılında Breiman, Friedman, Olshen ve Stone ( BFOS ) tarafından geliştirilen algoritma ikili karar ağaçları oluşturur. C&RT, iki çocuk düğümü oluşturup bütün açıklayıcı değişkenleri kullanarak veriyi alt gruplara ayırmak üzerine kurulmuştur. En iyi açıklayıcı değişken safsızlık ve değişim ölçülerindeki değişkenliği kullanarak seçilir. Burada amaç hedef değişkene ilişkin mümkün olabilen en homojen veri alt gruplarını oluşturmaktır. Bağımsız değişken tektir ve sınıflayıcı, sıralayıcı ve sürekli türde; açıklayıcı değişken bir ya da daha fazla olabilir ve sınıflayıcı, sıralayıcı ve sürekli türde ölçülmüş olabilir. Kayıp veri olması durumunda bu kayıp verilerin yerine koyma kuralları uygulanır, bütün kayıp veriler aynı şekilde davranmazlar ve kayıp veriler yerine konan (vekil) değişkene bağlı olarak bir gruba dahil edilir. Örneğin, mevki değişkenine göre veri ayrılmak istensin, mevki değişkeni bazı gözlemler için kayıp veri olsun. Bölge ofisi kayıp veriler yerine en iyi mevki grubunu atayacaktır. [ <http://dollar.biz.uiowa.edu/~street/Trees.ppt> ] C&RT algoritması limitleri ağacın ikili ayırmalar ile sınırlandırılması ve değişken seçerken maliyet matrisi dikkate alınması şeklinde özetlenebilir. C&RT metodunda ağaç oluşturulurken ön olasılıklar ( prior probabilities) kullanılır. Ön olasılıklar deney ünitelerinin dahil olacağı sınıfın belirlenmesini etkiler.\*

C&RT algoritmasının adımları şu şekildedir:

1. İlk düğümden başlayarak bölünmeyi sağlayacak tüm mümkün adayların içinden bir tane ayıraç seçilir,
2. Hedef değişkenin tipine göre safsızlık ölçütü hesaplanır,
3. Açıklayıcı değişkenler safsızlık ölçütlerine göre karşılaştırılır,
4. Safsızlık ölçütünü maksimum olan değişkene göre ayrıştırma yapılır,

---

\* Orekici Temel,G., Çamdeviren,H., ve Akkuş,Z., ( 2005 ), “ Sınıflama Ağaçları Yardımıyla Legs Syndrome Hastalarına Tanı Koyma ”, İnönü Üni. Tıp Fak. Dergisi, 12 (2):113.

5. Ayıraç belirleme sürecine diğer açıklayıcı değişkenler içinde devam edilir,
6. Herhangi bir durdurma kuralına rastlayana kadar ağaç büyütülür.

Bu algoritma ile elde edilen maksimum ağaçtan optimum ağacı oluşturmak için budama yapılmalıdır. C&RT ile oluşan ağaçların budanmasında ayrı bir değerlendirme veri setine gerek vardır. Optimum ağaç hem hatalı ayırma riskini hem de ağacın karmaşıklığını ölçen bir indeks ( maliyet-karmaşıklık ölçüsü ) kullanarak bu indeks değerini minimum yapan ağaçtır.

### 2.1.2 QUEST ( Quick, Unbiased, Efficient, Statistical Tree )

Hızlı, yansız, etkili istatistiksel ağaç (Quick, Unbiased, Efficient, Statistical Tree) olarak bilinir. 1997 yılında Loh ve Shih tarafından geliştirilmiştir. C&RT algoritmasında olduğu gibi ikili karar ağaçları oluşturmak üzerine kurulu bir yapısı vardır. Fakat CHAID ve C&RT' dan farklı olarak değişken seçimi ve ayırma noktası seçimi işlemlerini ayrı ayrı ele almaktadır. En çok bilinen ayrıntılı araştırma metotları C&RT gibi, ağaç büyüme sürecinde daha fazla bölünmeyi sağlayan fazla kesikli değer ile değişken seçme yoluna gidilmektedir. Bu durum sonuçların genelleştirilmesini azaltarak modelin yanlış olmasına neden olmaktadır. Eğer bağımsız değişken için açıklayıcı değişkenler aynı bilgi vericilikte iseler QUEST bu açıklayıcı değişkenlerden herhangi birini eşit olasılıkla seçecektir. Ayrıca hesaplamalardaki etkinliği açısından da kuvvetli yapıdadır. C&RT' nın sağladığı bütün avantajlara sahip olmasına rağmen yine de C&RT gibi ağaç hantal yapıdadır. Bağımlı değişken tek ve nominal türde; bir ya da daha fazla açıklayıcı değişken sınıflayıcı, sıralayıcı ve sürekli türde ölçülmüş olabilir. Kayıp verilerde yerine koyma kuralları uygulanır. QUEST, bağımlı değişkenler kategorik yapıda ise, yansız ağaç önemli ise, büyük ve karmaşık yapıda veri seti varsa ve ağaç tahmininde etkili bir algoritmaya gerek varsa, ağaç ikili bölünme ile sınırlandırılmışsa kullanılır. [ <http://engr.smu.edu/~mhd/8331f03/AT.pdf>.] C&RT' de olduğu gibi ön olasılıklar kullanılmaktadır. QUEST algoritması şu şekildedir; her ayırma için her açıklayıcı değişken ve bağımlı değişken arasındaki ortaklık Anova F testi ya da Levene testi ( sıralayıcı ve sürekli açıklayıcı değişkenler için ) ya da Pearson Ki-Kare değeri ( sınıflayıcı açıklayıcı değişkenler için ) hesaplanarak bulunur. Değişken seçimi için küçük p değerine sahip olan değişken seçilir. Ayırma işleminde ise eğer bağımlı değişken ikiden fazla kategoriye sahip ise iki süper sınıf bulabilmek için iki ortalama kümeleme algoritması kullanılır. Aksi takdirde açıklayıcı değişkenin en iyi bölünmesini bulmak için Kuadratik Diskriminant Analizi ( QDA) kullanılır. Bu süreç herhangi bir durdurma kuralına rastlayıncaya kadar devam eder. C&RT olduğu gibi

maksimum-karmaşıklık ölçüsü ile ağaç budanır. CHAID' de olduğu gibi maliyet matrisi direkt olarak ağaç büyütme sürecine katılmaz. Ancak, simetrik maliyet matrisi için, maliyet bilgileri modele maliyet matrisindeki ön olasılıkları kullanarak dahil edebilir.

### 3. AID ANALİZİ

#### 3.1 AID Analizinin Genel Yapısı

Ağaç temelli sınıflandırma işlemlerinin ortaya çıkışı 60'lı yılların başlarına kadar gider. AID tekniği yani otomatik etkileşim belirleyicisi tekniği ilk kez Michigan Üniversitesi İnceleme ve Araştırma Merkezi'nde 1964 yılında Morgan ve Sonquist tarafından geliştirilmiştir. Bu çalışmalar sonra da devam etmiştir. Morgan ve Sonquist AID ile, veri kümesini nicel bir bağımlı değişkene göre, sınıflayıcı (nominal) ya da sıralayıcı (ordinal) açıklayıcı değişkenler kullanarak ikiye bölünmesi yöntemini ortaya çıkarmışlardır. AID gruplar arası kareler toplamalarını, her bölümde maksimize ederek ikili ağaçlar oluşturur.\* Holmes' a göre AID tekniğinin geliştirilmesini nedeni standart çapraz sınıflama (cross-tabulation) analizlerinin veri içindeki kompleks etkileşimleri açıklama kabiliyetinin tatmin edici olmamasıdır. AID, analistlerin ortalamalar yardımı ile verideki etkileşimleri sistematik bir şekilde incelemelerini sağlar. Pazar araştırmalarında bu tür incelemeler hedefe en uygun pazar segmentinin belirlenmesinde rol oynar. \*\*

AID tekniğine alternatif analiz metotları ile arasındaki farklılıklar ve tekniğin sağladığı üstünlükler şu şekilde açıklanabilir.

Veri seti ile ulaşılacak istenilen asıl amaç aşağıdaki formül bileşenlerinin arasındaki ilişkiyi saptamaya yöneliktir.

$$y = f(x_1, \dots, x_n) + e \quad (3.1)$$

Formülde y bağımlı değişken,  $x_1, \dots, x_n$  açıklayıcı değişkenleri ya da y' yi etkileyen bağımsız değişkenleri ve e hata terimini ifade etmektedir. İyi bir model önemli korelasyona sahip değişkenleri doğru yapıda formüle etmek şeklinde tanımlanabilir. Ön olasılıklar veya teoriler değişkenlerin seçiminde çeşitli varsayımlar sunar, fakat genel olarak bağımsız

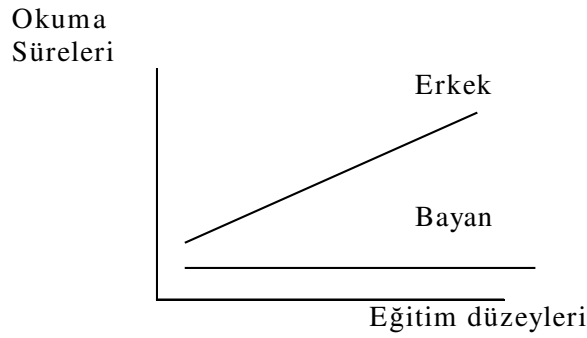
---

\* Kass,G.V., (1980), “ An Exploraty Tecnique For Investigating Large Quantaties Of Categorical Data”, Applied Statistic,29(2):119.

\*\* Baron,S., ve Phillips,D.,(1994), “Attitude Survey Data Reduction Using CHAID: An Example in Shopping Center Market Research”, Journal Of Marketing Manegement, X :75.

değişkenlerdeki lineer olmayan etkilerin ve etkileşimlerin saptanması daha zor olur. Bağımlı ve bağımsız değişkenler arasındaki bağıntı doğrusal veya eğrisel olabilir. Yanlış belirlenmiş bir model, daha büyük açıklanmayan varyansa ve bileşenlerin etkilerine dair yanlış tahminlere neden olabilir.

Lineer olmayan sonuçlar genellikle doğrusal grafik prosedürler ya da dönüşümlerle tanımlanmaktadır. Etkileşimler daha karmaşık problemlerin tanıtımında kullanılır. Etkileşim bir bileşenin etkisinin ( $x_i$ ) bir diğerinin değerine bağlı olduğu durumlarda ortaya çıkar. Örneğin eğitim, erkekler arasında okuma için harcanan zamanı etkileyebilir fakat bu kadınlar için geçerli değildir. Bu durum grafiksel olarak iki eğriyle Şekil 3.1’de gösterilmektedir. Eğer bu etkileşim göz ardı edilirse, eğitimin etkisi hesaba katılmamış yani daha düşük bir tahmin yapılmış olacaktır. Morgan ve Sonquist’e göre sosyal araştırmalarda bu tarz etkileşimler çok rastlanan, yaygın bir durumdur. \*



Şekil 3.1 Okuma süreleri üzerinde; Cinsiyet ve eğitim arasındaki etkileşim

Çapraz sınıflama (cross-tabulation) tablolarının avantajı, bu tablolarda etkileşimlerin izole edilebilmesidir. Bu tablolar, yalınlık, kolaylık ve etkileşimlerin izole edebilme imkanı gibi belirleyici özelliklerinden dolayı veri analiz sonuçlarının sunulmasında en çok kullanılan metot haline gelmişlerdir. Çapraz sınıflama tabloları ile ilgili ana problem, bu metot da analizleri iki bileşenden öteye genişletmenin zor olmasıdır, böylece analiz çok kullanışsız hale gelmektedir ve analiz için kullanılan örnek ya da model kısa bir sürede (bölünme süresi) çalışmayı zorlaştıracak kadar küçük alt gruplara bölünmektedir. Önemli değişkenlerin dışarıda bırakıldığı model, varyansın önemli kısımlarını açıklamakta başarısız olacaktır. Bununla birlikte, eğer değişkenler arasında bir ilişki söz konusu ise, bunlardan birini ya da bir

\* Doyle,P., ve Fenwick,I.,(1975) “The Pitfalls of AID Analysis”,Journal Of Marketing Research,November 1975,XII:408.



kaçını model dışında bırakmak, dahil edilen değişkenlerin etkilerinin yanlış bir tahmininin yapılmasına neden olacaktır.

Regresyon analizi, örnekten hareket edildiği durumlarda, bilinen bir veya daha fazla değişkenden yararlanarak başka bir değişkenin tahmin edilmesi ve bu tahmin sürecinde yapılan hataların ölçülmesidir (Genceli,2001). Regresyon analizi, çoklu bileşenleri ve bunların önem seviyelerini kullanma ve kısmi katkılarını tahmin etme olanağı sağlar. Bununla birlikte, modelin biçimlenmesi söz konusu olduğunda, bazı sınırlamaları vardır. Bunun anlamı, bağımsız değişkenlerin etkilerinin lineer ve katkılı olduğu varsayımdır. Bu varsayımın gerçekleşmemesi durumunda ise birinci mertebeden modeller yetersiz kalır ve başka model arayışlarına gidilmelidir. Ayrıca araştırmacı hem bütün etkileşimli terimleri belirlemeli ve modele dahil etmelidir hem de bağımsız değişkenler arasında etkileşim olmamasını varsaymalıdır. Ancak deneysel olaylarda çok az çalışma sahası olduğundan etkileşimlerin önemsiz olduğu varsayımı ile model oluşturulur. Non-lineerlikler ve etkileşimler tahmin edilebilir olduğunda, bunların öncelikle açıkça belirtilmesi gerekir. Önemli bağımsız değişken ya da değişkenlerin model kapsamına alınmaması, ilgisiz bağımsız değişkenlere modelde yer verilmesi, model daha büyük olduğu halde ilişkilerin tek denklemle ifade edilmesi gibi matematiksel biçimlendirme hatalarından dolayı regresyon analizinden istenilen faydanın elde edilemeyebilir. Regresyon analizi ile ilgili olarak, Bass, Tigert ve Lonsdale tarafından kanıtlanan ikinci bir problem de bu analizin genellikle davranışların tahmini için uygun bir teknik olmadığıdır. Genelde analiz önemli davranış farklılıklarını kullanarak grupların belirlenmesi ile ilgilenir. Regresyon analizi, bunun yanında, gözlemin bir parçası olarak gruptan ziyade bireyi kullanmaktadır ve modeli kurma kriteri olarak da iki grup arası değişkenlik yerine grup içi değişkenliği kullanmaktadır. Neticede, regresyon katsayıları gruplanmış veriler arasındaki önemli farkları yansıtmayabilir. Assael ve diğerleri bu tarz problemlerin üstesinden gelmek için AID’i önermektedirler. AID, grup davranışlarının açıklanmasında çoklu bileşenlerin kullanılmasına olanak sağlar ve aynı zamanda modelin yapısıyla ilgili sınırlayıcı varsayımlardan kurtulur. \*

Araştırmacılar tanımlayıcı istatistiklerden fazlasına ulaşmak istediklerinde ve spesifik biçimsel hipotez testlerini kullanmak istemediklerinde AID analizinden yararlanabilir. AID

---

\* Doyle,P., ve Fenwick,I.,(1975) “The Pitfalls of AID Analysis”,Journal Of Marketing Research,November 1975,XII:409.

analizinin kullanılabilmesi için tek varsayım, bağımsız değişkenlerin sınıflayıcı ölçekle ve bağımlı değişkenlerin ise aralıklı ölçekle ölçülmüş olmasıdır. Bağımlı değişkenlerin dağılımı ya da bağımsız değişkenlerin fonksiyonel formu ile ilgili herhangi bir varsayım AID analizinde yoktur. \* Bütün bağımsız değişkenler kategorik yani koşulsuz olarak ele alınmakta ve kullanıcı, kategorilerin sayısını ve bunların her değişken için sıralarını açıkça belirleyebilmektedir.

AID analizinin basit stratejisi anakütleyi bağımlı değişken ile ilgili en mümkün istatistiksel homojen alt gruplara bölmektir. Belirli segment ya da gruplar için homojenlik gruptaki her bireysel grup ve grup ortalaması arasındaki farkın kareleri toplamının hesaplanmasıyla ölçülür. Bu da grup içi varyasyonların ölçüsü olan gruplar içi kareler toplamıdır(W). Küçük gruplar içi kareler toplamı büyük homojenliği göstermektedir.

### 3.2 AID Analizinin Algoritması

AID tekniğinin amacı, (3.1) formülü ile gösterilen problemi çözmek değil, basitçe söylemek gerekirse formülün yapısını tanımlamaktır. Regresyon analizinden farklı olarak, AID tekniği, bileşenlerin nisbi önemi ya da istatistiksel anlamı hakkında güvenilir bir bilgi vermemektedir. AID tekniğinin başlıca fonksiyonu, değişkenlerin arasındaki ilişkilerin doğasını, özelliklerini yani katkılı ya da etkileşimli olup olmadıklarını araştırmaktır. Bu özelliğiyle, AID açıkça, regresyon analizi ya da benzer teknikler için başlangıç niteliğindedir. Temel prensip bağımlı değişken ve açıklayıcı değişkenler arasındaki mümkün tüm ilişkileri araştırmış önceki çalışmalarla bağımlı değişkenin varyansını açıklamaya çalışmaktır. Araştırmanın sonuçları ikili ağaçta gösterilir ve dallandırmanın yanı sıra her bölümün genişliği ve modelin açıklayıcı gücü ağaçta yer alır. Algoritmayı özetleyecek olursak, anakütle ile başlar önce anakütleyi iki gruba böler. Bu gruplardan her birini toplam kareler kriterine göre hem grup bölünemeyecek kadar küçük hem de grup homojen oluncaya kadar iki gruba daha böler.

Adım adım algoritma şu şekildedir:

1. Tüm değişkenler dikkate alındıktan sonra, i. değişken yani en büyük kareler toplamına sahip olan seçilir. (total sum of squares)

---

\* Staelin,R.,(1971), “Another Look at AID”, Journal of Advertising Research, October 1971, XI,(5):.25.

$$TSS_i = \sum_{a=1}^{N_i} Y_a^2 - \left( \frac{\sum_{a=1}^{N_i} Y_a}{N_i} \right)^2$$

$TSS_i$  = i. değişken ya da grup için toplam kareler toplamı

$Y$  = Bağımlı değişkene ait gözlemler

$Y_a$  = i. değişken ya da gruba birliktaki gözlem çifti için bağımlı değişken değeri

$N_i$  = i. değişken ya da gruba ait gözlem sayısıdır.

2. Bir değişken seçildikten sonra,; gözlemler aşağıdaki eşitliği maksimum yapacak iki alt gruba bölünür.

$$BSS_{ikp} = (n_1 \bar{Y}_1^2 + n_2 \bar{Y}_2^2) - n_i \bar{Y}_i^2$$

$n_1$  ve  $n_2$  bölünme sonrası her gruptaki gözlem sayısı

$\bar{Y}_1, \bar{Y}_2$  her bölünmüş grup için ortalama

$N_i$  = bölünme öncesi i. değişken ya da grup için gözlem sayısı

$\bar{Y}_i$  = bölünme öncesi i. değişken ya da grup için ortalama

$BSS_{ikp} (B) = k$  ve  $p$  gruplarına bölünen i. değişken ya da grup için gruplar arası kareler toplamıdır.

Bu adımlar sonucu oluşan ağaç şu bilgileri içermektedir:

- Her grup için bağımlı değişkenin ortalaması (her kutunun sol alt köşesinde bulunur ).
- Her grup içindeki gözlem sayısı (her kutunun sağ alt köşesinde bulunur ).
- Her değişken bölünmesine ait birleştirici güce dayalı sıralayıcı dereceler oluşturan bölünmelerin mertebesi ( her kutunun üstünde bulunur ).
- Her bir bölünme için t değeri.
- Analizin toplam açıklayıcı gücünü gösteren  $R^2$  (determinasyon katsayısı).

AID analizinde ağacın dallanmalarının önemi büyüktür. Eğer simetrik dallanma var ise (aynı değişkenler ağacın yüksek ya da düşük dallarında beraber yer alıyorsa ) bağımsız değişkenler arasında güçlü bir etkileşim yoktur. Eğer ağaçtaki dallanma simetrik değilse, açıklayıcı değişkenler arasında etkileşim vardır ve daha kompleks bir modele ihtiyaç duyulabilir. AID analizinden elde edilmiş biçimlendirilmemiş model, çıktı verinin kalitesi üzerinde oldukça etkilidir. Tüm parametrelerin birlikte dağılımları, ayrı ayrı etkilere sahiptir ve doğru kullanılarak analizden büyük bilgiler elde edilebilir.

AID ağacının her bir dalı 3 durdurma kuralından bir tanesi ile son verilene kadar çatallanır. Çatallanma;

- 1) Bölünemeyecek kadar küçük sınıflar oluştuğunda,
- 2) Grubun homojen hale gelmesiyle daha fazla bölünmenin gereksiz olduğunda,
- 3) Artık gruplar arası kareler toplamını küçültecek hiçbir olası bölünme kalmadığında durur.

### 3.3 AID Analizinin Limitleri ve Dezavantajları

AID genellikle yanlış kullanılmakta ve yanlış yorumlanmaktadır. Bu yüzden veri analizinin başlangıç kısmı olarak doğru kullanıldığında önemli olan, yanlış tahminlerin ne kadar kabul edilebileceğinin belirlenmesi ve doğrulama testleri olmadan kullanımının tatmin edici olmayacağını göz önünde bulundurulmasıdır.

AID tekniğinin popülerliğinin nedeni, sezgisel ve görsel cazibesidir ancak uygulamalar AID tekniğinin bazı katı sınırlamaları olduğunu fark etmemektedirler. Bu limitleri incelersek;

**Örnek Boyutu ( Sample Size ):** Bireyden ziyade grubu analizin bir parçası olarak kullanılması nedeniyle büyük örnek boyutları tercih edilmektedir. Songquist, Baker ve Morgan, eğer AID analizinden anlamlı sonuçlar elde edilmek isteniyorsa en az 1000 vakanın incelenmesini önermektedirler. Hatta bu rakam bile, araştırma tekniklerinin her zaman için onaylanması gerekliliğinden dolayı, yetersiz kalmaktadır. Bu durumda eğer bölünme-örnekleme (split-sample) analizi kullanılacaksa, yararlı sonuçlar elde edilmesi için 2000 vakanın incelenmesi tercih edilmektedir.

**Korelasyonlu Açıklayıcı Değişkenler (Intercorrelated Predictors) :** Çok değişkenli tekniklerden farklı olarak AID birbiriyle bağlantılı bileşenleri dikkate almamaktadır. Aşağıda açıklandığı gibi, birbiriyle karşılıklı ilişkisi olan bileşenlerin olduğu durumlarda, bunlardan birinin özel olarak seçilmesi muhtemeldir. Yalnızca değişkenlerden birinin ilk olarak seçilecek olması değil, seçilmiş olsa dahi, bunun korelasyonlarının seçilmesi ihtimali daha azdır. Dolayısıyla, değişkenlerin görünüm sıralanışı, nispi önem için bir belirti değildir ve dışında tutma mutlaka değersizlik anlamına gelmez. Bu son ikaz, AID ağacı ile ilgili olan basit ve yaygın yorumu son derece cesur hale getirmektedir.

**Çarpıklaştırılmış Değişkenler ( Skewed Variables ) :** Eğer bağımlı değişken yüksek bir çarpıklığa sahipse program defalarca küçük gruplara bölünmeye eğilimlidir. Sonquist açıklayıcı değişkenlerin çarpıklığının onların güçlerini azaltarak yüksek BSS'nin ortaya çıkmasına neden olduğunu bu durumdaki AID ağacında istenmeyen bir görünüm

oluşturduklarını söylüyor. Aslında, BSS’ deki bu etki çarpıklığın yönüne bağlıdır. Eğer açıklayıcı değişken bağımlı değişkenin yüksek değerlerine doğru çarpıklaştırılmışsa, BSS çarpıklaştırılmış değişkendeki bölünmelerle çok arttırılabilir. Bu nedenle, çokça çarpıklaştırılmış değişkenler, özellikle de analizin ilerleyen safhalarında çıkarılma eğilimindedir. Her iki durumda da bir logaritma ya da karekök alma bazen yardımcı olabilmektedir.

**Gürültü ( Noise ) :** Veride korelasyonlu açıklayıcı değişkenler ve gürültünün aynı anda varlığında, aynı ana küleden alınan farklı örneklerle aynı ağaç elde edilemez. Sonquist, değişkenlerin görünümünün düzeni değişse bile son grupların düzenlemesini kararlı bulmaktadır. Bununla birlikte Doyle, birbiriyle ilişkili bileşenlerle yaptığı çalışmalarda, kararlı olmayan ağaçlarla ilgili problemler bulmuştur. Ağaç durağanlığı, her bölünme için bir sonraki en iyi BSS’ yi veren değişkeni kaydederek test edilebilir, program değişkeni tekrar ortaya çıkarmak için zorlayabilir. Eğer sonuçta elde edilen ağaçlar karasız, dengesizlerse, bu durumda analizi yapan kimse ya birbiriyle ilişkili bağların verilerini faktör analizini kullanarak temizler ya da ağaç takımları tarafından gösterilen kapsamlı yapı son regresyon denklemine birleştirilebilir.

**Önem Testleri ( Significance Tests ) :** AID tekniğinin önceki versiyonları bölünmeler için önem testlerini veriyor olsa da, bunlar algoritmanın izlediği temel stratejileri görmezden gelmektedirler. Sonquist, Baker ve Morgan kabul etmektedirler ki “incelenen mümkün bölünmelerin büyük sayıda olmasından dolayı, istatistiksel önemi sorgulayan hiçbir nokta yoktur”.

**Etkileşimler ( Interactions ) :** Yapıyı tanımlama amacının net olmasına rağmen, AID, etkileşimlerin çeşitli biçimlerine karşı duyarsız bulunmuştur. AID’ nin sadece açıklayıcı değişkenin BSS üzerindeki en yakın etkisini incelemesi ve gelecekteki olası bir bölünmelerin incelememesi nedeniyle, bir aşamalı olmayan herhangi bir etkileşim tanımlanamayacaktır. Böylece,  $X_1$ ’in karakteristiğinin varlığını ve  $\bar{X}_1$ ’in karakteristiğinin yokluğunu ifade ederse formülde eğer  $E ( Y / X_1 \& X_2 ) \neq E ( Y / \bar{X}_1 \& \bar{X}_2 )$  ise; fakat  $E ( Y / X_1 ) = E ( Y / \bar{X}_1 )$  ve  $E ( Y / X_2 ) = E ( Y / \bar{X}_2 )$  bu durumda  $X_1$  ve  $X_2$  birbirini ve  $Y$ ’ yi etkileseler bile, AID analizinde görünmeyeceklerdir, çünkü bunların bireysel hiçbir etkisi yoktur. AID’ in en son versiyonunda ( AID-III ) bu etkileşimlerden bazılarını inceleyen, ileriye dönük bir tercih bulunmaktadır fakat bu avantaj bilgisayarda çalışma zamanını arttırmaktadır.

**Durdurma Kuralları ( Stopping Rules):** Birbiriyle ilişkili bileşenlerden kaynaklanan problemlerden dolayı, durdurma kurallarının belirleyen kararlar önem kazanmaktadır. Eğer kurallar ağacın erken budanmasına neden oluyorsa o zaman önemli değişkenler

kullanılmayabilir. Diğer taraftan zayıf durdurma kuralları ile gürültünün yol açtığı sahte bölünmelerin olasılığı artar.

## 4. CHAID ANALİZİ

### 4.1 CHAID Analizinin Genel Yapısı

CHAID ( Otomatik Ki-Kare Etkileşim Belirleme ) tekniği kategorik bağımlı değişkenler için tasarlanmış AID analizinin bir uzantısıdır. Bu analizde amaç veriyi daha homojen alt gruplara bölmektir. Araştırmacı oldukça homojen bir veri kümesi ile çalışmak ister. Büyük çapta bir veri kümesinin homojen bir alt gruba indirgenmesi problemi demek; bağımlı değişkeni mümkün olduğunca tutumlu bir şekilde açıklayan diğer değişkenleri ve bunlarla ilgili verileri ortaya koymak demektir. İşte CHAID analizi, kategorik değişkenlere ilişkin veri kümesini, bağımlı değişkeni en iyi açıklayacak şekilde detaylı homojen alt gruplara böler. Bu alt gruplar küçük tahmin edici gruplardan oluşur. Seçilen tahmin ediciler daha sonraki ileri analizlerde bağımlı değişkenin tahmininde kullanılacaktır. CHAID, regresyon problemlerinde kullanılabileceği gibi karar ağaçlarının oluşturulmasında etkilidir. Değişkenler arasındaki ilişki lineer yapıdan daha karmaşık ise veride gizli olan bu ilişkiyi bulmak için verinin belli kısımlarını eleme yöntemi olan CHAID kullanılır. “Ki-kare” ismini almasının nedeni algoritmasında birçok çapraz tablonun kullanılması ve istatistiksel önem oranları ile çalışmasıdır. \*

CHAID analizi;

- ❖ Sınıflama ölçme düzeyinde ölçülmüş bir bağımlı değişkeni en iyi şekilde açıklamak için kullanılır,
- ❖ Açıklayıcı değişkenler sınıflayıcı, sıralayıcı ve aralıklı ölçek ile ölçülmüş olabilir,
- ❖ Kayıp verileri yeni bir kategori gibi davranır ve bu kategoriyi p-değeri hesaplamalarına dahil eder,
- ❖ Kategorileri sıralanabilen ya da sıralanmayan, açıklayıcı değişkenlerin yer aldığı veri kümesini, bağımlı değişkene göre detaylı alt kümeler böler,
- ❖ Bu bölünme işlemini gerçekleştirirken, açıklayıcı değişkenlere ait kategorileri, bağımsız olarak yeniden düzenler, yani kategorileştirir.
- ❖ Daha sonraki her bölünmeyi yeniden bağımsız olarak gerçekleştirir.

---

\* Hoare,R.,(2004), “ Using CHAID for Classification Problems”, New Zealand Association 2004 Conference,Wellington,s.1

Yani CHAİD analizi, çok kategorili değişkenlerin yer aldığı büyük bir veri kümesini, benzer kategorileri birleştirerek, önemli sayılan değişkenlere göre bölerek, bir bakıma önceki durumuna oranla özet şekilde tanımlamış olur. Her bir açıklayıcı değişken için kategorilerin anlamlı bir şekilde birleştirilmesinden sonra, bağımlı değişkene göre kontenjans tabloları oluşturularak, Bonferroni p değerleri ile  $\chi^2$  istatistikleri hesaplanır. Açıklayıcı değişkenler birbirleri ile karşılaştırılıp, en küçük Bonferroni p değerine sahip olan açıklayıcı değişkenin kategorilerine göre, veriler alt gruplara ayrılır.

CHAID analizinde her bir açıklayıcı değişken için en iyi bölünme bulunur. Daha sonra açıklayıcı değişkenler en iyi seçilene kadar karşılaştırılır ve seçilen en iyi açıklayıcı değişkene göre yeniden bölünmeler yapılır. Tüm alt bölümler bağımsız olarak yeniden analiz edilir. Her bir açıklayıcı değişken kategorilerini izin verdiği mümkün bölünmeler gerçekleştirilerek  $\chi^2$  testindeki önem derecesine göre kontenjans tabloları oluşturulur. Buradan yola çıkarak CHAID analizi ki-kare istatistiklerini, Bonferroni yaklaşımını ve kategori birleştirme algoritmasını kullanarak araştırmacının ağaç diyagramı ile en önemli açıklayıcı değişkenleri ve bağımlı değişken ile olan etkileşimleri elde etmesini sağlar.

#### 4.2 CHAID Analizinin Algoritması

Bağımlı değişken kategori sayısı  $d \geq 2$  olsun. Analiz edilecek olan belirli bir açıklayıcı değişken  $c \geq 2$  sayıda kategoriye sahip olsun. Analizdeki amaç,  $c \times d$  kontenjans tablosunu açıklayıcı değişkenindeki uygun kategorileri birleştirme yolu ile en anlamlı  $j \times d$  tablosuna indirmektedir. Kavramsal olarak ilk olarak  $T_j^{(i)}$  istatistiğini hesaplarız.  $T_j^{(i)}$   $j \times d$  tablosunu oluşturmadaki  $i$ . metod için,  $\chi^2$  istatistiğidir. ( $j:2,3,4,\dots,c$ ;  $i$ ' nin değişim aralığı açıklayıcı değişkenin tipine bağlıdır.)  $T_j^{(*)} = \max_i T_j^{(i)}$  ise en iyi  $j \times d$  tablo için,  $\chi^2$  istatistiği elde edilmiş olur. Yani, en önemli  $T_j^{(i)}$  seçilir.

Monotonik ya da dichotomous serbest açıklayıcı değişkenin varlığında  $T_j^{(i)}$  Fisher metoduna göre bulunabilir. Bu dinamik program  $c^2$  hesaplarına dayanır.  $d \geq 3$  ve açıklayıcı değişken sıralı kategorilere sahip değilse Fisher metodundan yararlanılamaz. Dreyfus 1977' de dinamik programlarda standart uygulamaların permütasyon tipi problemlerde uygulanabilir olduğunu göstermiştir. Bu çözüm ise  $2^c$  kadardır.\*

\* Kass, G.V., (1980), "An Exploratory Technique For Investigating Large Quantities Of Categorical Data", Applied Statistic, 29(2):121.



Algoritma 3 aşamadan oluşmaktadır; birleştirme, dağıtma ve durdurma.

#### 4.2.1 Birleştirme

**1.Adım:** Her bir açıklayıcı değişken için sırasıyla, açıklayıcı değişkenin kategorileri ile bağımlı değişkenin kategorilerinin çapraz tablosu bulunur ve adım 2 ile 3 uygulanır.

**2.Adım:** Sadece açıklayıcı değişkenin tipi tarafından belirlenen uygun çiftler göz önüne alınarak ,  $2 \times d$  alt tablosunda anlamlılığı düşük olan açıklayıcı değişken kategori çiftleri bulunur. Eğer önem derecesi kritik bir değere ulaşmıyorsa, bu iki kategori birleştirilir. Ve bu birleşim tek bir kategori olarak ele alınır ve bu adım tekrarlanır. Bu işlem açıklayıcı değişkenin kendi içindeki birleşmeleri anlamsız oluncaya kadar devam eder.

**3.Adım:** Açıklayıcı değişkenin tipi tarafından oluşturulan ve orijinal kategorilerin 3 veya daha fazlasının birleştirilmesi ile meydana gelen; her bir bileşik kategori için, birleşmenin tekrar ayrılabilceği en önemli ikili bölünme bulunur. Eğer önem derecesi kritik değerin üzerindeyse, bölünme gerçekleştirilir ve 2.adıma dönülür.

#### 4.2.2 Dağıtma

**4.Adım:** Optimal bir şekilde birleştirilmiş olan, her bir açıklayıcı değişken için önem derecesi hesaplanarak, en büyük önem derecesine sahip olan, diğerlerinden ayrılır. Eğer bu önem derecesi, verilen kriter değerlerlerde büyük ise, veri kümesi seçilen açıklayıcı değişkenin birleştirilmiş kategorilerine göre alt gruplarına bölünür.

#### 4.2.3 Durdurma

**5.Adım:** Verinin analiz edilememiş her bir grubu için, 1. adıma dönülür. Bu adımda en az sayıda gözleme sahip olan gruplar göz ardı edilebilir.

### 4.3 Açıklayıcı Değişkenler Önemliliği

Algoritmanın 4.adımında, indirgenmiş olan kontenjans tablosunun önem derecesinin, test edilmesi gerekir. Eğer orjinal kontenjans tablosunda herhangi bir indirgenme yoksa  $\chi^2$  testi kullanılabilir.  $\chi^2$  testi açıklayıcı değişkenin kategori sayısına bağlıdır, aksi halde çok dikkatli bir şekilde uygulanmalıdır. Kesin sonuçlar bilinmiyorsa ya da orijinal olasılık tablosu indirgenmemiş ise Bonferroni sonuçlarını kullanılması tercih edilir.

Orijinal olasılık tablosunun indirgenmesi; her bir açıklayıcı değişken için , kendi içinde kategorileri anlamlı bir şekilde birleştirilip, en iyi bölünmenin bulunmasından sonra , bağımlı değişkene göre kontenjans tablosunun oluşturulması demektir. Daha sonra  $\chi^2$  ile Bonferroni düzeltilmiş p değerleri hesaplanır. c sayıda kategoriye sahip bir açıklayıcı değişkenin ( $1 \leq r \leq c$ ) olmak üzere , r gruba kaç farklı yolla indirgeneceğini bulmak için CHAID analizindeki her üç tip açıklayıcı değişken için bu Bonferroni çarpanlarının formülleri aşağıdaki gibidir.

#### 4.3.1 Monoton Açıklayıcı Değişkenler

Monoton açıklayıcı değişkenler, kategorileri sıra belirten bir ölçek üzerinde olan açıklayıcı değişkenlerdir. Bu da bitişik kategorilerin sadece birbiri ile gruplandırılabilceği anlamına gelir. Bonferroni çarpanı,kolayca elde edilen binomial çarpandır.

$$B_{\text{monotonik}} = \binom{c-1}{r-1} \quad (4.1)$$

Yaş değişkeni ile eğitim düzeyi açıklayıcı değişkenler olsun, eğitim düzeyi için, bütün kategori düzeyleri ilkokuldan başlayarak sırası ile ortaokul,lise ve üniversite şeklinde sıralanarak devam ediyor ise monoton açıklayıcı değişkene örnek gösterilebilirler.

#### 4.3.2 Serbest Açıklayıcı Değişkenler

Kategorileri tamamen nominal olan açıklayıcı değişkenlere serbest açıklayıcı değişkenler denir. Bu da kategorilerin her şekilde gruplandırılabilceği anlamına gelir. Aşağıdaki gibi formüle edilir.

$$B_{\text{serbest}} = \sum_{i=0}^{r-1} (-1)^i \frac{(r-1)^c}{i!(r-i)!} \quad (4.2)$$

Meslek grupları değişkeni ya da bireyin kimlerle birlikte yaşadığını gösteren açıklayıcı değişkenler bu tip açıklayıcı değişkenlere örnek olabilir.

#### 4.3.3 Hareketli Açıklayıcı Değişkenler

Hareketli açıklayıcı değişkenler, kategorileri sıralayıcı ölçek ile ölçülmüş olup hem sıraya uymayan hem de sıralayıcı ölçekteki pozisyonun bilinmediği tek bir kategorinin olduğu açıklayıcı değişkenlerdir. Kayıp ya da bilinmeyen bir kategori olması durumunda, bu durum ortaya çıkar. Kayıp bilgi Preece tekniği gibi bazı tekniklerle yeniden kodlanabilir. Preece tekniği, kayıp bilginin yerine muhtemel bir sayı ya da değer atama yoluyla kayıp bilgiyi yeniden kodlar.

Hareketli kategori dışındakiler aynı monoton açıklayıcı değişkende olduğu gibi grublandırılır. Hareketli kategorilerde gruplama yapılırken, hareketli kategori tek başına bırakılabilir veya başka bir kategori ya da kategori grubuyla birleştirilebilir. Bonferroni çarpanı, monoton durumdaki çarpan genişletilerek, basit bir şekilde elde edilir.

$$B_{\text{hareketli}} = \binom{c-2}{r-2} + \binom{c-2}{r-1} = \frac{r-1+r(c-r)}{c-1} B_{\text{monotonik}} \quad (4.3)$$

Kategorileri, sıra belirten bir ölçek üzerinde olan eğitim düzeyi monoton açıklayıcı değişkeni için, bir eğitim düzeyine ait, örneğin ortaokul mezunu olanlar ile ilgili veriler kayıp ise bu açıklayıcı değişken ise bu açıklayıcı değişken artık hareketli açıklayıcı değişken grubuna girer.

## 5. AID ANALİZİ ile CHAID ANALİZİNİN KARŞILAŞTIRILMASI

Bu çalışmada ağaç temelli sınıflandırma işlemlerinde kullanılan AID ve CHAID analizlerine yer verilmiştir. Birbirleri ile oldukça benzer yönleri olmasına rağmen aralarında bazı farklılıklarda bulunmaktadır. AID ve CHAID analizlerini karşılaştırsak;

- AID analizinde, açıklayıcı değişkenlerin sahip olduğu kategorilerde, herhangi bir birleştirme işlemi yapılmadan, açıklayıcı değişkenlerin en açıklayıcı olduğu düşünülenlerden herhangi biri kullanılarak (nicel bağımlı değişken), veri başarılı olacak şekilde iki eşit parçaya ayrılır.  
CHAID analizinde, açıklayıcı değişkenlerin kategorileri arasında uygun olan birleştirmeler yapılarak istatistiksel değerler yardımıyla bağımlı değişken ile ilişkili olan açıklayıcı değişkenin kategorisi iki ve ya daha fazla bölünür.
- AID analizinde, sıra belirten açıklayıcı değişkenler “monotonik” olarak adlandırılır. Tamamen nominal ölçümlü diğer açıklayıcı değişkenler “serbest” olarak adlandırılır.  
CHAID analizinde, serbest ve monoton açıklayıcı değişkenlere ek olarak hareketli açıklayıcı değişkenleri de göz önüne alarak kayıp bilgi sorunu giderilmeye çalışılır. Kass 1980 yılında yaptığı çalışmada “floating predictor” ( yüzen, hareketli ) adında yeni bir açıklayıcı değişken tanımlamış ve bunu da CHAID analizine dahil etmiştir.
- AID, aralıklı ölçekle ölçülmüş bir bağımlı değişken üzerinde kullanılmaktadır ve her ikili parçada gruplar arası kareler toplamı yani F-istatistiği en büyük değerine ulaştırılır.  
CHAID, sınıflayıcı ( nominal ) ölçekle ölçülmüş bağımlı değişkene uygulanır ve mutlaka ikili bölünme olması gerekmeyen her bölünmede ki-kare istatistiğini maksimize eder.
- AID analizinde, veriyi bölmek için herhangi bir önem testi kullanılmaz.  
CHAID analizinde, bir önem testi kullanılmaktadır.
- AID analizi hatalı kullanımları olduğu ( Doyle, 1973 ) ve veri kümesinin doğal yapısındaki değişkenliği dikkate almadığından ötürü ciddi kısıtlamalara sahip olduğu

gerekçesiyle eleştirilmektedir. MacLachlan ve Johansson (1981) göre bazı sınırlamalar dikkatli analizler, büyük örneklemeler ve tekrarlanan alt örneklemeler ile çözülebilir.

CHAID analizi, bu problemi bölünmeler esnasında önem testlerini kullanarak aşar. Böylelikle kritere göre açıklayıcı değişkenin bölünemez olduğuna karar verilmesini sağlayan çoklu bölünmelerin oluşumuna ve kontrolüne izin verir. Bu da deneyimsiz araştırmacılar için uygun analiz olmasını ve AID ile ilgili şüphelerin yok olmasını sağlamaktadır.

- AID analizi, daha fazla kategoriye sahip olan açıklayıcı değişkenleri destekleyerek bölünmelerde yanıltıcı sonuçlar elde edilmesini sağlayabilir.

CHAID analizi, ise önem testlerini kullanarak AID tekniğinin bu eğilimlerini hükümsüz hale getirir.

CHAID analizi, AID analizinin limitlerini dikkate alarak geliştirilmiştir. Kısaca, AID ve CHAID analizleri analistlerin ihtiyaçları göz önüne alınarak karşılaştırıldığında, CHAID, AID' den daha uygun ve daha esnek yapıda olduğu söylenebilir. Açıkça CHAID kullanımı kolay, sezgisel sonuçlar üreten ve istatistiksel olarak geçerli olan bir analizdir. Bu nedendir ki CHAID hem çalışmanın içindeki denekler hem de analistlerin ihtiyaçları için tatmin edicidir.

## 6. UYGULAMA

Çalışmanın uygulaması bir şirketin personeline gerçekleştirilen çalışmalardan memnuniyetini ve önemini belirlemek amacıyla yapılan anket çalışmasının verileri üzerinden yapılmıştır. Sorular toplam 3231 personel tarafından cevaplanmıştır. Anket formunda 31 soru vardır ve sorular eğitim, işe alma, performans değerlendirme, ücretlendirme, yan faydalar, kariyer planlanma, özlük işleri ve bordrolama, kurum içi iletişim, insan kaynakları departmanı hakkındaki görüşler, faaliyet alanı ve personelin şirketteki unvanı konuları hakkındadır. Eğitim, işe alma, performans değerlendirme, ücretlendirme, yan faydalar, kariyer planlanma, özlük işleri ve bordrolama, kurum içi iletişim soruları için cevaplar bu konu benim için önemli ve şirketimizdeki bu tür çalışmalardan memnuniyet şeklinde iki alt başlıkta 5 seçenekli likert ölçeği üzerinden alınmıştır. İnsan kaynakları departmanı hakkındaki görüşler sorularına ise 5 seçenekli likert ölçeği üzerinden katılıp katılmadıkları sorulmuştur. Faaliyet alanı holding ve bölge diye ayrılmış sınıflayıcı ölçekle, unvan ise müdür ve üzeri, uzman/şef, satıcı, motokurye/sürücü, asistan/görevli olmak üzere sınıflayıcı ölçekle ölçülmüştür. Ek 1 'de anket soru formu yer almaktadır.

Bölgeler bazında bir ayrıma gidilmiştir ve toplam 7 bölge üzerinden analiz yapılmıştır. CHAID ile amaçlanan bu 7 bölge arasında mevcut sorular üzerinde nasıl bir gruptamanın olacağıdır. Bağımlı değişken bölgeler olarak alınmış ve sınıflayıcı (nominal) ölçek ile ölçülmüştür. Açıklayıcı değişkenlerden faaliyet alanı ve unvan değişkenleri sınıflayıcı ölçekle, diğer açıklayıcı değişkenler 5 kategoriden oluşan sıralayıcı (ordinal) ölçekle ölçülmüştür. Çalışmada yer alan bölgelerin dağılımı şu şekildedir;

Çizelge 6.1 Personelin bölgelere göre dağılımı

| Bölge Adı    | Frekans | Yüzde | Kümülatif Yüzde |
|--------------|---------|-------|-----------------|
| AKDENİZ      | 262     | 8,1   | 8,1             |
| EGE BÖLGESİ  | 454     | 14,1  | 22,2            |
| GÜNEYANADOLU | 600     | 18,6  | 40,7            |
| İST. AVRUPA  | 738     | 22,8  | 63,6            |
| İST.ANADOLU  | 324     | 10,0  | 73,6            |
| KARADENİZ    | 410     | 12,7  | 86,3            |
| ORTAANADOLU  | 443     | 13,7  | 100,0           |
| TOPLAM       | 3231    | 100,0 |                 |

SPSS paket programında CHAID 3.0 modülü kullanılarak analizler yapılmıştır. Bölgelerin bağımlı değişken olduğu analizlerde risk matrisi yani yanlış sınıflama oranı çok yüksek çıkmaktadır. Yaklaşık olarak % 35 yanlış sınıflama yapılmaktadır. Bu yüzden çalışmada bağımlı değişken olarak faaliyet alanı alınmasının daha uygun olduğuna karar verilmiştir. Böylelikle faaliyet alanı bağımlı değişken diğer değişkenler ise açıklayıcı değişken olarak alındığında risk matrisi % 94 doğru sınıflama yapıldığını göstermektedir. Veriler üzerinde daha farklı sonuçları görmek için insan kaynakları departmanı ile ilgili görüşleri kendi içinde alt gruplara ayırmaya; diğer eğitim, işe alma, performans değerlendirme, ücretlendirme, yan faydalar, kariyer planlanma, özlük işleri ve bordrolama, kurum içi iletişim gibi konularındaki memnuniyet ve önem kendi içinde alt gruplara ayrılmaya çalışılmıştır.

## 7. SONUÇLAR VE DEĞERLENDİRME

### 7.1 CHAID Analizi 1

Bağımlı değişken faaliyet alanı olarak alınmıştır, bağımlı değişken 2 kategoriden oluşan sınıflayıcı ölçekle ölçülmüştür. Açıklayıcı değişkenler ise 5 kategoriden oluşan sıralayıcı ölçekle ölçülmüş eğitim, işe alma, performans değerlendirme, ücretlendirme, yan faydalar, kariyer planlanma, özlük işleri ve bordrolama, kurum içi iletişim ile ilgili sorular, 7 kategoriden oluşan sınıflayıcı ölçekle ölçülmüş bölge adı ve 6 kategoriden oluşan sınıflayıcı ölçekle ölçülmüş unvanlardır, bu değişkenler dikkate alınarak faaliyet alanı değişkeni alt gruplara ayrılacaktır. Ağaç genişliği 10 olarak belirlenmiştir. Anakütle büyüklüğünün küçük olması nedeniyle tüm değişkenlerin değerlendirilmesi için çocuk düğümünün büyüklüğü 10 olduğunda ve terminal düğümün sayısının 20 olması durumunda ağacın durdurulmasına karar verilmiştir.

Faaliyet alanı öncelikle (p değeri= 0) diğer değişkenlere göre en küçük olan bölgeler ile alt gruba ayrılmıştır, bölgelerde Ege ile Karadeniz, Akdeniz ile İstanbul Anadolu birleştirilerek 5 kategori olarak değerlendirilmiştir. Bölgeleri tek tek incelersek;

Orta Anadolu bölgesindeki personel kurum içi iletişim ile ilgili şirkette yapılan çalışmalara (C8b) dikkat etmektedir. Orta Anadolu’da bulunan 443 personelden bu soruya 14’ü cevap vermemiş, 223’ü 1 ile 2 derece, 206’sı 3,4 ve 5 vererek değerlendirmişlerdir. Orta derece memnun ile çok memnun arasındaki personel şirketteki eğitim ve geliştirme çalışmalarına göre 160’ı 1,2,3 ve 4 ile cevaplayanlar; 46’sı 5 ile cevaplayanlar ve cevapsızlar olarak gruplandırılmıştır. Bu çok memnun ve cevapsız bırakanların bir arada gruplandırıldığı alt kümede %40’ı bölgede ve %60’ı holdingde görev yapmaktadır.

Akdeniz ve İstanbul Anadolu bölgelerinin birleştirilmesi ile oluşan alt kümede unvanlara (S31) göre 3 gruba ayrılmıştır. 124’ü unvan belirtmemiştir, yani anakütlenin %3,84’dür, 118’i motokurye/sürücüdür ve bu personel işe almanın kendisi için önemine (C2a) değinmiş, 12’si hiç memnun değilim ile orta derece memnunum aralığında yer almaktadır bu grubun hepsi holdingde görev yapmaktadır ve 106’sı memnun, çok memnun ve cevapsız bırakanların



birleşmesi ile oluşan alt kümededir ve yine hepsi holdingde görev yapmaktadır. Asistan, müdür ve üzeri, satıcı ve uzman/şef unvanındakilerin birleşmesi ile oluşan alt grupta 334 kişi bulunmaktadır ve kurum içi iletişimin kendisi için önemine (C8a) dikkat etmektedir ve tüm personel içinde %10,28'si önemli olduğunu düşünmektedir.

Ege ve Karadeniz bölgelerinden birleşmesi ile oluşan alt kümede 864 kişi yer almaktadır ve anakütlenin % 27'si bu gruba girmektedir. Şirketin sunduğu yan faydalardan memnuniyet (C5b) ile bu bölgenin önemli derecede ilişkili olduğu görülmüştür. 658'si hiç memnun değilim ve memnunum aralığında cevaplayanlar ile cevapsızlardan oluşmakta bu grubunda 340'si Ege bölgesindedir, tüm personelin %10'u; 318'i ise Karadeniz bölgesindedir yine bu grupta tüm personelin %10'una denk gelmektedir. Yan faydalardan çok memnunum diyen personelin 65'i asistan ve müdür-üzeri, 96'sı motokurye/sürücü ve unvan belirtmeyenler, 45'i satıcı,uzman/şeftir. Anakütlenin %1,4'ünü içine alan satıcı ve uzman/şef grubunda yer alanların 27'si Ege, 18'i Karadeniz bölgesindedir. Egede bulunan 27 kişilik personel özlük işlerinden memnuniyeti değerlendirmiş ve çok memnunum diyen 17 personelin hepsinin bölgede görev yaptığı görülmüştür.

Güney Anadolu Bölgesindeki personel tüm personelin %19'dur. 218'i asistan ve uzman/şef unvanı ile çalışmaktadır, kurum içi iletişim çalışmalarından 50'si hiç memnun olmadığını belirtmiş bu gruptakilerin eğitim ve geliştirme çalışmalarına da dikkat ettiği gözlenmiştir, 172'si hiç memnun değilim ve cevapsızlar, 14'ü çok memnunum ve geri kalan 19'u 2,3,4 ile değerlendirme yapmıştır. 274'ü müdür-üzeri, motokurye/sürücü ve satıcı unvanına sahiptir bu gruptakilerde yine asistan ve şef alt grubundakiler gibi eğitim ve geliştirme çalışmalarına dikkat emektedir ve 55'i hiç memnun olmayanları (tüm personelin %1,7'si) ve 219'u memnun değilim, çok memnunum diyenler ve cevapsız grubunu (tüm personelin %6,8'i) oluşturmaktadır.

İstanbul Avrupa Bölgesi ise en fazla personelin bulunduğu bölgedir ve ağacın en fazla dallandığı bölgedir. Bu bölgedeki 738 personelin 424'ü asistan ve satıcı, 139'u müdür-üzeri ve unvan belirtmeyenler, 175'i Motokurye/sürücü ve uzman pozisyonundadır. Anakütlenin %13'üne denk gelen asistan ve satıcı alt grubu işe alınmanın önemine (C2a) dikkat çekmiş ve çoğunluğu yani 225'i çok memnundur ve bu grubun %74'ü holdingde, %36'sı bölgede çalışmaktadır. İşe alınmanın kendisi için hiç önemli olmadığını söyleyenler ile cevapsız bırakanlar ise 24 kişidir. 175'i ise aynı soruya 2,3,4 ile cevap vermiş kariyer planlama

çalışmalarına (C6b) dikkat ettiği görülmüştür; 30'u hiç memnun olmadığı, 46'sı orta derecede memnun olduğu, 99'u yani çoğunluğu çok memnun olanlar ile cevapsız bırakanların olduğu gruplara ayrılmıştır. Tüm personelin %1,4'ünü oluşturan orta derece memnun olanlar performans değerlendirme konusunun kendisi için önemine göre çok memnun olanlar 16 kişiden, hiç memnun olmayanlar ile memnun olanlar arasında cevaplayanlar 30 kişiden (%93'ü bölgede görev yapıyor) oluşan alt kümeye ayrılmıştır. Çoğunluğun oluşturduğu çok memnun ve cevapsız grubunda ise 37'si satıcı, 62'si asistan pozisyonundadır. 11 asistan kurum içi iletişim çalışmalarından (C8b) hiç memnun olmadığını belirtmiştir. %4,3 oranındaki müdür-üzeri ile unvan belirtmeyenler grubunda ise eğitim ve geliştirme konusunda yapılan çalışmalara göre 53'ü hiç memnun değilim ve memnun değilim kategorilerin birleşmesi ile oluşan alt grupta, 86'sı orta derecede memnunlar ile çok memnunlar ve cevapsızların oluşturduğu alt grupta yer almaktadır. Motokurye/sürücü ve uzman/şef kategorilerinin birleşmesi ile oluşan grupta 175 kişi bulunmaktadır ve işe alınma konusunda yapılan çalışmalara dikkat ettikleri görülmüştür, bu konuda çok memnun olan 21 kişi şirket tarafından sağlanan yan faydalara değinerek çoğunluğu çok memnun olduğunu belirtmiştir.

Çizelge 7.1 C1a ile C8b arasındaki cevapların ve bölge adı, unvanların açıklayıcı değişken olması durumunda risk matrisi

| Yanlış Sınıflama Matrisi        |                    |            |     |        |
|---------------------------------|--------------------|------------|-----|--------|
|                                 | Gerçek Kategoriler |            |     |        |
|                                 |                    | 1          | 2   | Toplam |
| Tahmini Kategoriler             | 1                  | 2881       | 292 | 3173   |
|                                 | 2                  | 19         | 39  | 58     |
|                                 | Toplam             | 2900       | 331 | 3231   |
| Risk İstatistikleri             |                    |            |     |        |
| RiskTahmini                     |                    | 0,101517   |     |        |
| Risk Tahmininin Standart Hatası |                    | 0,00531319 |     |        |

Çizelge 7.1 göre CHAID analizi sonunda yanlış sınıflama % 10,15 dır, yani % 89,85 oranında doğru sınıflama yapılmıştır.

## 7.2 CHAID Analizi 2

Bağımlı değişken faaliyet alanı olarak alınmıştır, 2 kategoriden oluşan sınıflayıcı ölçekle ölçülmüştür. Açıklayıcı değişkenler ise 5 kategoriden oluşan sıralayıcı ölçekle ölçülmüş insan kaynakları hakkındaki personelin görüşleri ve işinden memnuniyeti, 7 kategoriden oluşan sınıflayıcı ölçekle ölçülmüş bölge adı ve 6 kategoriden oluşan sınıflayıcı ölçekle ölçülmüş unvanlardır, bu değişkenler dikkate alınarak faaliyet alanı değişkeni alt gruplara ayrılacaktır. Ağaç genişliği 10 olarak belirlenmiştir. Anakütle büyüklüğünün küçük olması nedeniyle tüm değişkenlerin değerlendirilmesi için diğer analizde olduğu gibi çocuk düğümünün büyüklüğü 10 ve terminal düğümün sayısının 20 olması durumunda ağacın durdurulmasına karar verilmiştir.

CHAID Analizi 1 bölümündeki analiz gibi faaliyet alanı ile bölgeler arasında anlamlı bir ilişki olduğuna ( $p=0$ ) karar verilmiş ve en küçük p değerine sahip olması nedeniyle CHAID alt gruplara ayırmayı ilk olarak bölgelere göre yapmıştır. Buna göre teke tek incelersek;

443 personelin 375'nin holdingde, 65'nin bölgede görev yaptığı Orta Anadolu bölgesinde çoğunluğun genel olarak işinden 2 ile 5 arasında memnun olduğu ortaya çıkmıştır ve tüm personelin %9,9'u demektir. 52'si hiç memnun olmadığını söylemiştir.

İstanbul Anadolu bölgesinde toplam 738 kişinin 569'u holdingde ve 168'i bölgede görev yapmaktadır. Unvanlara göre alt gruplara ayrıldığında 424'ü asistan ve satıcı, 139'u müdür-üzeri ve unvan belirtmeyenler, 175'i ise motokurye/sürücü ve uzman/şef pozisyonundadır. Müdür-üzeri pozisyonu ile unvan belirtmeyenler insan kaynaklarının şirkette yeni uygulamalar gerçekleştiği konusunda (S20), 22 kişi, hiç katılmadığını söylemiştir bu grubun yarısı bölgede yarısı holdingde görevlendirilmiştir. 117'i ise yeni uygulamalar gerçekleştirdiği konusunda katılmıyorum ile bütünüyle katılıyorum aralığında cevap verenler ve cevapsız bırakanlardır. % 13 oranındaki asistan ve satıcı grubundakilerin çoğunluğu işinden genel olarak memnuniyet konusunda (S29) çok memnun ile cevapsızların oluşturduğu gruptadır, bu grubun 169'u holdingde, 47'si bölgede görev yapmaktadır. Bu alt kümenin elemanları çalışanların insan kaynakları departmanına saygı duydukları konusunda (S15) 28'si cevap vermemiş, 16'sı hiç katılmadığını ve 173'ü katılmıyorum ile bütünüyle katılıyorum aralığında

cevap vermiştir. Bu 173 kişinin %75'i holdingde, %25'i bölgede görevlidir ve 79'u satıcı, 94'u asistandır. İşinden genel olarak memnuniyet konusunda hiç memnun olmayanlarla memnun olanlar 207 kişidir ve %57'si holdingde, %43'ü bölgede görevlidir.

%18,57 oranındaki Güney Anadolu bölgesindekilerin 108'i unvan belirtmemiş, 274'ü müdür-üzeri, satıcı ve motokurye/sürücü kategorilerinin birleştiği kümede ve 218'i ise asistan ve uzman/şef kategorilerinin birleşmesiyle oluşan kümededir. %8,5 oranındaki müdür-üzeri, satıcı ve motokurye/sürücü unvanındakileri insan kaynakları departmanının olanaklar konusunda bilgi verici olması konusuna (S12) dikkat etmektedir ve 22'si bu konuyu cevapsız bırakmış, 33'ü çok memnun olduğunu ve 219'u yani çoğunluk ise 1 ile 4 aralığında cevap vererek aynı grupta yer almıştır. Asistan ve uzman/şef alt grubundakiler iş tanımları hakkında insan kaynakları departmanının bilgili olması hakkında 114'ü hiç memnun ve memnun değilim kategorilerin birleşmesi ile oluşan gruptadır ve bu grupta 88 kişi şirket için gerekli bir departman olduğunu (S9) belirtmektedir. İş tanımları ile ilgili soruya 94 kişi 3 ile 5 arasında cevap vermiştir ve bu grup üyeleri sorunlarının insan kaynakları departmanı tarafından bilindiğine (S24) inanmaktadır ve personele karşı takdir edici olması hakkında ise 11'i bütünüyle katıldığını söylemiştir.

Ege ve Karadeniz bölgesinin birleşmesi ile oluşan grupta insan kaynakları departmanının gerekli olduğuna 864 kişinin 829'u bütünüyle katılanlar ile ne katılıyorum ne katılmıyorum aralığında cevap verenler ve cevap vermeyenlerden oluştuğu, 35 kişinin ise hiç katılmayanlar ve katılmıyorum aralığında cevap verenlerden oluştuğu görülmüştür. %1,1 oranındaki hiç katılmıyorum ve katılmıyorum grubundakiler 13'ü işinden genel anlamda hiç memnun olmadığını ve memnun olmadığını belirtmiştir. Şirket için gerekli bir departman olduğu konusunda %25,7 oranındaki bütünüyle katılanlar ile ne katılıyorum ne katılmıyorum aralığında cevap verenler ve cevap vermeyenler alt grubu İK departmanının şirkette söz sahibi olması konusunda çoğunluk, 543 kişi, orta derecede ve olumlu yanıt vermiştir, bu alt kümenin hemen hemen tamamı holdingde görevlidir.

3231 personelin 586'sı CHAID analizi ile Akdeniz ve İstanbul Anadolu bölgelerinin birleşmesi ile oluşan alt gruptadır. Unvanlara göre 124'ü unvan belirtmeyenlerden, 118'si motokurye/sürücü ve 344 asistan, müdür-üzeri, satıcı ve uzman/şef kategorilerin birleşmesiyle oluşan alt gruptadır. Motokurye/sürücü pozisyonundakiler genel anlamda işinden memnuniyet konusunda 104'ü yani çoğunluğu olumlu yanıt verenler ve cevapsızların oluşturduğu alt

kümededir. Tüm personelin %10,7'sinin yer aldığı asistan, müdür-üzeri, satıcı ve uzman/şef alt grubundakiler İK departmanının şirketin başarılı olmasına olumlu yönde katkısı olduğu (S16) konusunda çoğunluk 2 ile 4 aralığında cevap vermiştir. Bu soruya 101 kişi bütünüyle katıldığını söyleyenler içinde 34'ü Akdeniz ve 67'si İstanbul Anadolu bölgesinde görev yapmaktadır.

Çizelge 7.2 S9 ile S29 arasındaki cevapların ve bölge adı, unvanların açıklayıcı değişken olması durumunda risk matrisi

| Yanlış Sınıflama Matrisi        |                    |            |     |        |
|---------------------------------|--------------------|------------|-----|--------|
|                                 | Gerçek Kategoriler |            |     |        |
|                                 |                    | 1          | 2   | Toplam |
| Tahmini Kategoriler             | 1                  | 2888       | 316 | 3204   |
|                                 | 2                  | 12         | 15  | 27     |
|                                 | Toplam             | 2900       | 331 | 3231   |
| Risk İstatistikleri             |                    |            |     |        |
| RiskTahmini                     |                    | 0,112349   |     |        |
| Risk Tahmininin Standart Hatası |                    | 0,00555568 |     |        |

Çizelge 7.2 göre CHAID analizi sonunda yanlış sınıflama % 11,23 dür, yani % 88,77 oranında doğru sınıflama yapılmıştır. Örneğin 1 grupta yer alması gerekirken sadece 12 tanesi yanlış sınıfta yer almıştır. Bu değerler oluşan ağacın güvenilirliği açısından tatmin edici sonuçlardır.

### 7.3 CHAID Analizi 3

Çalışmanın son bölümünde tüm değişkenleri kullanarak bağımlı değişkenin yine faaliyet alanı olması durumundaki CHAID analizi sonuçları yorumlanacaktır. Ayrıca ağaç durdurma durumları çocuk ve terminal düğümün büyüklüğü dikkate alınarak ilişkili değişkenlerin değişip değişmediği de incelenecektir. İlk olarak terminal ve çocuk düğüm büyüklüğü sıfır (0) alınmıştır. Buna göre bölgeleri tek tek incelersek;

Orta Anadolu bölgesin genel anlamda işinden memnuniyete göre ayırma yapılmıştır, böylece 318'i çoğunluğu 3 ile 5 arasında cevap vermiştir, 52'si hiç memnun ile memnun değilim diyerek bu grup üyeleri şirket tarafından sunulan yan faydalara (C5b) dikkat çekerek çoğunluk hiç memnun olmadığını belirtmiştir. İşinden memnuniyet konusunda 73 kişi cevap vermemiştir bu gruptakiler kurum içi iletişimi ile ilgili çalışmalardan 38 kişisi 1 ile 2 aralığında cevap vermiş ve bu grup üyeleri insan kaynaklarının yeni çalışmalara öncülük etmesi konusunda çoğunluk olarak memnun olduğunu belirtmiştir. Ancak bu grupta memnun değilim diyen 5 kişiden işe alım çalışmalarından çoğunluğunun memnuniyetsiz olduğu görülmüştür. Kurum içi iletişim alt grubunda 3 ile 5 arasında cevap veren 34 kişiden 29'u insan kaynaklarının düzenli ve titiz çalıştığına (S19) 2 ile 5 arasında cevap vermiştir. Eğer ağaç durdurma kurallarında çocuk düğümün 5, terminal düğümün 10 alınması durumunda Orta Anadolu bölgesinde ilk değişken aynı oranlarıyla işinden genel anlamda memnuniyet sorusunun alt gruplara ayrılması ile olmuştur. Sadece işinden memnun olmayanlar alt gruplara ayrılmamış, ağacın bu dalı daha fazla alt gruplara ayrılmamıştır. İşinden memnunum sorusunu cevapsız bırakanlar ise kurum içi iletişime göre değil İK'nın yeni uygulamalar gerçekleştirir sorusuna değinmişlerdir.

Güney Anadolu bölgesinde 108'i unvan belirtmeyen, 274'ü müdür-üzeri, motokurye/sürücü,satıcı, 218'i asistan, uzman/şef olmak üzere 3 alt gruba ayrılmıştır. Çoğunluğun bulunduğu müdür-üzeri, motokurye/sürücü,satıcı grubunda performans değerlendirmenin kişi için önemi (S3b) personelin dikkat ettiği konudur ve çoğunluk yani 271'i 2 ile 5 cevaplayanlar ve cevapsız kategorilerinin birleşmesi ile oluşan alt kümedir. Ayrıca bu gruptakiler insan kaynaklarının personel tarafından sevilmesi (S22) konusunda 11'i hiç katılmıyorum grubunu oluşturmuş ve bu grup üyelerinin çoğunluğu iş tanımları hakkında bilgi sahibi (S13) olmadığını söylemiştir. %6,8'nin asistan, uzman/şef olduğu alt grupta insan kaynaklarının takdir edici olması konusunda (S23) 39'u çok memnun olduğunu söylemiş ve bu grup üyeleri özlük işleri çalışmalarından memnuniyetlerini 7'si 1 ile 2 aralığında, 32'si 3 ile 5 aralığında cevap vererek ifade etmiştir. Bu alt gruptaki 1 ile 2 aralığında cevap verenlerden işe alınma konusunda kendisi için (C2a) çok önemli olduğunu söyleyen 2 kişidir. Özlük işleri çalışmalarında 3 ile 5 arasında cevaplayanlar ve cevapsız kategorilerinin birleşmesi ile oluşan 32 kişi performans değerlendirmenin kendisi için önemi (C3a) hakkında çoğunluk yani 26 kişi çok önemli olduğunu ifade etmiştir. İnsan kaynaklarının başarıları

takdir etmesi konusunda cevap vermeyenler, 8 kişi, işe alınma çalışmalarının memnun edici (C2a) olduğunu söylemiştir. Asistan, uzman/şef pozisyonundaki 218 kişi içinde 171'nin İK'nın takdir ediciliği konusunda 1 ile 4 arasında cevap verdiği görülmüştür ve bu alt gruptakiler kurum içi iletişime dikkat çekerek 39 kişi hiç memnun olmadığını söylemiştir. Hiç memnun olmadığını söyleyenler İK'nın kariyer hedeflerle ilgili olması (S18) hakkında çoğunluk 1 ile 3 aralığında cevap verenler ile cevap vermeyenlerin birleştiği alt grupta yer almaktadır. Ağaç durdurma işleminde terminal düğümün 10, çocuk düğümünün 5 olması halinde Güney Anadolu bölgesinde unvanlara göre yine aynı şekilde alt gruplar ayrılmış ancak müdür-üzeri, motokurye/sürücü ve satıcı grubunda performans değerlendirme yerine İK'nın olanaklar hakkında personele bilgi vermesi değişkeni üzerinden ayırma yapılmıştır.

Ege ve Karadeniz bölgesinin birleşmesinin öngören CHAID analizi unvanlar değişkenine göre alt gruplara ayırmıştır. 394 kişinin bulunduğu müdür-üzeri, motokurye/sürücü ve unvan belirtmeyenlerden oluşan alt grupta İK'nın personel tarafından saygı görmesi konusunda (S15) 27'si 2 ile 4 aralığında cevap vererek İK'nın yönetim ile iletişim sağlaması hakkında (S27) çoğunluk yani 18 kişi hiç katılmıyorum, katılmıyorum ve cevapsız kategorilerinin birleşmesi ile oluşan alt grup üyeleridir ve çoğunluğu 16'sı İK departmanının yapılan iş hakkında bilgi sahibi (S13) olmadığını ifade etmiştir. 339 kişiden oluşan İK'ya saygı duyulduğuna bütünüyle katılanların ve cevap vermeyenlerin oluşturduğu alt grup ise ücretlendirmenin kendisi için önemli olduğunu söylemiş ve genel anlamda işinden memnun olduğunu ifade etmiştir. Bu grupta işinden memnun olmayanlar arasında bölgeler göre ayırım yapıldığında 7 kişiden 5'i Ege, 2'si Karadeniz bölgesinde görevlidir. 864 kişiden 470'i asistan, satıcı, uzman/şef unvanı ile çalışmaktadır. İK departmanının gerekli olduğu konusunda (S9) olumsuz düşünen 22 kişiden şirket tarafından sağlanan yan faydalar (C5b) konusunda 10 kişi çok memnun olduğunu ifade etmiştir. İK gerekli diyenlerden 448'i, büyük bir kısmı, şirket tarafından sağlanan yan faydalar hakkında çok memnun olduğunu söyleyen 93 kişi İK'nın sözü geçen bir departman olması konusunda 17 kişi ile hiç katılmıyorum ve katılmıyorum grubunda olduğu ve bu grubun 9'nunun Ege, 10'nun Karadeniz bölgesinde çalıştığı ayırımı yapılmıştır. Yan faydalarla ilgili çalışmalar sorusuna 1 ile 4 arasında cevap verenler ile cevapsızların oluşturduğu alt grup üyelerinin, 335 kişidir, 166'si Ege'de ve 189'u Karadeniz'de görevlidir. Ege'de görev yapanlar kariyer planlama çalışmalarının kendisi için önemli olduğunu düşünen 163 kişi işe alınma çalışmalarını değerlendirerek 156'sı 1 ile 4 aralığında cevap verenler ve cevapsızların oluşturduğu alt gruptur ve çok memnunun diyen 7 kişiden 6'sı işinden genel anlamda hiç memnun olmayan ya da cevap vermeyenlerin

grubudur. Karadeniz’de ise İK’nın şirket içi iletişimde (S11) başarısı hakkında çok memnun olan ya da cevap vermeyenler grubunda 32 kişiden 13 tanesi kurum içi iletişim çalışmalarından çok memnun olduğunu söylemiştir. S11’e 1 ile 4 aralığında cevap verenlerden 151’i işinden genel anlamda memnun olduğunu belirterek eğitim ve geliştirme çalışmalarının önemine dikkat ettiği görülmüştür. Eğitim ve geliştirme çalışmaları hakkında çoğunluk önemli olduğunu düşünen gruptadır ve ayrıca bu grup üyeleri içinde eğitim ve geliştirme çalışmalarından 65 kişinin memnun olduğu görülmüştür. Ege ve Karadeniz bölgesinde terminal düğümün 10, çocuk düğümün 5 olması halinde yine ilk önemli değişken unvanlar olmaktadır. Ancak unvan belirtmeyen 8 kişi alt gruplara ayrılmamıştır. Müdür-üzeri, motokurye/sürücü olan alt grup üyeleri ağacın son dallarında İK’nın yönetim ile personel arasında iletişim sağlar sorusuna verilen cevaplara göre alt gruplara ayrılması olmamıştır. Aynı şekilde bu grupta İK’nın saygı duyulan departman olduğuna çok katılıyorum diyen ve cevaplamayanlar ücretlendirmenin önemine göre değil işinden genel olarak memnuniyete göre ayırım yapılmıştır. Asistan, satıcı ve uzman/şef alt grubunda da sadece alt dallarda bölgelere göre ayırımdan sonra kariyer planlamanın önemine göre değil ücretlendirmenin önemine göre ayırım yapılmıştır.

Akdeniz ve İstanbul Anadolu bölgelerinin birleştirildiği grupta 344’ü asistan, müdür-üzeri, satıcı, uzman/şef, 124’ü unvan belirtmeyenler ve 118’i motokurye/sürücü grubunda yer alacak şekilde ayırım yapılmıştır. Buradan yola çıkarak motokurye/sürücü grubu üyeleri ücretlendirmenin önemi (S4a) konusunda 107 kişi önemli değil ve çok önemli aralığında cevap verenler ya da cevapsız bırakanlardır ve bu grup İK’nın şirket başarısına olumlu katkısı olduğunu (S16) çoğunluk olarak 2 ile 5 aralığında cevaplayarak ya da cevapsız bırakarak, personele karşı dürüstlüğüne yine 102 kişi yani çoğunluğu 2 ile 5 aralığında cevaplayarak ya da cevapsız bırakarak şeklinde alt gruplara ayrılmıştır. Unvan belirtmeyenler ise ücretlendirmeden çok memnun olan 19 kişiden 17’si ücretlendirmenin çok önemli olduğunu söylemiştir. Ücretlendirmeden memnuniyetini 1 ile 4 aralığında cevaplayanların 71’i İstanbul Anadolu bölgesindedir, eğitim ve geliştirmenin çok önemli olduğunu söyleyerek İK’nın yönetim-çalışan dengesi sağladığına (S17) inanmaktadır. Ayrıca İK’nın olanaklar hakkında bilgi verici olduğu konusunda çok memnun olan grup 8 kişidir ve bu kişilerden 6’sı işinden genel anlamda çok memnun olduğunu söyleyen ya da cevapsız bırakanların oluşturduğu alt gruptur. Ücretlendirmeden memnuniyet konusunda 27 kişi Akdeniz Bölgesindedir, ücretlendirmeden memnuniyet konusunda hiç memnun değilim diyen 9 kişiden 5’i işe alınma çalışmalarından memnuniyetsiz olduklarını söyleyerek kariyer planlamanın kendisi için



önemine dikkate çekmiştir. Akdeniz’de ücretlendirmeden 2 ile 5 aralığında cevap veren 18 kişiden 16’sı kariyer planlama çalışmalarından 2 ile 5 aralığında cevap verenlerin grubudur. Asistan, müdür-üzeri, satıcı ve uzman/şef alt grubundakiler özlük işlerinin önemi (C7a) hakkında cevap vermeyenlerden, 13 kişidir, 8 kişi ücretlendirmenin kendisi için çok önemli olduğunu söylemiştir. C7a sorusuna 2 ile 5 aralığında cevap veren 330 kişiden 320’si İK’nın personele dürüst yaklaştığını (S10) söylemektedir, hiç dürüst olmadığını söyleyen 10 kişiden 6’sı İK’nın çalışanların sorunlarından habersiz olduğunu düşünmektedir. Bu bölgede terminal düğüm 10, çocuk düğüm 5 olması halinde ilk olarak yine unvanlara göre aynı oranlarda alt gruplar oluşmuş, ancak unvan belirtmeyen grubu ücretlendirmeden çok memnun olanlar alt grubuna ayrılmamıştır. Ücretlendirmeden memnuniyet konusunda Akdeniz bölgesindekilerin grubunda dallanma ücretlendirmeden memnuniyete göre alt gruplara ayrılarak durdurulmuştur. İstanbul Anadolu bölgesinde ise dallanma hiç olmamıştır. Asistan, müdür-üzeri, satıcı ve uzman/şef grubunda ise özlük işlerinin önemi değil İK’nın şirkete sağladığı olumlu katkılara göre ayırım yapılmış ardından bölgelere göre alt gruplar belirlenmiştir. Motokurye/sürücü grubu ise ücretlendirmeden memnuniyete göre değil İK’nın yönetim ile personel iletişimini sağladığına değinerek işinden genel olarak memnuniyetini belirtenlere göre alt gruplar oluşturulmuştur.

İstanbul Avrupa bölgesinde 738 kişi yer almaktadır ve 175’i motokurye/sürücü ve uzman/şef pozisyonundadır. Bu gruptakiler içinde işe alınma konusunda çok memnun olan 21’i şirket tarafından sağlanan yan faydalardan da çok memnun olduğunu söylemiştir. 139’u müdür-üzeri ve unvan belirtmeyenlerin oluşturduğu gruptur ve 134’u genel anlamda işinden memnuniyetini 2 ile 5 aralığında cevaplayanlar ya da cevapsız bırakanlardır, bu grup üyeleri İK’nın yeni uygulamalar (S20) gerçekleştirmesi konusuna değinerek 2 ile 4 arasında cevap verenler çalışanlara karşı dürüst olduğu konusuna dikkat çekenlerdir. S20’ye çok katılıyorum cevabı veren 100 kişiden eğitim ve geliştirme çalışmaları konusunda memnun olanlar ile olmayanlar hemen hemen yarı yarıya alt gruplara ayrılmıştır. İstanbul Avrupa bölgesinde 424 kişi asistan ve satıcı pozisyonundadır ve işinden genel anlamda memnuniyetini “çok” olarak nitelendiren ve cevap vermeyenlerden oluşan alt grupta 217 kişi İK’nın personel tarafından saygı gördüğü konusuna (S15) değinmiştir. S15’e cevap vermeyen 28 kişiden 25’i eğitim ve geliştirme çalışmalarının kendisi için önemine 4 ile 5 aralığında cevap verenler ya da cevapsız bırakanların grubudur. Yine İK’nın hiç saygı görmediğini düşünen 16 kişiden 11’i İK’nın sorunlar hakkında bilgi sahibi olmasını 1 ile 4 aralığında cevaplayarak İK’nın sevilen bir departman olmadığını söylemiştir. İK’nın saygınlığı sorusuna çoğunluk 2 ile 5 aralığında

cevaplayanlardan 94'ü asistan ve 79'u satıcıdır. Buna göre asistan alt grubundakiler İK'nın düzenli ve titiz olduğuna değinerek 2 ile 5 aralığında cevap verenler ile cevapsızların birleşmesi ile oluşan grup üyeleri performans değerlendirme çalışmalarından hemen hemen yarı yarıya memnun ve memnun olmayan diye iki gruba ayrılmıştır. İşinden memnuniyetini 1 ile 4 aralığında cevaplayan 207 kişiden 108'i işe alımın kendisi için çok önemli olduğunu söyleyenler ya da cevapsızların oluşturduğu alt grubundadır. İşe alımın kendisi için önemini 1 ile 4 aralığında cevaplayan 99 kişi kariyer planlama çalışmalarından memnuniyete göre 4 alt gruba ayrılmıştır. Hiç memnun olmadığını söyleyen 24 kişiden eğitim ve geliştirmeden hiç veya memnun değilim diyen 17 kişiden İK'nın başarıları takdir edici olduğu konusunda hiç katılmıyorum cevabı verenler 5 kişidir; eğitim ve geliştirme çalışmalarından çok, orta ya da memnun olan 7 kişi İK'nın personelin saygı duyduğuna dikkat ettiği görülmüştür. Kariyer planlama konusunda "memnun değilim" diyen 35 kişi kurum içi iletişimden memnuniyetsiz olduğu ve bu grubun İK'nın dürüstlüğünü 2 ile 5 aralığında değerlendirdiği görülmüştür. Kariyer planlama çalışmalarından orta derece memnun olan ve cevapsız bırakanların grubunda 25 kişi işe alınma çalışmalarının kendisi için önemli olduğunu orta, iyi ve çok derecelerinde nitelendirenlerdir ve şirket tarafından sağlanan yan faydalara değinmişlerdir. Kariyer planlama konusunda çok ve iyi derece memnun olan 14 kişiden 8'i asistan ve 62'si satıcıdır. Asistan grubu işe alınma esnasındaki değerlendirmelerin önemine değinerek çoğunluk önemli olduğunu söylemiştir, bu kişiler şirketteki özlük işleri çalışmalarından memnun olup olmadıkları şeklindeki ayırimda çoğunluk yani 25 kişiden 24'ü 2 ile 5 aralığında cevap verenlerin ya da cevapsızların oluşturduğu alt gruptadır. İstanbul Anadolu bölgesi en çok dallanmanın olduğu bölge idi ancak terminal düğüm büyüklüğü 10, çocuk düğüm büyüklüğü 5 olduğunda ağacın alt dallanmalarının olmadığı görülmüştür. Yine unvanlara göre ayırım yapılarak asistan, satıcı grubundakiler işinden memnuniyetini 5 ile cevaplayanlar ve cevapsız bırakanlar İK'nın saygı gördüğü konusunda hiç katılmıyorum diyen grup 16 kişidir ve dallanma durmuştur. İK'nın saygı gören departman olduğuna cevap vermeyenler ise eğitim ve geliştirmenin kendisi için önemi sorusuna göre değil özlük işlerinden memnuniyetine göre alt gruplara ayrılmıştır.

## KAYNAKLAR

- Baron, S. ve Philips, D., (1994), “Attitude Survey Data Reduction Using CHAID: An Example in Shoppig Centre Market Research”, Journal of Marketing Management, 1994, 10:75-88.
- Doyle, P ve Fenwick,I.,(1975), “The Pitfalls of AID Analysis”, Journal of Marketing Research, November 1975,XII:408-413.
- Erbaş, S. ve Güneş, A., (1998), “CHAID Analizi”, İstatistik Konferansı, 26-27 Ekim 1997, Gazi Ün.Ankara.
- Hair,J.F. vd., (1995), Multivariate Data Analysis with Readings Prentice Hall, New Jersey.
- Hoare, R., (2004), “Using CHAID for Classification Problems”, New Zealand Statistical Association,2004, New Zealand.
- Kass, G.V.,(1980), “An Exploratory Technique for Investigating Large Quantities of Categorical Data”, Applied Statistic, 1980,29(2):119-127.
- Mano, C. ve Rasa, E.,(2002), “ A Discussion of Modelling Technique for Personal Lines Pricing”,Trans 27<sup>th</sup> ICA 2002, Brazil.
- Morgan, E. W. ve Paananen, O., (1972), “ An Analysis of Socioeconomic Change- A further Examination of Economic Base Theory”, Annals of Regional Science, June 1972,6(1):124-134.
- Orekici Temel, G., Çamdeviren, H. ve Akkuş, Z., (2005), “Sınıflama Ağaçları Yardımıyla Legs Syndrome Hastalarına Tanı Koyma”, İnönü Üniversitesi Tıp Fakültesi Dergisi, 12(2):111-117.
- Özmen, Ş., Yurtkoru, S. ve Sipahi, B., (2002), “Application of Datamining Tool for Classification of Organizational Change Expectation”, Yöneylem Araştırma ve Endüstri Mühendisliği XXIII, 3-5 Temmuz 2002, Yeditepe Üni. İstanbul.
- Staelin, R.,(1971), “ Another Look At AID”, Journal of Advertising Research, October 1971,XI(5):23-28.
- Yağız, Z.,(2003),CHAID Analizi, Yüksek Lisans Tezi, Gazi Üni. Fen Bilimleri Ens.

## INTERNET KAYNAKLARI

- [1][www.bilgiyonetimi.org/cm/pages/mkl\\_gos.php?nt=538](http://www.bilgiyonetimi.org/cm/pages/mkl_gos.php?nt=538).
- [2][www.consumer.gov.uk/homesafetynetwork/pdf/diyapp2](http://www.consumer.gov.uk/homesafetynetwork/pdf/diyapp2).
- [3][www.cs.uiuc.edu/class/fa05/cs412/chaps/6.pdf](http://www.cs.uiuc.edu/class/fa05/cs412/chaps/6.pdf).
- [4]<http://dollar.biz.uiowa.edu/~street/Trees.ppt>.
- [5][www.fatih.edu.tr/~nsolakoglu/courses/spring2005/ISLE502-DERS4.ppt](http://www.fatih.edu.tr/~nsolakoglu/courses/spring2005/ISLE502-DERS4.ppt).
- [6][www.hrs.co.nz/downloads/UsingCHAIDforclassificationproblems](http://www.hrs.co.nz/downloads/UsingCHAIDforclassificationproblems).
- [7]<http://webhost.ua.ac.be/eume/workshops/reallife/janssens.pdf>.

### Ek 1 Anket Soru Formu

Değerli Arkadaşımız,

Bu anketteki tüm sorular sizin görüşlerinizi de alarak, insan kaynakları departmanımızın çalışmalarını düzenlemek amacıyla hazırlanmıştır. Böylece şirket içinde birlikte düşünerek eksiklerin giderilmesi ve keyifli bir iş ortamı amaçlanmaktadır. İlginize teşekkür ederiz.

Saygılarımızla,  
İnsan Kaynakları Departmanı

**S1:** İşini daha iyi yapması için çalışanların eğitim ihtiyaçlarının belirlenmesi, ihtiyacı olan konularda eğitim verilmesi, eksiklerini tamamlayabilmesi için eğitim programlarına katılması konusunda ne düşünüyorsunuz ?

C1a: Bu konu benim için,

|                    |                 |                       |          |              |
|--------------------|-----------------|-----------------------|----------|--------------|
| 1.Hiç Önemli Değil | 2. Önemli Değil | 3. Ne Önemli Ne Değil | 4.Önemli | 5.Çok Önemli |
|--------------------|-----------------|-----------------------|----------|--------------|

C1b: Şirketimizdeki bu tür çalışmalardan,

|                      |                   |                           |            |                |
|----------------------|-------------------|---------------------------|------------|----------------|
| 1.Hiç Memnun Değilim | 2. Memnun Değilim | 3. Ne Memnunum Ne Değilim | 4.Memnunum | 5.Çok Memnunum |
|----------------------|-------------------|---------------------------|------------|----------------|

**S2:** Personel eksikliğinin hızla giderilmesi, işe uygun kişilerin bilgi ve becerilerinin geçerli testlerden geçilerek işe alınması konusunda ne düşünüyorsunuz?

C2a: Bu konu benim için,

|                    |                 |                       |          |              |
|--------------------|-----------------|-----------------------|----------|--------------|
| 1.Hiç Önemli Değil | 2. Önemli Değil | 3. Ne Önemli Ne Değil | 4.Önemli | 5.Çok Önemli |
|--------------------|-----------------|-----------------------|----------|--------------|

C2b: Şirketimizdeki bu tür çalışmalardan,

|                      |                   |                           |            |                |
|----------------------|-------------------|---------------------------|------------|----------------|
| 1.Hiç Memnun Değilim | 2. Memnun Değilim | 3. Ne Memnunum Ne Değilim | 4.Memnunum | 5.Çok Memnunum |
|----------------------|-------------------|---------------------------|------------|----------------|

**S3:** Başarılı ve başarısız personelin birbirinden ayrılması, başarılı sayılma kriterlerinin açık ve net olması, performans değerlendirmelerin tarafsız yapılması konusunda ne düşünüyorsunuz ?

C3a:Bu konu benim için,

|                    |                 |                       |          |              |
|--------------------|-----------------|-----------------------|----------|--------------|
| 1.Hiç Önemli Değil | 2. Önemli Değil | 3. Ne Önemli Ne Değil | 4.Önemli | 5.Çok Önemli |
|--------------------|-----------------|-----------------------|----------|--------------|

C3b: Şirketimizdeki bu tür çalışmalardan,

|                      |                   |                           |            |                |
|----------------------|-------------------|---------------------------|------------|----------------|
| 1.Hiç Memnun Değilim | 2. Memnun Değilim | 3. Ne Memnunum Ne Değilim | 4.Memnunum | 5.Çok Memnunum |
|----------------------|-------------------|---------------------------|------------|----------------|

**S4:** Ücret dengesizliklerinin olmaması, personelin performansına göre ücret belirlenmesi konusunda ne düşünüyorsunuz ?

C4a: Bu konu benim için,

|                    |                 |                       |          |              |
|--------------------|-----------------|-----------------------|----------|--------------|
| 1.Hiç Önemli Değil | 2. Önemli Değil | 3. Ne Önemli Ne Değil | 4.Önemli | 5.Çok Önemli |
|--------------------|-----------------|-----------------------|----------|--------------|

C4b: Şirketimizdeki bu tür çalışmalardan,

|                      |                  |                           |            |                |
|----------------------|------------------|---------------------------|------------|----------------|
| 1.Hiç Memnun Değilim | 2.Memnun Değilim | 3. Ne Memnunum Ne Değilim | 4.Memnunum | 5.Çok Memnunum |
|----------------------|------------------|---------------------------|------------|----------------|

**S5:** Şirket tarafından sunulan servis, yemek, sosyal olanaklar gibi ücret dışındaki faydalar konusunda ne düşünüyorsunuz ?

C5a: Bu konu benim için,

|                    |                 |                       |          |              |
|--------------------|-----------------|-----------------------|----------|--------------|
| 1.Hiç Önemli Değil | 2. Önemli Değil | 3. Ne Önemli Ne Değil | 4.Önemli | 5.Çok Önemli |
|--------------------|-----------------|-----------------------|----------|--------------|

C5b: Şirketimizdeki bu tür çalışmalardan,

|                      |                  |                           |            |                |
|----------------------|------------------|---------------------------|------------|----------------|
| 1.Hiç Memnun Değilim | 2.Memnun Değilim | 3. Ne Memnunum Ne Değilim | 4.Memnunum | 5.Çok Memnunum |
|----------------------|------------------|---------------------------|------------|----------------|

**S6:** Mevcut pozisyondan yükselme şansının olması, pozisyonda yükselme şartların açık ve net olması, personelin üst pozisyonlar için yetiştirilmesi, açık pozisyonlarda öncelikle şirket içinden personelin değerlendirilmesi konusunda ne düşünüyorsunuz ?

C6a: Bu konu benim için,

|                    |                 |                       |          |              |
|--------------------|-----------------|-----------------------|----------|--------------|
| 1.Hiç Önemli Değil | 2. Önemli Değil | 3. Ne Önemli Ne Değil | 4.Önemli | 5.Çok Önemli |
|--------------------|-----------------|-----------------------|----------|--------------|

C6b: Şirketimizdeki bu tür çalışmalardan,

|                      |                   |                           |            |                |
|----------------------|-------------------|---------------------------|------------|----------------|
| 1.Hiç Memnun Değilim | 2. Memnun Değilim | 3. Ne Memnunum Ne Değilim | 4.Memnunum | 5.Çok Memnunum |
|----------------------|-------------------|---------------------------|------------|----------------|

**S7:** Nema, yıllık izin gibi kayıtların düzenli takip edilmesi ve talep edilmesi durumunda bilginin ve belgelerin sunulması, SSK ile ilgili işlemlerin ve bordroların açık ve net yapılması konusunda ne düşünüyorsunuz ?

C7a: Bu konu benim için,

|                       |                 |                          |          |                 |
|-----------------------|-----------------|--------------------------|----------|-----------------|
| 1.Hiç Önemli<br>Değil | 2. Önemli Değil | 3. Ne Önemli Ne<br>Değil | 4.Önemli | 5.Çok<br>Önemli |
|-----------------------|-----------------|--------------------------|----------|-----------------|

C7b: Şirketimizdeki bu tür çalışmalardan,

|                         |                      |                              |            |                   |
|-------------------------|----------------------|------------------------------|------------|-------------------|
| 1.Hiç Memnun<br>Değilim | 2. Memnun<br>Değilim | 3. Ne Memnunum<br>Ne Değilim | 4.Memnunum | 5.Çok<br>Memnunum |
|-------------------------|----------------------|------------------------------|------------|-------------------|

**S8:** Personelin kültürel ve sosyal alanlarda desteklenmesi için yemekli geceler, tiyatro, gezi gibi aktivelerin yapılması, şirket ile ilgili değişiklikler hakkında personelin bilgilendirilmesi konusunda ne düşünüyorsunuz ?

C8a: Bu konu benim için,

|                       |                 |                          |          |                 |
|-----------------------|-----------------|--------------------------|----------|-----------------|
| 1.Hiç Önemli<br>Değil | 2. Önemli Değil | 3. Ne Önemli Ne<br>Değil | 4.Önemli | 5.Çok<br>Önemli |
|-----------------------|-----------------|--------------------------|----------|-----------------|

C8b: Şirketimizdeki bu tür çalışmalardan,

|                         |                      |                              |            |                   |
|-------------------------|----------------------|------------------------------|------------|-------------------|
| 1.Hiç Memnun<br>Değilim | 2. Memnun<br>Değilim | 3. Ne Memnunum<br>Ne Değilim | 4.Memnunum | 5.Çok<br>Memnunum |
|-------------------------|----------------------|------------------------------|------------|-------------------|

Aşağıdaki sorularda insan kaynakları departmanımızı değerlendirmeniz istenmektedir.

Değerlendirme şu ölçeğe göre yapılmalıdır.

|                       |                 |                                      |               |                            |
|-----------------------|-----------------|--------------------------------------|---------------|----------------------------|
| 1.Hiç<br>Katılmıyorum | 2. Katılmıyorum | 3. Ne Katılıyorum<br>Ne Katılmıyorum | 4.Katılıyorum | 5.Bütünüyle<br>Katılıyorum |
|-----------------------|-----------------|--------------------------------------|---------------|----------------------------|

**S9:** Gerekli bir departmandır ( )

**S10:** Personele yaklaşımı dürüsttür. ( )

**S11:** Şirket içi iletişimde başarılıdır. ( )

**S12:** Olanaklar konusunda personele bilgi verir. ( )

**S13:** Şirketteki iş tanımları hakkında bilgisi vardır. ( )

**S14:** Kaliteli hizmet sunar. ( )

**S15:** Personelden saygı görür. ( )

**S16:** Şirketin başarılı olmasına katkıda bulunur. ( )

**S17:** Yönetim ile personel isteklerini dengeler. ( )

**S18:** Kariyer hedeflerimizle ilgilidir. ( )

**S19:** Çalışmalarında düzenli ve titizdir. ( )

**S20:** Şirkette yeni uygulamalar gerçekleştirir. ( )

**S21:** Şirket içinde söz sahibidir. ( )

**S22:** Personel tarafından sevilen bir departmandır. ( )

**S23:** Personele karşı takdir edicidir. ( )

**S24:** Sorunlarımız hakkında bilgi sahibidir. ( )

**S25:** Çözüm üretme kabiliyeti vardır. ( )

**S26:** Beklentilerimize duyarlıdır. ( )

**S27:** Yönetim ile iletişimi sağlar. ( )

**S28:** Yeni çalışmalarda öncüdür. ( )

**S29:** Genel anlamda işimden memnunum ( )

**S30:** Şirketteki faaliyet alanınız hangisidir ?

Holding Merkezi ( )

Bölge ( )

**S31:** Şirketteki unvanınız nedir ?

Müdür ve üzeri ( )

Uzman / Şef ( )

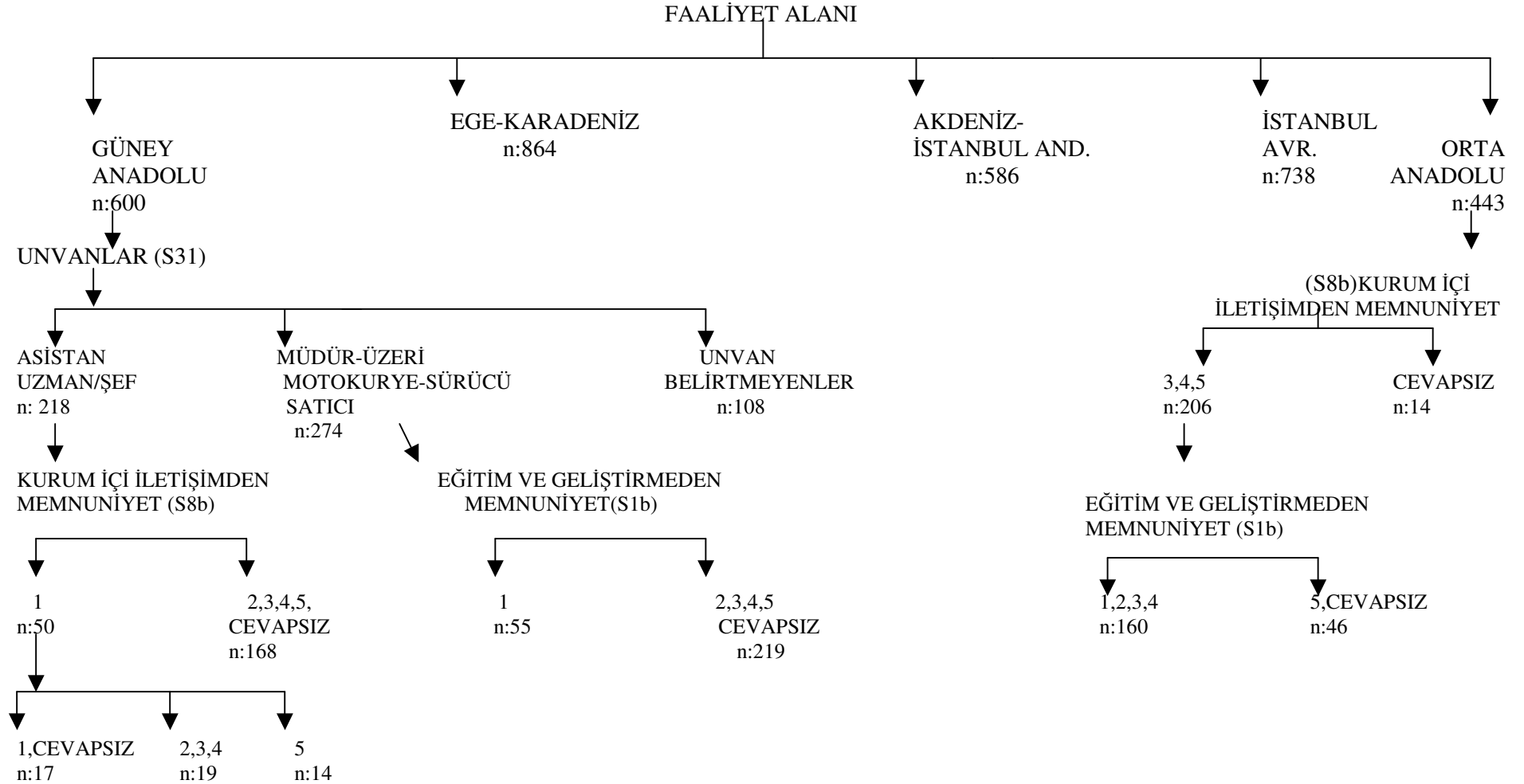
Satıcı ( )

Motokurye/ Sürücü ( )

Asistan / Görevli ( )

Cevaplarınızı içtenlikle verdiğiniz için teşekkür ederiz.

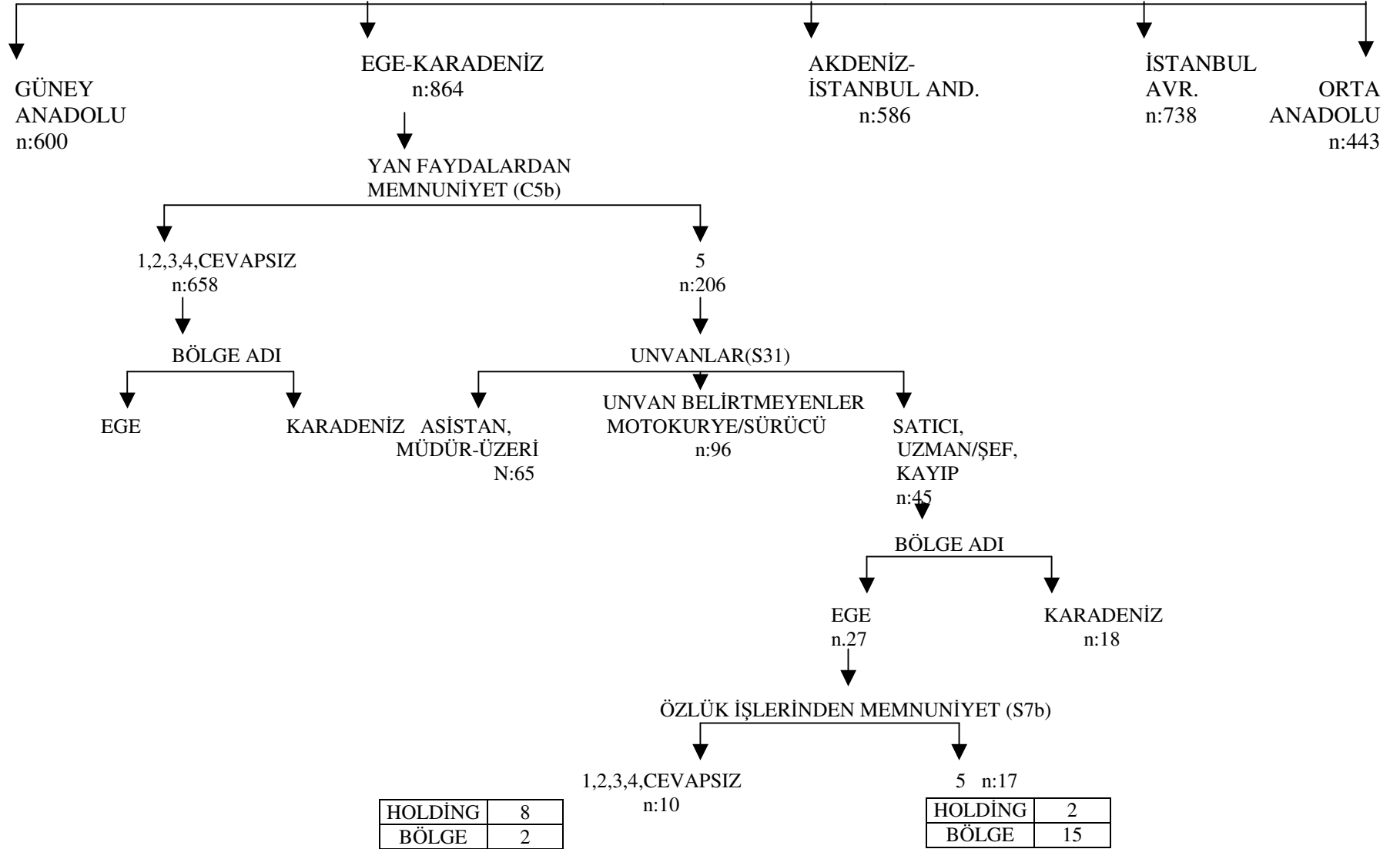
**EK2 Eğitim, İşe Alma, Performans Değerlendirme, Ücretlendirme, Yan Faydalar, Kariyer Planlanma, Özlük İşleri Ve Bordrolama, Kurum İçi İletişim, Bölge Adı ve Unvanlar ile Faaliyet Alanına Göre Ağaç Diyagramı**



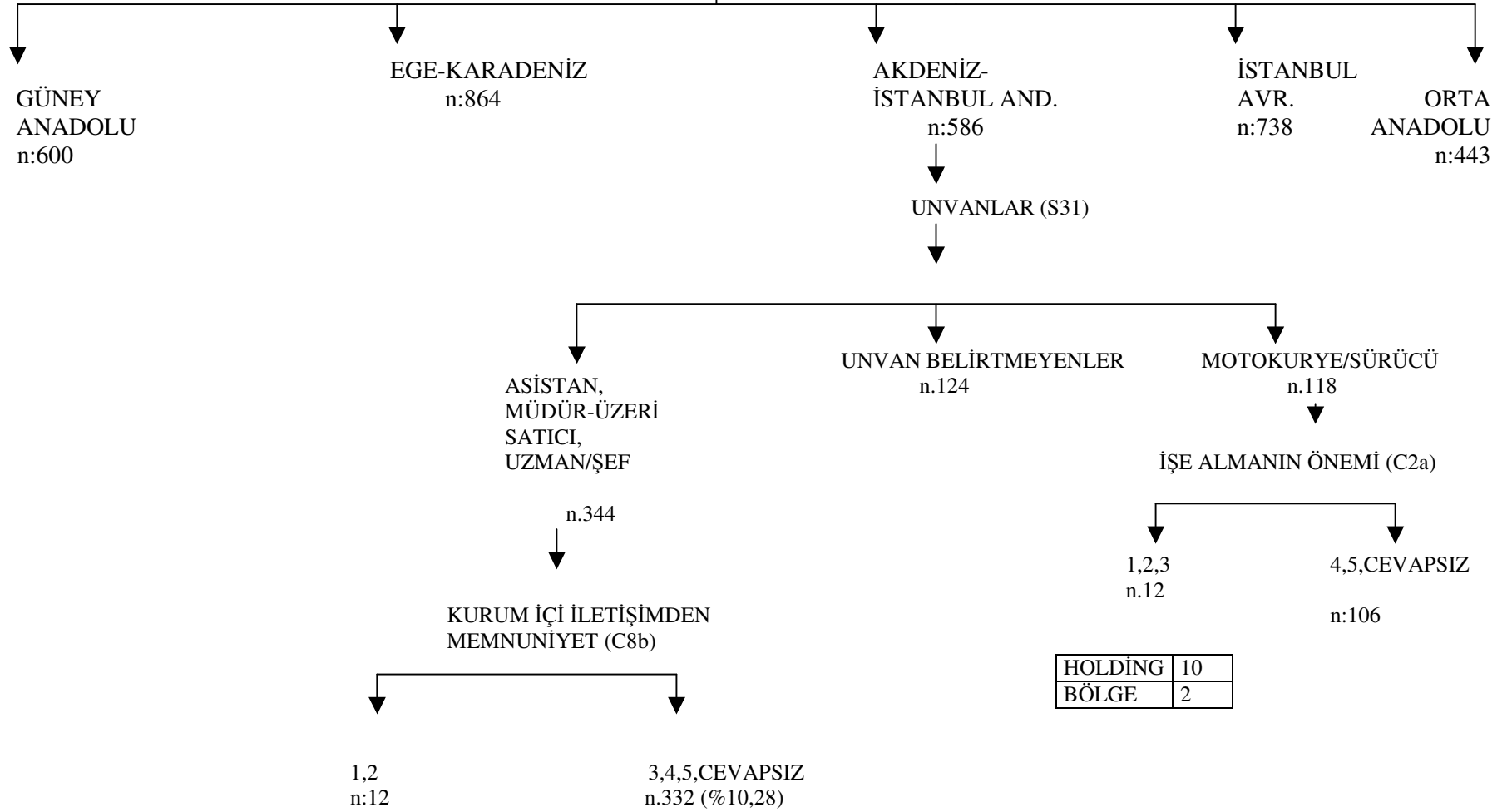


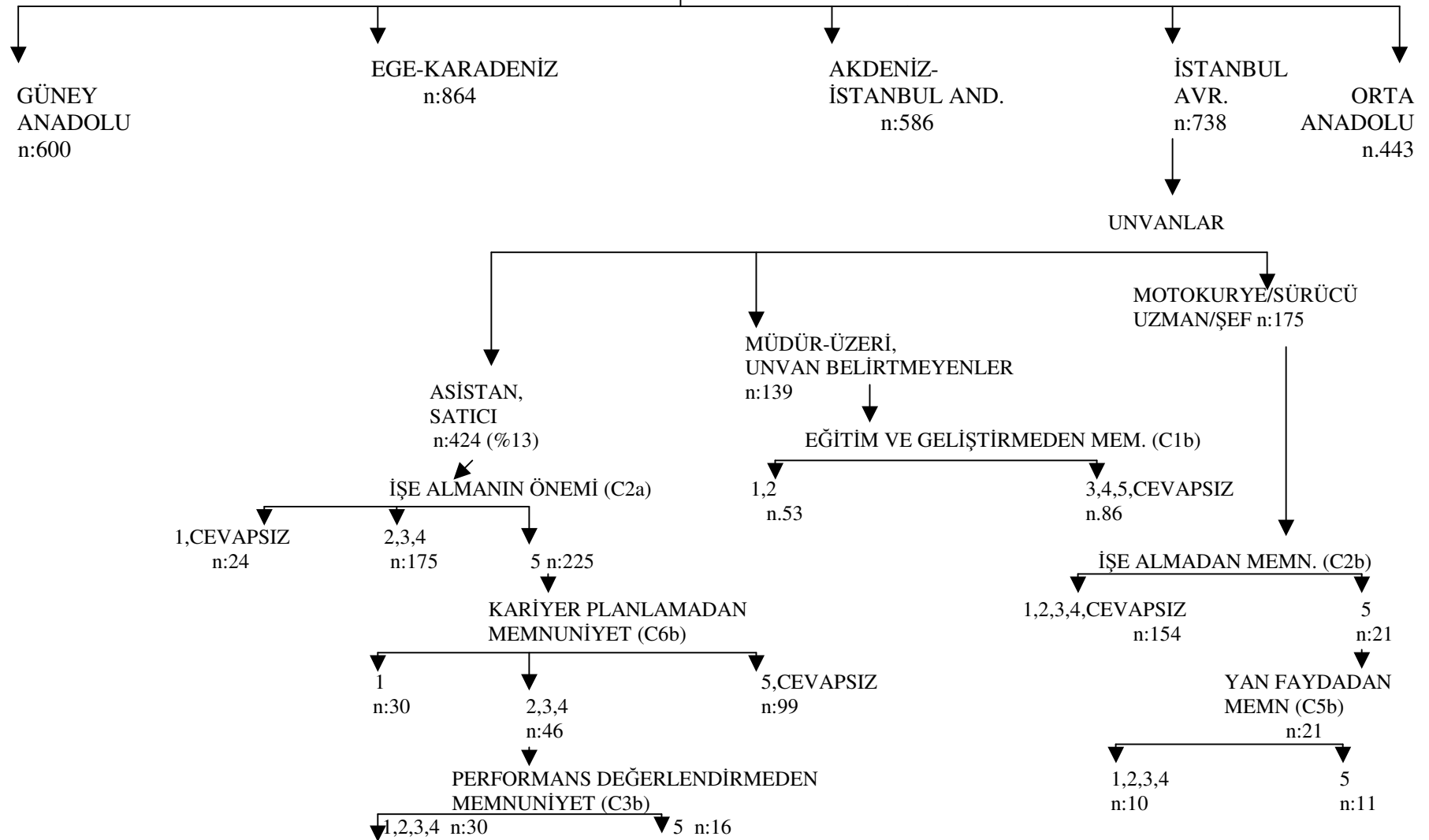
## Ek 2 (devam)

## FAALİYET ALANI(S30)

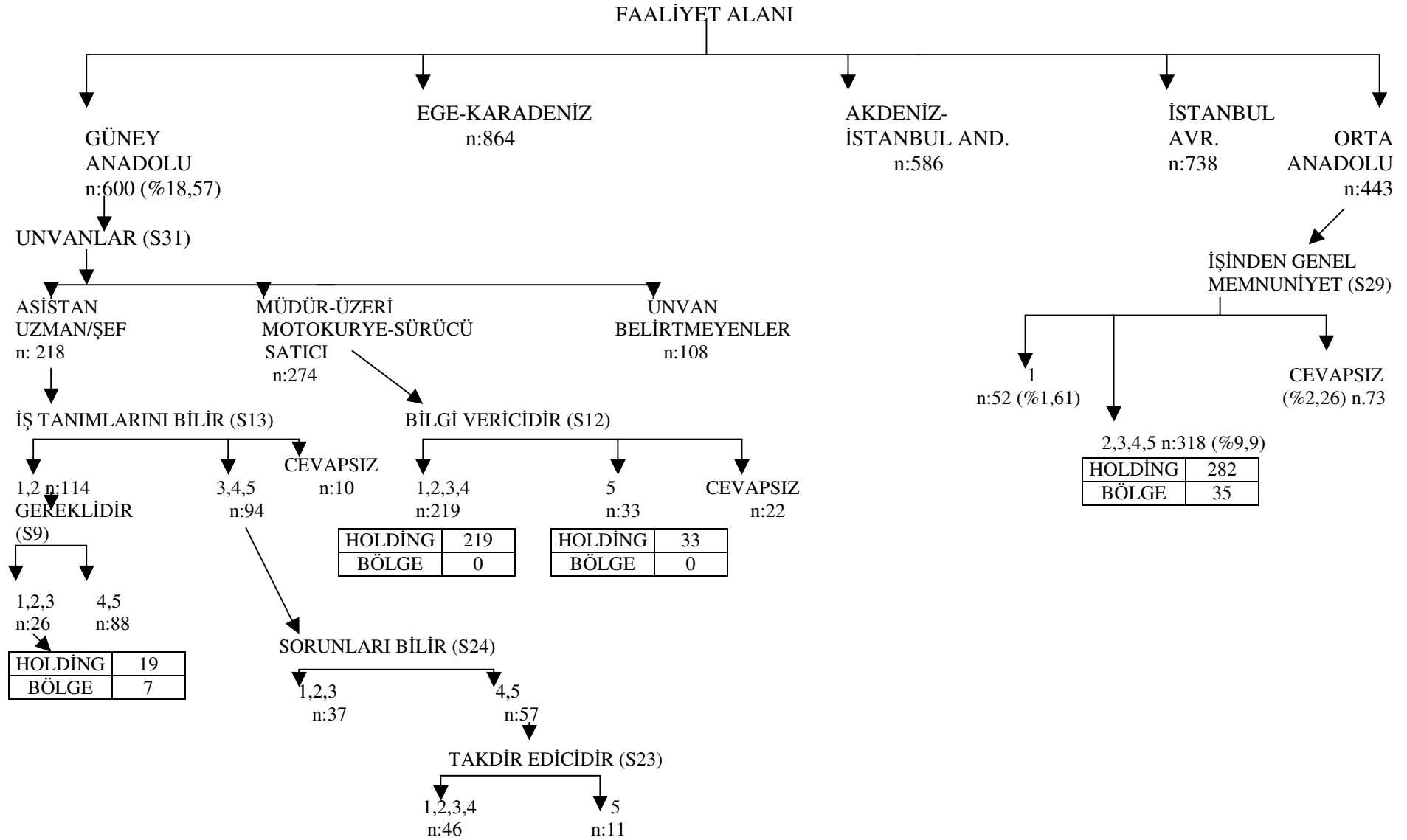


**Ek 2 (devam)**  
FAALİYET ALANI(S30)



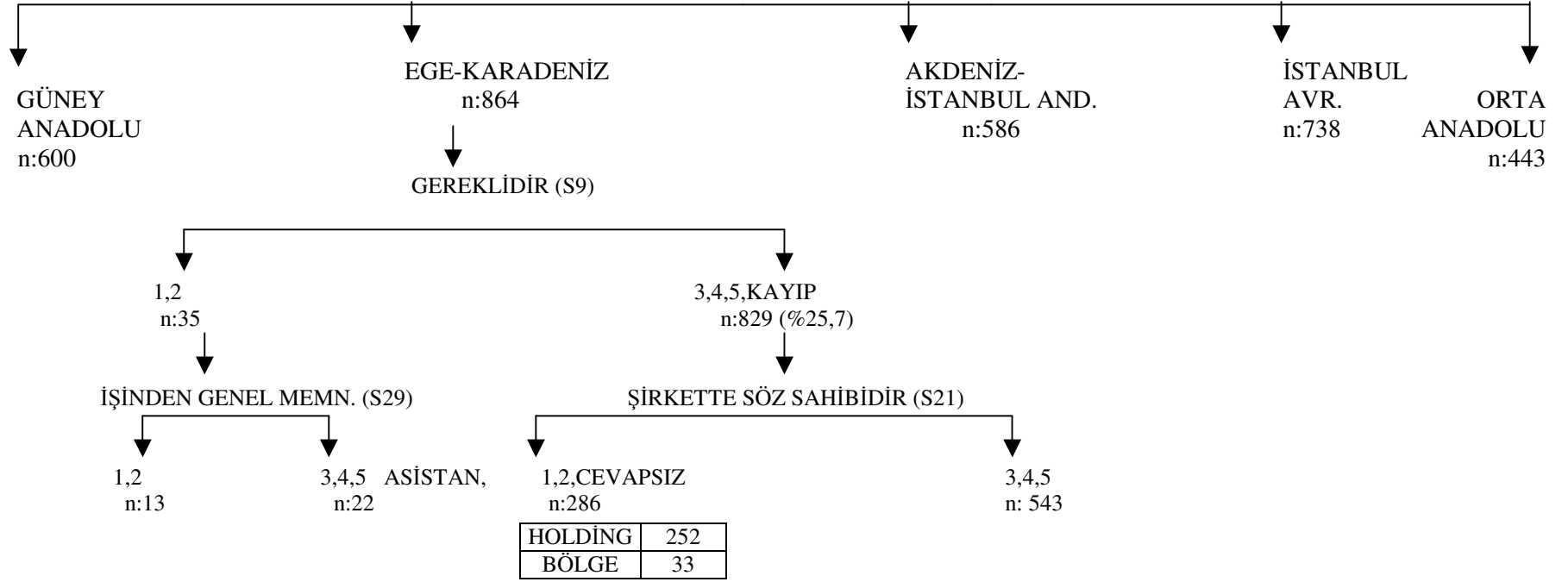


### EK3 İnsan Kaynakları ile İlgili Görüşler, Bölge Adı ve Unvanlar ile Faaliyet Alanına Göre Ağaç Diyagramı



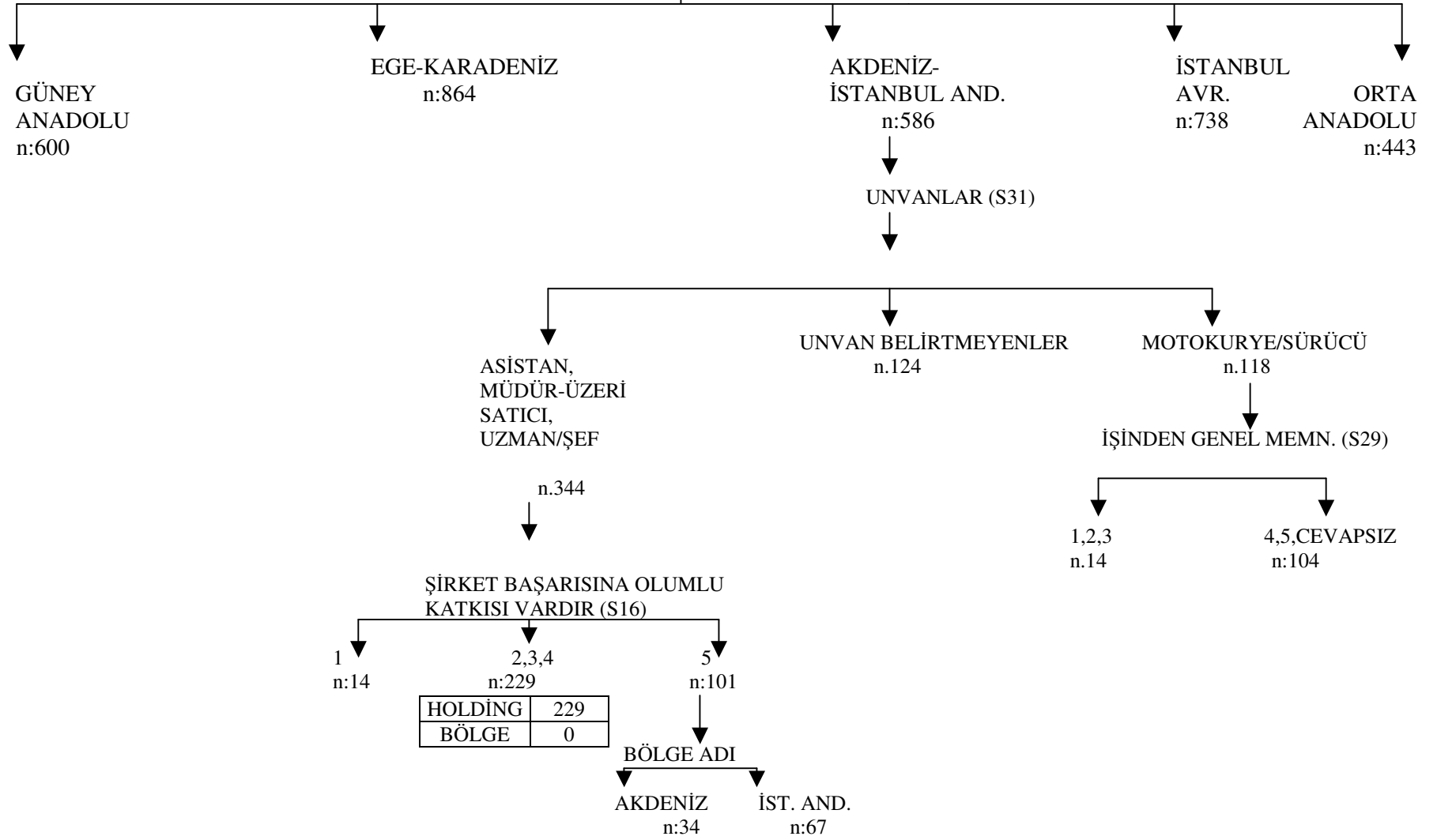
**Ek 3 (devam)**

FAALİYET ALANI(S30)



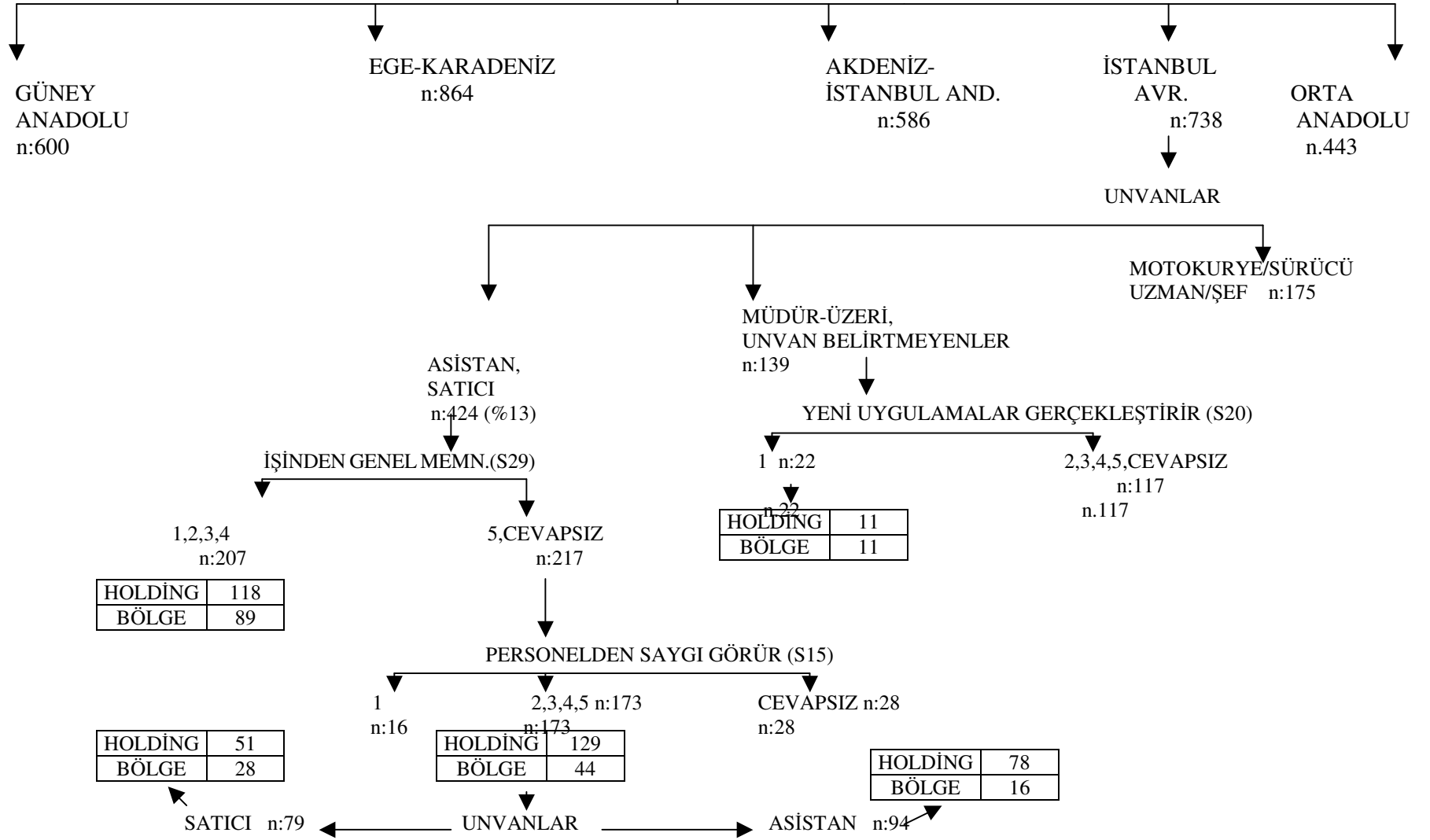
**Ek 3 (devam)**

FAALİYET ALANI(S30)

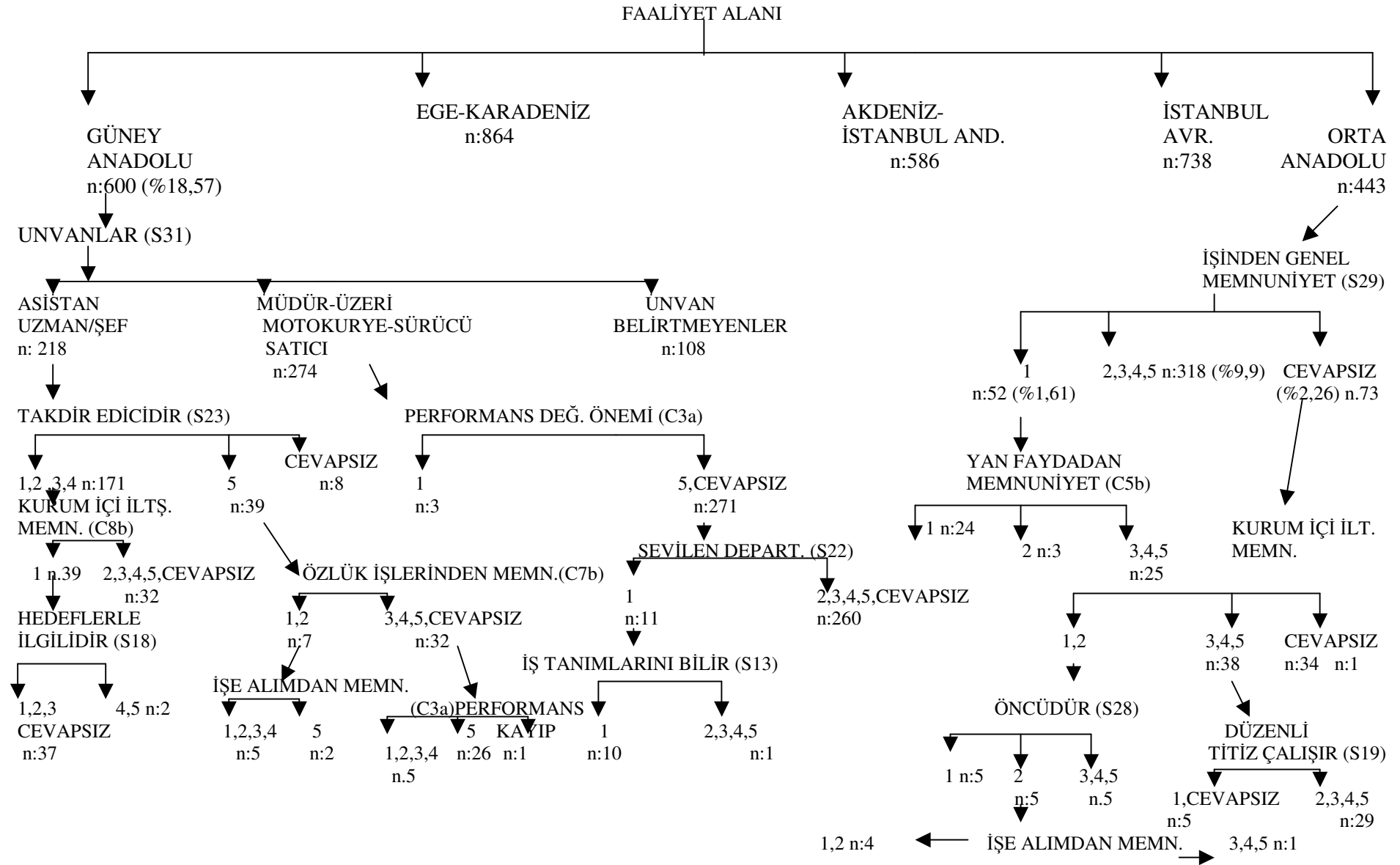


## Ek 3 (devam)

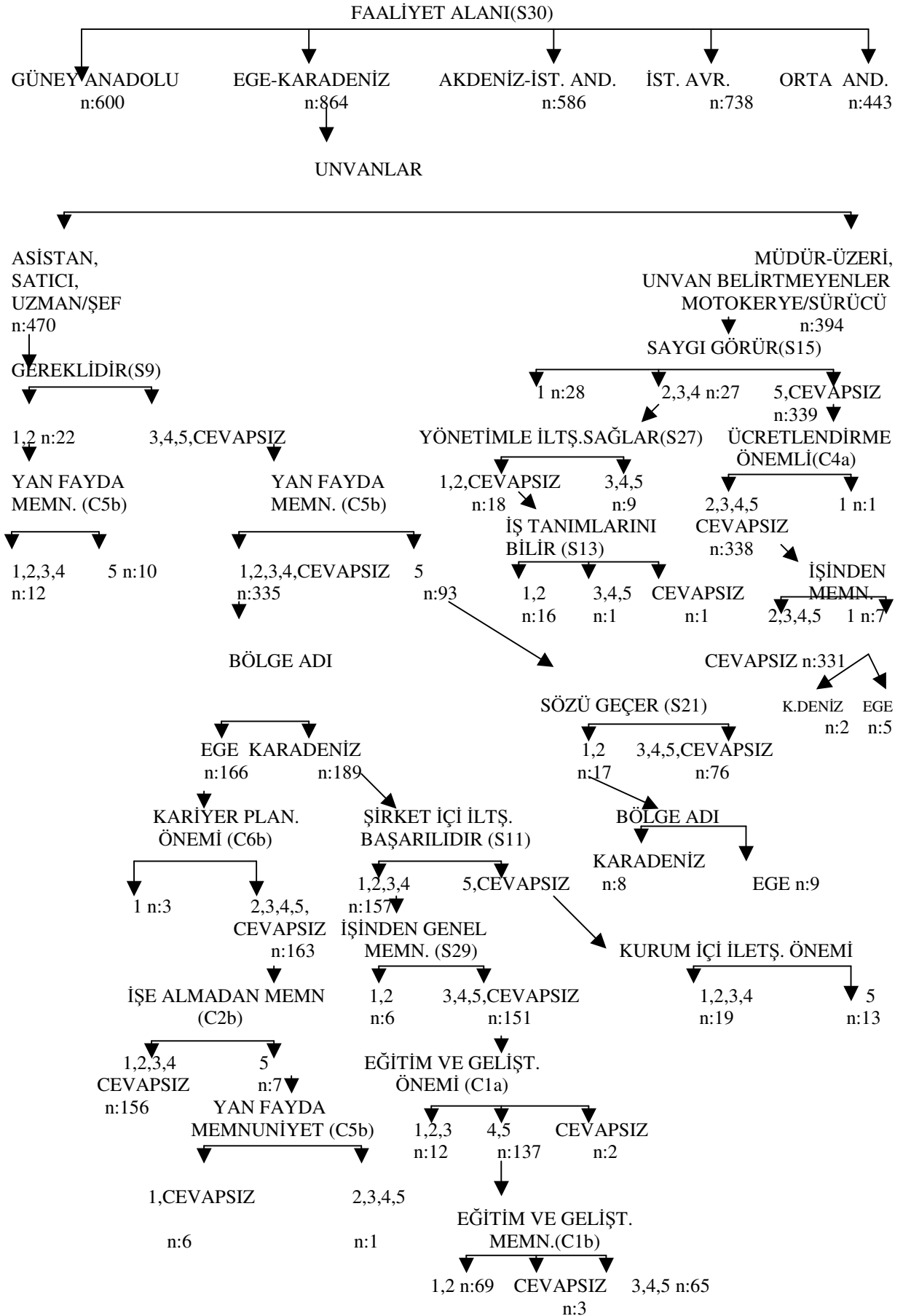
FAALİYET ALANI(S30)

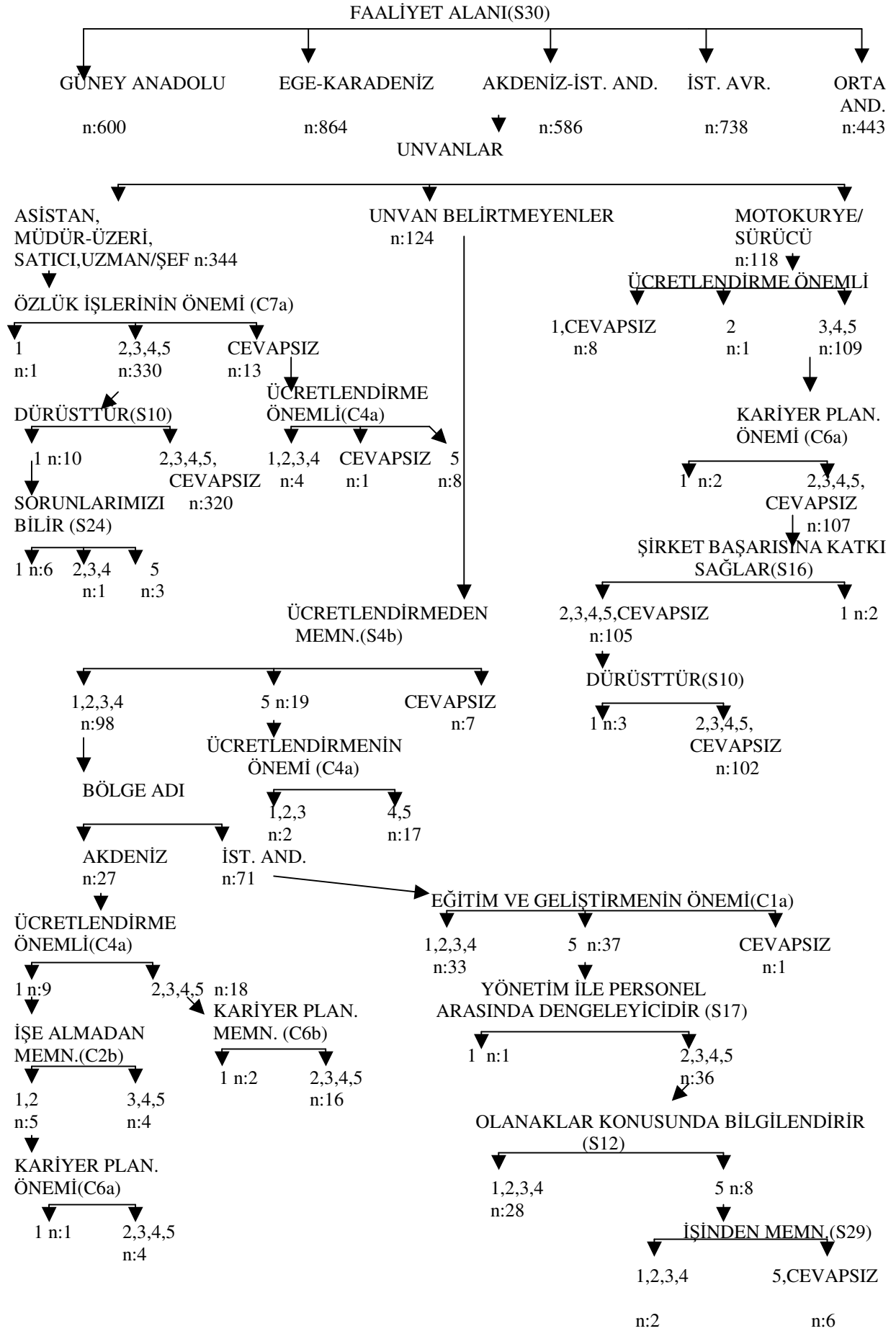


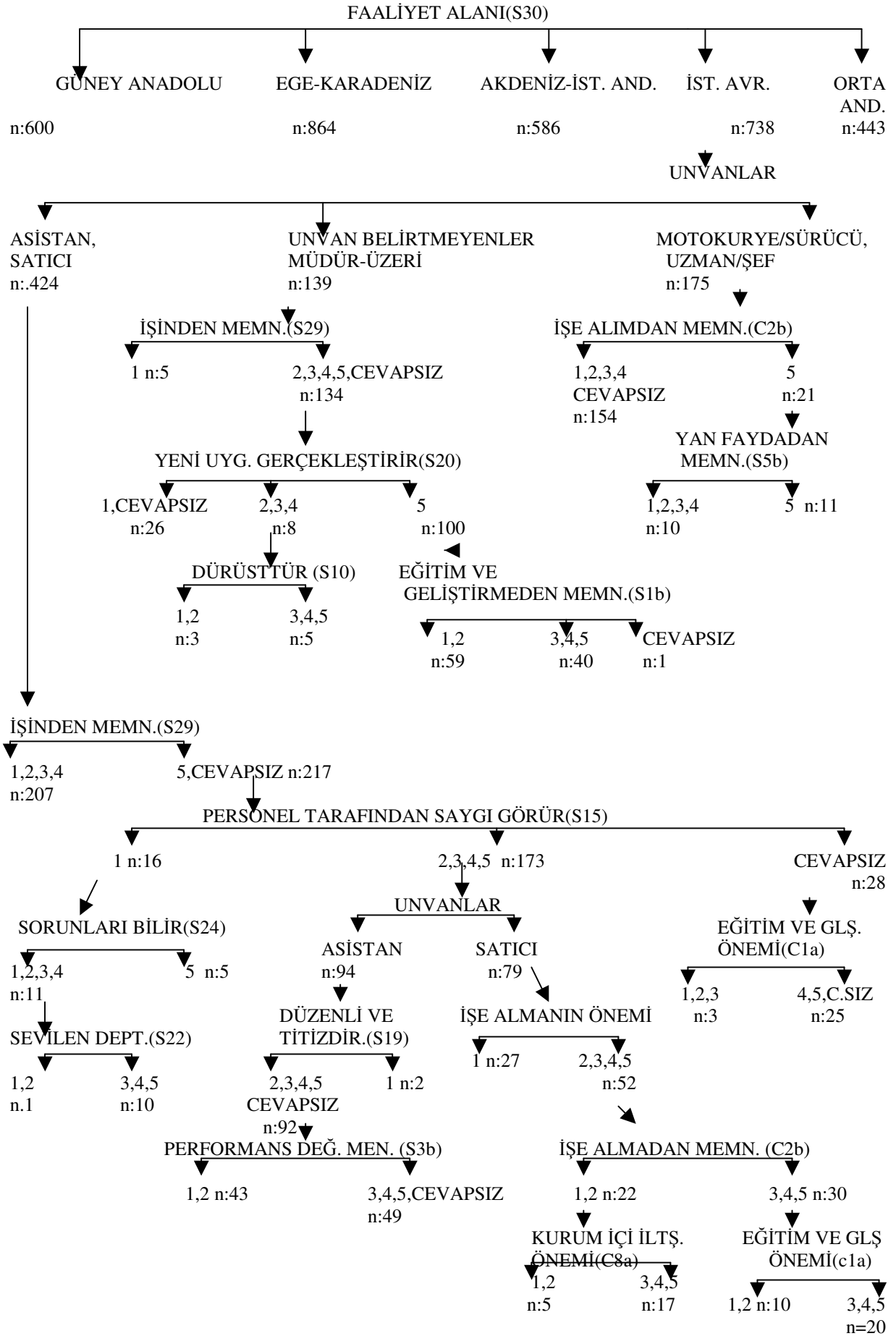
# EK4 Bütün Değişkenler ile Faaliyet Alanına Göre Ağaç Diyagramı

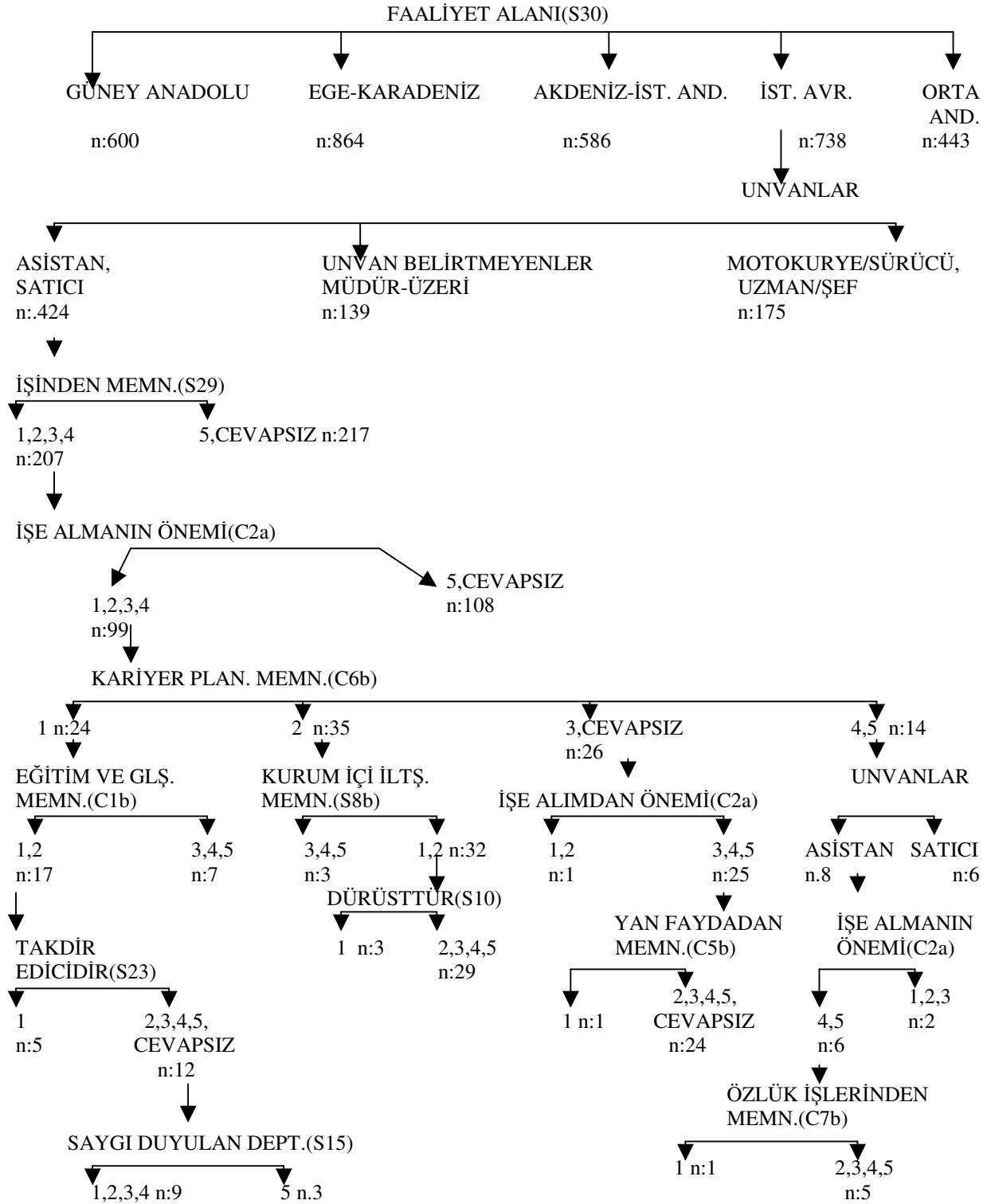




**Ek 4 (devam)**

**Ek 4 (devam)**

**Ek 4 (devam)**

**Ek 4 (devam)**

## ÖZGEÇMİŞ

|               |            |                                                                                                                 |
|---------------|------------|-----------------------------------------------------------------------------------------------------------------|
| Doğum Tarihi  | 29.05.1979 |                                                                                                                 |
| Doğum Yeri    | Merzifon   |                                                                                                                 |
| Lise          | 1994-1997  | Merzifon Anadolu Lisesi                                                                                         |
| Lisans        | 1998-2002  | Yıldız Teknik Üniversitesi Fen-Edebiyat Fak.<br>İstatistik Bölümü                                               |
| Yüksek Lisans | 2003-2006  | Yıldız Teknik Üniversitesi Fen Bilimleri Ens.<br>İstatistik Anabilim Dalı<br>İstatistik Yüksek Lisans Öğrt. Prg |

## Çalıştığı kurumlar

|                 |                                                                               |
|-----------------|-------------------------------------------------------------------------------|
| 04/2003-06/2003 | Devlet İstatistik Ens. İst. Bölge Md.                                         |
| 01/2004-04/2006 | İtalyan Dış Ticaret Merkezi                                                   |
| 07/2006- Devam  | Fintek Finans Bilgi Teknolojileri A.Ş<br>(Ziraat Bankası Bilgi İşlem Merkezi) |